Macroeconomic Uncertainty Prices when Beliefs are Tenuous

Lars Peter Hansen*,[a] and Thomas J.Sargent†,[b]

[a]University of Chicago, United States
[b]New York University, United States

March 23, 2020

Abstract

Investors face uncertainty *over* models when they do not know which member of a set of well-defined "structured models" is best. They face uncertainty *about* models when they suspect that all of the structured models might be misspecified. We refer to worries about the first type of ignorance as *ambiguity concerns* and worries about the second type as *misspecification concerns*. These two types of ignorance about probability distributions of risks add what we call uncertainty components to equilibrium prices of those risks. A quantitative example highlights a representative investor's uncertainties about the size and persistence of macroeconomic growth rates. Our model of preferences under concerns about model ambiguity and misspecification puts nonlinearities into marginal valuations that induce time variations in market prices of uncertainty. These reflect the representative investor's fears of high persistence of low growth rate states and low persistence of high growth rate states.

Keywords— Risk, robustness, model uncertainty, asset prices, relative and Chernoff entropy

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

*E-mail: lhansen@uchicago.edu, corresponding author.
†E-mail: thomas.sargent@nyu.edu.

1 Introduction

This paper describes prices of macroeconomic uncertainty that depend on how investors evaluate consequences of alternative specifications of state dynamics. We construct a quantitative example in which a representative investor who is uncertain about prospective macroeconomic growth rates impersonates "the market." Adverse consequences for discounted expected utilities make the representative investor fear high persistence of macroeconomic growth rates in times of weak growth and low persistence in times of strong growth, fears that induce variations in values of assets.

Our representative investor has a family of structured models with possibly time-varying parameters that we represent with a recursive structure for continuous time models with Brownian motion information flows suggested by Chen and Epstein (2002). We say that the investor experiences *model ambiguity* concerns because he does not know which model is best. Since the investor distrusts his structured models, he also cares about all unstructured models that reside within a statistical neighborhood of the set of structured models.[1] We formalize the investor's concerns about *model misspecification* by entertaining unstructured statistical models with similar observable implications and represent them with the preferences proposed by Hansen and Sargent (2020), a continuous-time version of the dynamic variational preferences of Maccheroni et al. (2006).[2]

Shadow prices that reflect aspects of model specifications that most concern the representative investor equal uncertainty prices that clear competitive securities markets. Multiplying an endogenously determined vector of worst-case drift distortions by minus one gives a vector of local prices that are increments to expected returns associated with exposures to alternative shocks over an instant of time that compensate the representative investor for bearing model uncertainty.[3] A representative investor's concerns about the persistence of macroeconomic growth rates make uncertainty prices vary over time because they depend on the state of the economy. These findings extend earlier quantitative results that had indicated how investors' responses to modest amounts of model ambiguity can replace the implausibly large risk aversions during economic downturns that are required to explain

[1]By "structured" we mean more or less tightly parameterized statistical models. Thus, "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]Hansen and Sargent (2020) extends models of Hansen and Sargent (2001) and Hansen et al. (2006) that surround a single structured baseline probability model with an infinite dimensional family of difficult-to-discriminate unstructured models.

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

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. We exploit the technical result that positive martingales with means equal to unity represent likelihood ratios that alter baseline probabilities. To express his model *ambiguity*, a representative investor uses a restricted set of such martingales that describe what we call structured models. To express his suspicion that all of his structured models are *misspecified*, a representative investor uses a substantially larger family of positive, mean one, martingales to describe what we call unstructured models. Section 3 describes a statistical measure of discrepancy between martingales called discounted relative entropy and derives a counterpart to it for the divergence between a given probability and a set of probabilities associated with the structured models. When used in conjunction with the set of unstructured models, this divergence concept provides a convenient way for a decision maker to explore the consequences of potential misspecifications of the structured models. In section 4, we apply this set-based divergence to positive martingale to formulate a robust decision problem that incorporates both model ambiguity and model misspecification concerns.

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 discrepancies 1) to assess plausibility of worst-case models as recommended by 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 competitive equilibrium quantities and prices for an economy with a representative robust investor. Then it links worst-case probabilities that emerge from a robust planning problem to uncertainty compensations that the representative investor receives in competitive equilibrium. Section 8 offers concluding remarks.

2 Martingales and probabilities

This section describes convenient mathematical representations of positive martingales that alter a baseline probability model. For the decision makers within our economic model, these martingales are Radon-Nikodym derivatives, i.e., likelihood ratios, between two measures, just as they are for statisticians and econometricians. Starting from a decision maker's baseline probability measure, we use martingales to represent probabilities that are on a decision maker's radar as plausible alternative models, and we use them to explore the consequences of misspecification. In section 3, we formalize these two uses of martingales and the distinct roles that they play in confronting uncertainty.

For concreteness, we use the following *baseline* model of a stochastic process $Z \doteq \{Z_t : t \geqslant 0\}$ that governs the exogenous dynamics.[4]

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

where W is a multivariate standard Brownian motion.[5] We also study endogenous state dynamics that can be altered by the actions of a fictitious planner. With this in mind, 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, who in this paper is an investor, entertains alternative models that are represented as likelihood ratios, which in our setting are positive martingales having unit expectations. Within this continuous-time Brownian information environment, these martingales have a well known and convenient characterization, courtesy of the Girsonov Theorem and related results from probability theory. Using these, we can describe a likelihood ratio such as M^U with its evolution with respect to a 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.

[4]We let Z denote a stochastic process, Z_t the process at time t, and z a realized value of the process.

[5]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.

Imposing the initial condition $M_0^U = 1$, we express the solution of stochastic differential equation (2) as a so-called 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.[6]

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 $\mathcal{F} = \{\mathcal{F}_t : t \geqslant 0\}$.*

In what follows, we use the processes U to represent alternative martingales of interest. We describe probabilities implicitly by delineating the family of conditional expectations associated with each such U process, namely,

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

for any $t \geqslant 0$ and any bounded $\mathcal{F}_t$-measurable random variable B_t. This representation uses the positive random variable M_t^U as a Radon-Nikodym derivative for the date t conditional expectation operator $E^U\left(\,\cdot\,\mid \mathcal{F}_0\right)$. The martingale property for M^U assures that the Law of Iterated Expectations applies to the constructed probability measures. In what follows, we will refer to this probability measure as being affiliated with the martingale M^U.

Under baseline model (1), W is a standard Brownian motion; but under the alternative U model, it follows from the Girsanov Theorem that 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. In light of (6), we can write model (1) as:

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

While (3) expresses the evolution of $\log M^U$ in terms of increment dW, the evolution in

[6] 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.

terms of dW^U is

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

An important property for us is that the drift of the log-likelihood under this alternative measure is the quadratic term: $|U_t|^2/2$. Under the M^U implied probability, for each $t \geqslant 0$, the rate at which the log-likelihood grows locally is the squared norm of the drift U_t.

In summary, the Brownian information structure leads us to explore probabilities that include drift processes U in the specification of W. Our investor is particularly interested in U processes that represent history-dependent changes in the local mean of the Brownian increments.[7]

3 Statistical discrepancies

As an example of the decision theory developed in Hansen and Sargent (2020), an investor in our model evaluates decisions in light of a family of what we call *structured probability models*. Each such model is represented by a process S used to construct a martingale M^S. An investor also acknowledges that each structured model could be misspecified by introducing statistically nearby *unstructured models* that also concern her. In this section, we provide convenient ways to represent these two components of uncertainty that confront decision makers in our model.

3.1 Relative entropy divergence

To model formally how the investor acknowledges misspecification of each structured probability model, we introduce a relative entropy divergence that measures the discrepancy between a structured probability affiliated with a martingale, M^S, and another probability affiliated with a martingale M^U. Specifically, for horizon t the divergence is

$$E\left[M_t^U \left(\log M_t^U - \log M_t^S\right) \mid \mathcal{F}_0\right] \geqslant 0,$$

which is the expected log likelihood ratio computed using the probability measure that is affiliated with martingale M^U. Using a counterpart to formula (7), from formulas (3) and

[7]Here we maintain absolute continuity in the probability measures over finite time intervals. We discuss possible extensions later.

(6) it follows that

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

From this evolution equation, it follows that

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

which is one half of an average norm squared deviation between drift processes U and S, where the average is computed under the probability affiliated with the martingale M^U. This illustrates a simplification that occurs in our continuous-time formulation. To represent model specification concerns, we focus on conditional mean distortions in Brownian increments and measure these distortions using a weighted integral of $|U - S|^2$.

We entertain potential model misspecification using relative entropy as in the robust control theory contributions of Jacobson (1973), Whittle (1981), James (1992), Hansen and Sargent (2001) and many others. The decision-maker/investor in that literature uses relative entropy to penalize probability deviations from a *single* baseline probability measure. Here we extend this approach by entertaining a *set* of models of particular interest to the decision maker. We will elaborate on this point later. While expanding the collection of structured models relative to that earlier work, the approach here retains relative entropy penalties as a way to represent concerns that structured probability models are misspecified.

We study decision problems with infinite horizon objectives. We follow a standard practice in dynamic programming by assuming a time-discounted objective function. This is convenient analytically and has the material benefit of assuring dynamic consistency of preferences. To preserve these features under potential misspecifications of the decision maker's set of structured models, we also introduce discounting into the relative entropy that we shall penalize:

$$
\Delta\left(M^U; M^S \mid \mathcal{F}_0\right) = \frac{\delta}{2} E\left(\int_0^\infty \exp(-\delta\tau) M_\tau^U |U_\tau - S_\tau|^2 d\tau \mid \mathcal{F}_0\right)
$$

as in Hansen and Sargent (1995), Hansen and Sargent (2001), and Hansen et al. (2006). The scaling by δ makes this an exponentially weighted average of expectations of $\frac{1}{2}|U_\tau - S_\tau|^2$

over time.

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

In contrast to the earlier literature on robust control cited above, we start from a closed (in the sense of the relative entropy divergence) convex set of *structured* models that we represent as martingales $M^S \in \mathcal{M}^o$. Structured models in $\mathcal{M}^o$ are well articulated alternative specifications that particularly interest 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

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

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 know utility consequences of these nearby models too. The scaling parameter θ measures how an expected utility maximizing decision maker penalizes an expected utility minimizing agent for distorting probabilities relative to models in $\mathcal{M}^o$.

Although, unlike us, Chen and Epstein (2002) do not explore potential model specification using a likelihood-based discrepancy measure, we begin by following their lead when building a family of structured models. 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 and compact sets adapted to the filtration $\mathcal{F}$.[8] A convenient consequence of forming a set of structured models according to formula (9) is that the associated set of probabilities satisfies a property that Epstein and Schneider (2003) call rectangularity. This property of $\mathcal{M}^o$ ensures that a dynamic version of a Gilboa and Schmeidler (1989) max-min decision maker using this as the set of probabilities would have preferences over plans that are dynamically consistent.

We add concerns about misspecification of the family of structured models associated with $\mathcal{M}^o$ in a way that preserves dynamic consistency of preferences. To understand how,

[8] 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.

we revisit (8) and note that

$$\begin{aligned}\Theta(M^U|\mathcal{F}_0) &= \theta \inf_{M^S \in \mathcal{M}^o} \Delta\left(M^U; M^S \mid \mathcal{F}_0\right) \\ &= \frac{\theta\delta}{2}\int_0^\infty \exp(-\delta t) E\left[M_t^U \xi_t(U_t) \mid \mathcal{F}_0\right] dt \end{aligned} \tag{10}$$

where

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

Equation (10) provides a particularly tractable representation of a relative entropy divergence between the probability affiliated with M^U and the family of structured models $\mathcal{M}^o$. Representation (10) exploits the separability over time of the constraints in (9) that are used to construct $\mathcal{M}^o$. It extends the entropy penalty used in the previously referenced robust control papers that assume a single structured model expressed as a single drift distortion S.[9]

Before describing how we specify the set of structured models, we make two general remarks about the restrictions embedded in (9).

Remark 3.1. *In general, sets of particular structured models that are of interest to a decision maker will not be represented as in* (9) *with a process of restrictions on the local mean processes S characterizing those models. To rescue dynamic consistency, Epstein and Schneider (2003) suggest embedding such a set of models into a larger set that can be represented for example as in* (9)*. Hansen and Sargent (2020) discuss the tension between dynamic consistency obtained in this way and using the statistical concept of admissibility to subject a worst-case model to the plausibility test recommended by Good (1952). These considerations lead us to justify* (9) *as containing the set of structured models of interest instead of as the outcome of a non-degenerate rectangular embedding.*

Remark 3.2. *We choose not to capture potential model misspecification by restricting local means to be in sets of the form* (9)*. For us, such an approach is far too constraining because we want to acknowledge possible misspecifications of the structured models using relative entropy. As Hansen and Sargent (2020) show formally, embedding relative entropy neighborhoods within a rectangular set of probabilities compels a decision maker to entertain virtually all alternative probabilities that can be represented by positive martingales with unit*

[9]Earlier models based on robust control correspond to the special case in which the set of martingales $\mathcal{M}^o$ defined in equation (9) is a singleton.

expectations, a vast set of models that includes ones with arbitrarily large relative entropy discrepancies. To avoid this extreme outcome, we move beyond the max-min preference specifications of Gilboa and Schmeidler (1989) and Epstein and Schneider (2003) and use penalties (10) *to explore the consequences of ambiguity about models within the set of structured models $\mathcal{M}^o$ as well as specification doubts about all models with that set of structured models. In section 4, we show that the resulting preferences fit within a continuous-time counterpart of the dynamic variational preference specification of Maccheroni et al. (2006).*

3.3 Using relative entropy to restrict structured models

In this subsection, we use a function $\rho(z)$ to construct a rectangular set of structured probability models that we regard as a refinement of the set of models that satisfies a relative entropy constraint. It is a refinement in the sense that it excludes many other models that also satisfy the relative entropy constraint, models that we want to exclude for reasons discussed in detail in Hansen and Sargent (2020). We construct our rectangular set of probability models by restricting the time derivative of the conditional expectation of relative entropy.[10] Formally, we restrict the drift (i.e, the local mean) of relative entropy via a Feynman-Kac relation. We use the resulting derivative constraint to build a family of structured models. In doing this we could discount as in discrepancy (10). However, constraining a set of structured models, we find it simplest not to discount. Nevertheless, remark 3.3 below provides a discounted version of the calculations that follow.[11]

The (undiscounted) entropy for a stochastic process M^S relative to the baseline model is:

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

Evidently $\varepsilon(M^S)$ is the limit as $t \to +\infty$ of a process of mathematical expectations of time series averages

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

under the probability measure implied by M^S. Suppose that M^S is defined by the drift distortion process $S = \eta(Z)$, where Z is the Markov process governed by (1) with transition probabilities that converge to a unique well-defined stationary distribution Q under the M^S

[10] Restricting a derivative of a function at every instant is in general far more constraining than restricting the magnitude of a function itself.

[11] For the small discount rates that we use in applications including ones studied in this paper, impacts of discounting in this part of the analysis are quantitatively very small.

probability. In this case, we can use Q to evaluate relative entropy by computing:

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

We represent the local evolution of a Markov process with an instantaneous generator that is the local counterpart to the one-period transition distribution depicted as a conditional expectation operator. We construct the generator by differentiating a family of conditional expectation operators with respect to an elapsed time gap. For a diffusion, the infinitesimal generator $\mathcal{A}^\eta$ of transitions under the M^S probability is the second-order differential operator:

$$\begin{aligned}
\mathcal{A}^\eta \rho &= \frac{\partial \rho}{\partial z} \cdot (\widehat{\mu} + \sigma\eta) + \frac{1}{2}\text{trace}\left(\sigma' \frac{\partial^2 \rho}{\partial z \partial z'}\sigma\right) \\
&= \mathcal{A}^0 \rho + \frac{\partial \rho}{\partial z} \cdot (\sigma\eta)
\end{aligned}$$

for $S_t = \eta(Z_t)$, where the test function ρ resides in an appropriately defined domain of the generator $\mathcal{A}^\eta$. For such a test function:

$$\int \mathcal{A}^\eta \rho dQ = 0, \tag{11}$$

which follows from a local counterpart to Law of Iterated Expectations or the Kolomogorov forward equation for a diffusion. Thus, by finding a function ρ and a corresponding q such that

$$\mathcal{A}^\eta \rho = \frac{\mathsf{q}^2}{2} - \frac{|\eta|^2}{2}, \tag{12}$$

it follows from (11) that[12]

$$\varepsilon(M^S) = \frac{1}{2}\int |\eta|^2 dQ = \frac{\mathsf{q}^2}{2}.$$

The positive number q is a mean-square measure of the size of the corresponding drift discrepancy. A function ρ that satisfies (12) allows us to impose a long-horizon refinement of the relative entropy constraint in the sense that

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

[12]The test function ρ stated here is evidently defined only up to translation by a constant.

To help appreciate distinct roles of q and ρ, we start by computing discounted relative entropy:

$$\varepsilon(M^S;\delta) \doteq \frac{\delta}{2} E\left(\int_0^\infty \exp(-\delta\tau) M^S_\tau |S_\tau|^2 d\tau \mid \mathcal{F}_0 \right)$$

for arbitrarily small discount rates. With discounting, we solve the counterpart to (12)

$$\mathcal{A}^\eta \rho = \delta\rho - \frac{|\eta|^2}{2} \tag{13}$$

for a function ρ that depends implicitly on δ. For each δ this is a Feynman-Kac partial differential equation. One way to compute $\varepsilon(M^S)$ is to take a limit of a discounted counterpart

$$\varepsilon(M^S) = \lim_{\delta \downarrow 0} \varepsilon(M^S;\delta).$$

In Feynman-Kac partial differential equations indexed by a discount rate δ, $\delta\rho$ usually has a well defined limit as δ tends to zero; for us this is the relative entropy limit $\mathsf{q}^2/2$. The fact that $\frac{\partial\rho}{\partial z}$ but not $\delta\frac{\partial\rho}{\partial z}$ typically has a well defined limit as δ tends to zero explains why there is a sharp distinction between the roles of q and $\frac{\partial\rho}{\partial z}$ in the $\delta = 0$ limiting case.

Having described how we compute relative entropy, $\mathsf{q}^2/2$, and a corresponding function, $\rho(z)$, that gives a refined characterization of relative entropy, we move on to tell how we restrict a family of potential structured models. In addition to specifying $\mathsf{q}^2/2$, we now also specify ρ *a priori* up to a translation term. For reasons discussed in Hansen and Sargent (2020), restricting q alone is insufficient to allow us to get a set of martingales expressible in the form (9). To bring us a representation of the form (9), we require that the S process belong to the sequence of sets

$$\Xi_t = \left\{ s : \mathcal{A}^0\rho(Z_t) + \frac{\partial\rho}{\partial z}(Z_t) \cdot [\sigma(Z_t)s] \leqslant \frac{\mathsf{q}^2}{2} - \frac{|s|^2}{2} \right\} \tag{14}$$

for a given choice of (q,ρ). The boundary of a set Ξ_t defined in this way includes models having the same long-horizon relative entropy $\mathsf{q}^2/2$ and also the same refinement $\rho(z) - \int \rho dQ$ of relative entropy. For a given sequence of sets Ξ_t defined by (14), there exist many S processes that have relative entropy $\varepsilon(M^S)$ less than or equal to $\mathsf{q}^2/2$ but that violate the inequality on the right side of definition (14). This is the sense in which, by using the pre-specified function $\rho(z)$ and the sequence of sets Ξ defined by equation (14) to form the set of probabilities defined in (9), we are refining (i.e., strengthening) a relative entropy constraint: many processes satisfy the relative entropy constraint but violate the rectangularity constraint incorporated in definition (14).

Remark 3.3. *Had we discounted at rate $\delta > 0$, we would have been led to specify a function ρ satisfying*

$$\Xi_t = \left\{ s : \mathcal{A}^0 \rho(Z_t) + \frac{\partial \rho}{\partial z}(Z_t) \cdot [\sigma(Z_t) s] \leqslant \delta \rho(Z_t) - \frac{|s|^2}{2} \right\}.$$

A stochastic process $\eta(Z) = S$ on the boundary of Ξ has a function ρ that satisfies a Feynman-Kac equation of the form:

$$\mathcal{A}^0 \rho + \frac{\partial \rho}{\partial z} \cdot (\sigma \eta) - \delta \rho + \frac{|\eta|^2}{2} = 0.$$

Absent discounting, ρ is well defined only up to a translation, but then q comes into play. With discounting, the function ρ alone determines the set Ξ_t and a counterpart to q is captured by the level of the function.

When we solve a robust planner's problem in section 4.2.2, it will turn out to be straightforward to characterize the set Ξ_t because it is constructed by constraining a quadratic function of s given Z_t. The set of possible s's is a disc with state-dependent center $-\sigma' \frac{\partial \rho}{\partial z}$ and radius $\mathsf{q}^2/2 - \mathcal{A}^0 \rho$. As mentioned above, if our decision maker were interested only in the set of models defined by (9) and (14), we could stop here and use a dynamic version of the min-max preferences of Gilboa and Schmeidler (1989). That way of proceeding could indeed lead to interesting applications and is worth pursuing. But the investor to be studied in this paper wants also to investigate the utility consequences of models not in the set defined by (9) because he understands that all of his structured models are best interpreted as misspecified statistical approximations.

4 Recursive Representations of Preferences and Decisions

This section prepares the way for the section 6 quantitative application by describing a set of structured models and a continuation value process over consumption plans. A scalar continuation value stochastic process ranks alternative consumption plans. 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 $\{C_t : t \geqslant 0\}$, the continuation value process $\{V_t : t \geqslant 0\}$ is

$$V_t = \min_{\{U_\tau : t \leqslant \tau < \infty\}} E\left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M^U_{t+\tau}}{M^U_t} \right) \left[\psi(C_{t+\tau}) + \left(\frac{\theta\delta}{2} \right) \xi_{t+\tau}(U_{t+\tau}) \right] d\tau \mid \mathcal{F}_t \right) \tag{15}$$

where ψ is an instantaneous utility function and $\xi_t(U_t) = \inf_{S_t \in \Xi_t} |U_t - S_t|^2$. For tractability, we set $\psi = \log$ in computations below. While logarithmic utility is indeed special, it is an enlightening specification in this and other settings. Equation (15) builds in a recursive structure that can be expressed as

$$\begin{aligned} V_t = \min_{\{U_\tau : t \leqslant \tau < t+\epsilon\}} \Bigg\{ & E\left[\int_0^\epsilon \exp(-\delta\tau) \left(\frac{M^U_{t+\tau}}{M^U_t} \right) \left[\psi(C_{t+\tau}) + \left(\frac{\theta\delta}{2} \right) \xi_{t+\tau}(U_{t+\tau}) \right] d\tau \mid \mathcal{F}_t \right] \\ & + \exp(-\delta\epsilon) E\left[\left(\frac{M^U_{t+\epsilon}}{M^U_t} \right) V_{t+\epsilon} \mid \mathcal{F}_t \right] \Bigg\} \end{aligned} \tag{16}$$

for $\epsilon > 0$. Heuristically, we can "differentiate" the right-hand side of (16) 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:

$$dV_t = \nu_t dt + \varsigma_t \cdot dW_t.$$

A local counterpart to (16) is

$$0 = \min_{U_t} \left[\psi(C_t) + \frac{\theta\delta}{2} \xi_t(U_t) - \delta V_t + U_t \cdot \varsigma_t + \nu_t \right], \tag{17}$$

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 $\frac{dM^U_t}{M^U_t}$ 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 (17), a decision maker evaluates utility consequences of a set of models that includes unstructured models that our relative entropy measure asserts are difficult to distinguish from members of the set of structured models $\mathcal{M}^o$. In particular, to construct a set of models, the decision maker:

1) Begins with a Markovian baseline model (1).

2) Creates from the baseline model a set $\mathcal{M}^o$ of *structured* models by naming a sequence of closed convex sets $\{\Xi_t\}$ that satisfy (14) 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.

We now describe how to implement some of these steps for the section 6 quantitative model. We begin by describing the baseline model used by a key decision maker in our application, a robust planner. For step 1, the planner uses a particular instance of the diffusion (1) as a Markovian baseline model. Step 2 adds other Markovian models. Step 3 includes statistically similar models that are not necessarily Markovian.

4.2.1 Step 1

For the decision maker's baseline model, we use a single capital version of an Eberly and Wang (2011) model with a long-term risk state z. The decision maker is a robust planner who faces an AK model subject to adjustment costs with capital evolution:

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

where ϕ is convex with $\phi(0) = 0$, K_t is the capital stock, 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. 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 evolution equations as follows:

$$\begin{aligned} d\log K_t &= \left[\widehat{\alpha}_k + \widehat{\beta}_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) - \frac{|\sigma_k|^2}{2}\right] dt + \sigma_k \cdot dW_t \\ dZ_t &= \left(\widehat{\alpha}_z - \widehat{\beta}_z Z_t\right) dt + \sigma_z \cdot dW_t. \end{aligned} \tag{18}$$

4.2.2 Step 2

A planner forms the following collection of structured parametric models:

$$\begin{aligned} d\log K_t &= \left[\alpha_k + \beta_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) - \frac{|\sigma_k|^2}{2}\right] dt + \sigma_k \cdot dW_t^S \\ dZ_t &= (\alpha_z - \beta_z Z_t)\, dt + \sigma_z \cdot dW_t^S, \end{aligned} \tag{19}$$

where parameters $(\alpha_k, \beta_k, \alpha_z, \beta_z)$ distinguish structured models (19) from the baseline model, (σ_k, σ_z) are parameters common to model (18) and all models (19), W^S is a 2×1 Brownian motion, and the Brownian motions W and W^S are related by

$$dW_t = S_t dt + dW_t^S, \tag{20}$$

where S_t is the drift distortion implied by parameter values $(\alpha_k, \beta_k, \alpha_z, \beta_z)$. Collection (19) nests baseline model (18).

We represent members of a parametric class defined by (19) 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 (20) 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

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

To compute relative entropy $\mathsf{q}^2/2$ and the function $\rho(z)$, we apply the method of undetermined coefficients to solve the following instance of differential equation (12):

$$\frac{d\rho}{dz}(z)[-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot \eta(z)] + \frac{|\sigma_z|^2}{2}\frac{d^2\rho}{dz^2}(z) - \frac{\mathsf{q}^2}{2} + \frac{|\eta(z)|^2}{2} = 0. \tag{21}$$

Under parametric alternatives (19), ρ is quadratic in $z - \bar{z}$:

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

We first 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 $\mathsf{q}^2/2$.

We assume that the robust planner's instantaneous utility function is logarithmic. Then guess that the value function takes the additively separable form $\Psi(x) = \log k + \widehat{\Psi}(z)$, 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:

$$\begin{aligned} 0 = \max_i \min_s \Big\{ & \delta \log(\kappa - i) - \delta\widehat{\Psi}(z) + \widehat{\alpha}_k + \widehat{\beta}_k z + i - \phi(i) + \sigma_k \cdot s \\ & + [-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot s]\frac{d\widehat{\Psi}}{dz}(z) + \frac{1}{2}|\sigma_z|^2 \frac{d^2\widehat{\Psi}}{dz^2}(z) \Big\}, \end{aligned} \tag{22}$$

where i is a potential choice of the investment-capital ratio and s is a potential choice of the structured drift distortion and where s satisfies restriction (14), which we rewrite as:

$$[\rho_1 + \rho_2(z-\bar{z})]\left[-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot s\right] + \frac{|\sigma_z|^2}{2}\rho_2 - \frac{\mathsf{q}^2}{2} + \frac{s \cdot s}{2} \leqslant 0, \tag{23}$$

an inequality implied by our quadratic $\rho(z)$ function.

By fixing $(\rho_1, \rho_2, \mathsf{q})$, we can trace out a one-dimensional family of parametric models having the same relative entropy. For example, given $(\rho_1, \rho_2, \mathsf{q})$, we can first solve equation (21) 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 as functions of z. In this way, nonlinear structured models are included in the set of structured models near the baseline model as measured by relative entropy. These nonlinear models also have relative entropy $\mathsf{q}^2/2$. We can represent the resulting nonlinear model as a time-varying coefficient model by solving

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

for η_0 and η_1, z by z, along the one-dimensional curve in η_0 and η_1. We provide the following example upon which we shall base calculations to be discussed at length later in this paper.

Illustration 4.1. *In order to focus structured uncertainty on how drifts for* (K, Z) *respond to the state variable* Z*, suppose that the decision maker sets*

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

In this case, $\rho_1 = 0$ *and formula* (21) *becomes*

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

or equivalently,

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

Notice that restriction (23) *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*

$$\frac{1}{2} \eta_1 \cdot \eta_1 + \left(\frac{\mathsf{q}^2}{|\sigma_z|^2} \right) \left(-\widehat{\beta}_z + \sigma_z \cdot \eta_1 \right) \leqslant 0. \tag{24}$$

Next form the boundary of the convex set of alternative parameter configurations constrained by (24)

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

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 (22) 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 $r = \sigma s$. The two distinct entries of $r = \sigma s$ alter evolution equations for the state variables k and z, both of which appear in the objective function on the right side of HJB equation (22). Evidently, from HJB equation (22), the first entry, r_1, shifts the log capital evolution equation and the second entry, r_2, shifts the evolution equation for the exogenous state z. The criterion appearing in HJB equation (22) remains linear in r with a translation; linearity pushes the minimizing r to an ellipse that is the boundary of the convex constraint set for each z. Under calibrated parameters for the baseline model that we present in section 6, figure 1 shows ellipsoids associated with two alternative values of z.

Notations for q's: The caption of figure 1 indicates values for two versions of q, a quantity $\mathsf{q}_{s,0}$ that indicates entropy of a structured model to the baseline model that we denote model 0; and a quantity $\mathsf{q}_{u,s}$ that denotes entropy of an unstructured model relative to a structured model. Subsection 5.1 describes how we define and compute $\mathsf{q}_{u,s}$. Later we also use $\mathsf{q}_{u,0}$ to denote entropy of an unstructured model to the baseline model.

For every feasible choice of r_2, two choices of r_1 satisfy the implied quadratic equation for the ellipse mentioned above. 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 ellipse where $\frac{dr_1}{dr_2} = -\frac{d\widehat{\Psi}}{dz}(z)$ and depends on z, as Figure 1 indicates.

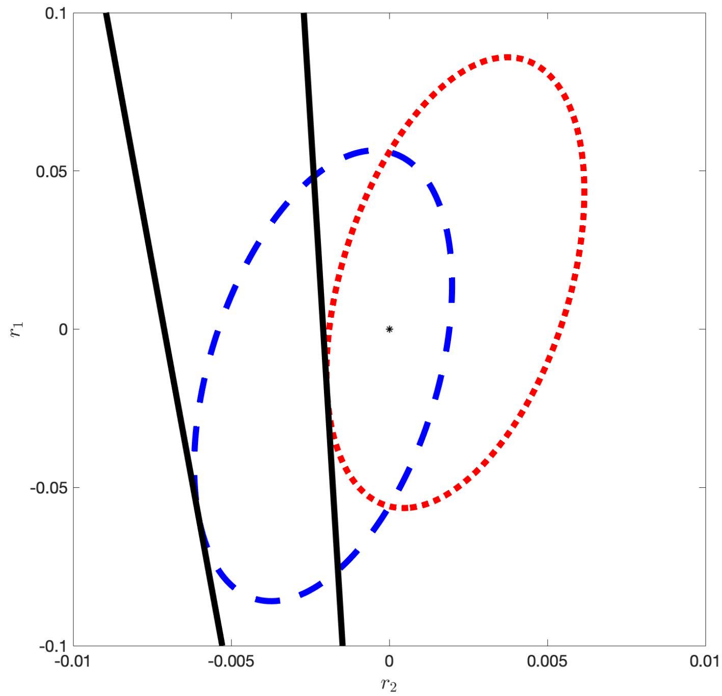


Figure 1: Two ellipsoids associated with alternative values of z. We used the same numerical values of parameters as in figure 3 from section 6 to the construct this diagram. The figure displays parameter contours for (r_1, r_2), holding relative entropies fixed at $\mathsf{q}_{s,0} = .1$ and $\mathsf{q}_{u,s} = .2$. The upper right contour, depicted with red dots, is for z equal to the .1 quantile of its stationary distribution under the baseline model and the lower left contour, depicted with blue dashes, 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.

4.2.3 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 (22):

$$
\begin{aligned}
0 = \max_i \min_{u,s} \Big\{ & \delta \log(\kappa - i) - \delta \widehat{\Psi}(z) + \widehat{\alpha}_k + \widehat{\beta}_k z + i - \phi(i) + \sigma_k \cdot u] \\
& + [-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot u] \frac{d\widehat{\Psi}}{dz}(z) + \frac{1}{2} |\sigma_z|^2 \frac{d^2 \widehat{\Psi}}{dz^2}(z) + \frac{\theta}{2} |u - s|^2 \Big\}
\end{aligned} \tag{25}
$$

where s is constrained by (23). Consider minimizing with respect to u. First-order conditions imply that

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

Substituting this choice of u into HJB equation (25) leads us to

Problem 4.2. *Robust planning problem*

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

where maximization and minimization are both 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 i^* for i is state independent, since the first-order conditions are:

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

Thus, the consumption-capital ratio is constant and logarithms of consumption and capital share a common evolution equation under the baseline 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 finding 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.[13]

[13] This outcome does not occur in environments with multiple capital stocks having different exposures

5 Alternative entropy measures

Preference orderings described in section 4 use the penalty parameter θ in conjunction with relative entropy to restrict a set of unstructured models that express the decision maker's fear that all of the structured models are misspecified. Good (1952) recommended that users of a max-min expected utility approach verify that a worst-case model is plausible.[14] We implement Good's suggestion here by characterizing both a worst-case structured model and a worst-case unstructured model and also exploring how the planner's 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 informative also to measure the magnitude of a worst-case adjustment for misspecifications of the structured models. Finally, 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 in HJB equation (25), we compute worst-case structured and unstructured models with the drift distortions

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

implied for example by the minimization that appears in the problem on the right side of equation (25).

to uncertainty. For a multiple capital stock example with a different specification of model ambiguity, see Hansen et al. (2018)

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

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 the 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 3.3 to compute entropy relative to the baseline model. In particular, instead of solving equation (21), we now solve

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

for $\mathsf{q}^2/2$ and for ρ, up to a constant of translation. We denote the solution for q as $\mathsf{q}_{u,s}$ to emphasize that it is relative entropy of an unstructured model relative to a structured model. In the application below, we report

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

as a convenient measure of the magnitude of the drift distortion of a worst-case model u relative to a worst-case model s.

Appendix A.2 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.

5.2 Chernoff entropy

A dynamic version of 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 explicit connection to a statistical decision problem makes it attractive, it is less tractable than relative entropy. To address this intractability, 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. We follow Newman and Stuck (1979) and proceed to characterize a long-run version of Chernoff entropy and show how to compute it. There are important quantitative differences when we measure Chernoff entropy globally instead of locally as in the approach of Anderson et al. (2003).[15]

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 checks whether $\log M_t^U - \log M_t^S \geqslant \mathsf{h}$. 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 based on the inequalities

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

for $0 < \gamma < 1$. 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 can bound this mistake probability for large t by following Donsker and Varadhan (1976) and Newman and Stuck (1979) and studying

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

for alternative choices of γ. Notice that the limiting rate does not depend on the choice of the threshold h, as is evident from the way that the first equality is established by bringing $\exp(-\gamma \mathsf{h})$ into play. Furthermore, the limit is often independent of the initial conditioning information. We apply these calculations for given specifications of U and S, checking that

[15]The local measure is more closely aligned with local uncertainty prices, a connection that Anderson et al. (2003) feature.

the limits are well defined.

To get the best bound, we compute

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

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

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

Setting $\Gamma(M^U, M^S) = 0$ would include only those alternative models M^U that cannot be distinguished from M^S on the basis of histories of infinite length.[16] 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:

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

so the half-life is approximately

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

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

[16] That is what is done in extensions of the rational expectations equilibrium concept to self-confirming equilibria that allow probability models to be wrong, but 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).

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, in appendix A.1 we can 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.[17]

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 sets of $\{M_t^U\}$ and $\{M_t^S\}$ processes. Some models included in these sets have troublesome but difficult to detect predictable components of consumption growth.[18]

6.1 Baseline model

We think of capital broadly 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 (18) that approximates responses of consumption to permanent shocks. Our VAR follows Hansen et al. (2008) in using several macroeconomic time series to infer information about long-term

[17] Appendix A describes how we evaluate both Chernoff entropy and relative entropy numerically for the nonlinear Markov specifications that we use in subsequent sections.

[18] 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.

consumption growth. We deduce a calibration of our baseline model (18) 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 levels of logarithms of consumption, business income, and personal dividend income be cointegrated additive functionals that share a single common martingale component that can be extracted using a method described by Hansen (2012). In Appendix B we describe our data and our method for estimating the discrete-time VAR that we use to deduce the following parameters for the baseline model (18):[19]

$$\begin{aligned} \widehat{\alpha}_c &= .484 \qquad & \widehat{\beta}_c &= 1 \\ \widehat{\alpha}_z &= 0 \qquad & \widehat{\beta}_z &= .014 \\ (\sigma_c)' &= \begin{bmatrix} .477 & 0 \end{bmatrix} \\ (\sigma_z)' &= \begin{bmatrix} .011 & .025 \end{bmatrix} \end{aligned} \tag{26}$$

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 (22) for two 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:

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

where here we use q as a synonym for $\mathsf{q}_{s,0}$. 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) associated with $\mathsf{q}_{s,0} = .1$ and $\mathsf{q}_{s,0} = .05$, respectively.

[19]We remind the reader that we set $.01\widehat{\beta}_c = \widehat{\beta}_k$, and $.01\sigma_c = \sigma_k$.

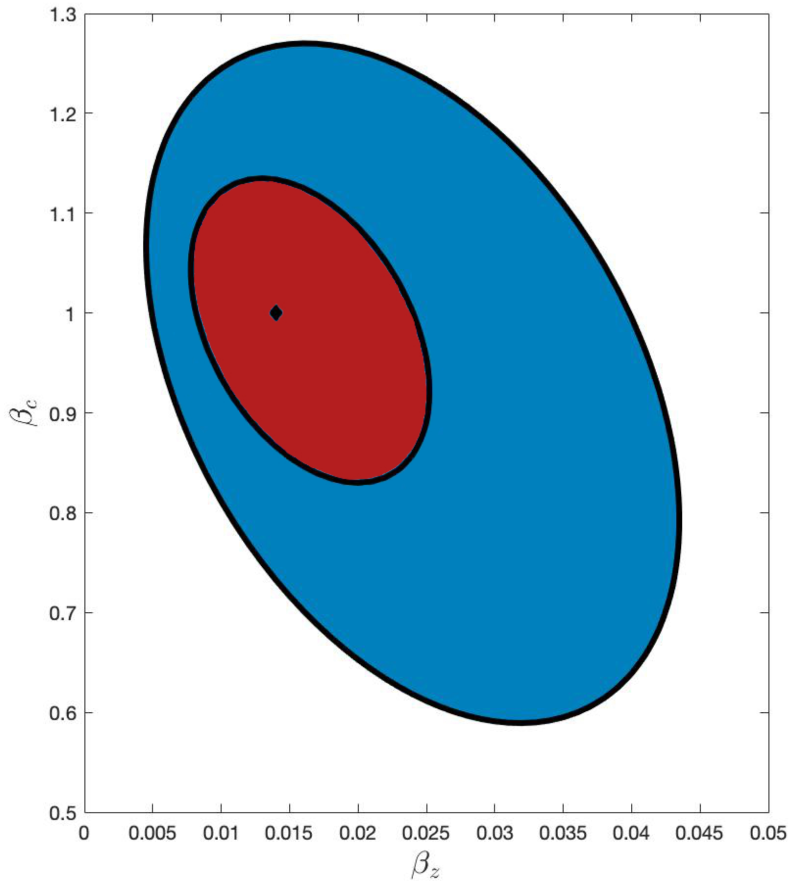


Figure 2: Parameter contours for (β_c, β_z) holding relative entropy $\mathsf{q}_{s,0}$ fixed. The outer curve is for $\mathsf{q}_{s,0} = .1$ and the inner curve is for $\mathsf{q}_{s,0} = .05$. The small diamond depicts the baseline model.

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 $r = \sigma s$ 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 in subsection 4.2 when we discussed Illustration 4.1, 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 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

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

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 that $\rho_1 = 0$, $\rho_2 = \mathsf{q}^2/|\sigma_z|^2$, and $\eta(z) = \eta_1(z - \bar{z})$, 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.

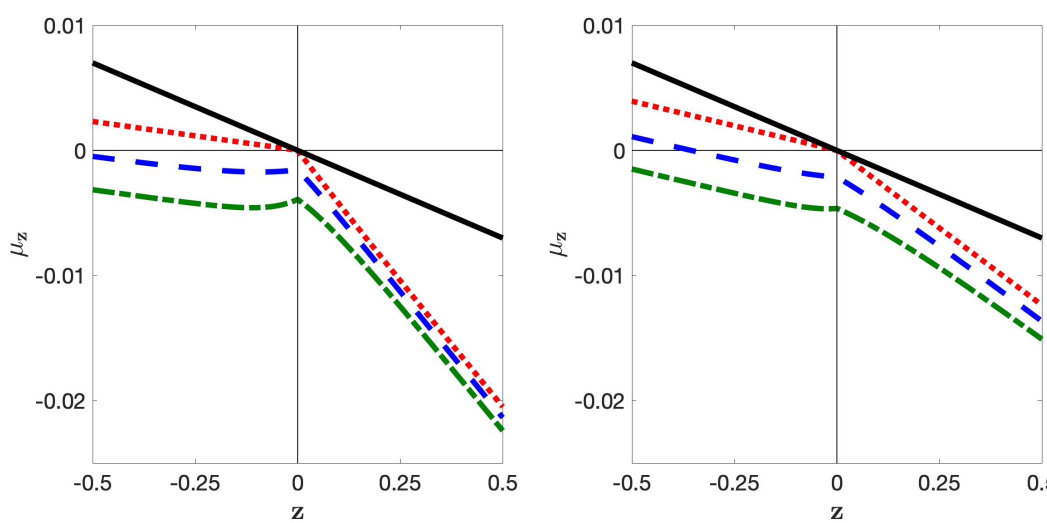


Figure 3: Worst-case structured model growth rate drifts. Left panel: larger structured entropy ($\mathsf{q}_{s,0} = .1$). Right panel: smaller structured entropy ($\mathsf{q}_{s,0} = .05$). The penalty parameter θ was reset to hit two different targeted values of $\mathsf{q}_{u,s}$. **Black** solid: baseline model; **red** dotted: worst-case structured model; **blue** dashed: $\mathsf{q}_{u,s} = .1$; and **green** dot-dashed: $\mathsf{q}_{u,s} = .2$.

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 in z 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 $\log k$ or equivalently $\log c$ is much more modest.

$\mathrm{q}_{s,0}$	$\mathrm{q}_{u,s}$	$\mathrm{d}_{u,s}$	half life u, s	$\mathrm{q}_{u,0}$	$\mathrm{d}_{u,0}$	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. $\mathsf{q}^2/2$ 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 report 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 peculiar 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. This happens because $\mathsf{q}_{u,s}$ and $\mathsf{q}_{u,0}$ are computed under the stationary probability measure implied by the worst-case unstructured model induced by U, while $\mathsf{q}_{s,0}$ 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 $(\mathsf{q}_{s,0}, \mathsf{q}_{u,s})$ configurations. The half lives associated with the $\mathsf{q}_{u,s}$'s that quantify potential model misspecification exceed 140 quarters. Even the smallest half-life associated with the $\mathsf{q}_{u,0}$ that expresses the overall discrepancy from the benchmark model equals 60 quarters. Discrimination is especially challenging when we limit the extent of model misspecification by setting $q_{u,s} = .1$.

How are the entropy measures related? We know of no formula that transforms relative entropy into long-run Chernoff entropy, but a formula from 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. It 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.

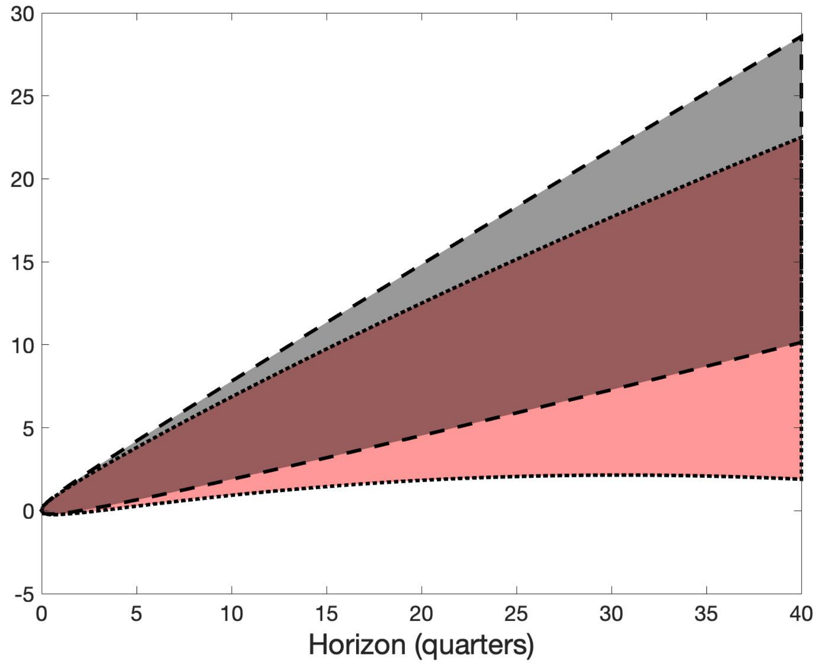


Figure 4: Distribution of $\log C_t - \log C_0$, the growth rate of consumption expressed in terms of logarithms, under the baseline model and worst-case model for $\mathsf{q}_{s,0} = .1$ and $\mathsf{q}_{u,s} = .2$. The gray shaded area between the dashed lines depicts the interval between the .1 and .9 deciles for every choice of the horizon under the baseline model. The red shaded area between the dotted lines gives the region within the .1 and .9 deciles under the worst-case model.

Figure 4 portrays effects 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 that the alternative structured models have no drift distortions for Z at $Z_t = \bar{z}$ by setting

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

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 now vary smoothly in the vicinity of $z = \bar{z}$.

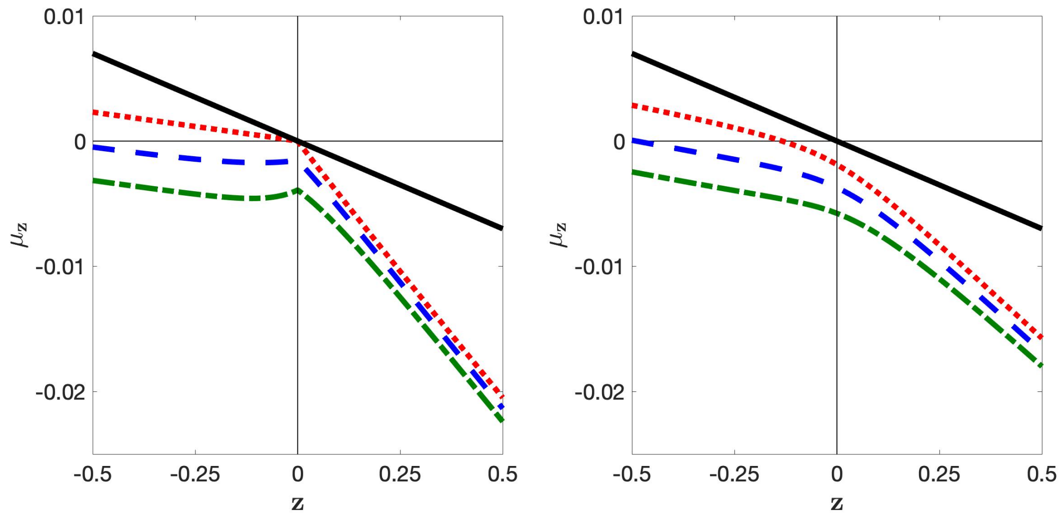


Figure 5: Distorted growth rate drift for Z. Relative entropy $\mathsf{q}_{s,0} = .1$. Left panel: $\rho_2 = \frac{(.01)}{|\sigma_z|^2}$. Right panel: $\rho_2 = \frac{(.01)}{2|\sigma_z|^2}$. **Black** solid: baseline model; **red** dotted: worst-case structured model; **blue** dashed: $\mathsf{q}_{u,s}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

Adding the restriction that $\rho_2 = 0$ makes the robust planner's value function become linear and makes 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 that 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 having model ambiguity and misspecification concerns receives for bearing risks. A representative investor cares about expected values of a discounted stream of future logarithms of consumption and chooses an instantaneous risk-free investment rate together with a portfolio of instant-by-instant exposures to dW_t by solving a continuous-time Merton portfolio problem. In equilibrium, the market compensates the investor for bearing exposure to the uncertain increment dW_t because the representative investor is averse to his ambiguity about his structured models and also to possible misspecifications of each of them. In equilibrium, these attitudes are embedded in market prices. We deduce equilibrium prices from the first-order conditions from the representative investor's portfolio problem together with restrictions on quantities that come from solving a robust planner's problem. In Appendix D we show that equilibrium market prices equal shadow prices for the robust planner's problem that we studied in section 4.

7.1 Local prices and stochastic discount factor evolution

Local prices are state dependent and compound across investment horizons in economically interesting ways. To capture these impacts, we price hypothetical cash flows over alternative investment horizons using equilibrium cumulative discount factors. Equilibrium local prices determine the local evolution of the stochastic discount factor process. Thus, the equilibrium stochastic discount factor process Sdf for our robust representative investor economy evolves as:

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

where the instantaneous risk-free interest rate is:

$$\delta + .01\left(\widehat{\alpha}_c + \widehat{\beta}_c Z_t\right) + \frac{1}{2}\left|.01\sigma_c\right|^2 + .01\sigma_c \cdot \eta^*(Z_t)$$

and $U_t^* = \eta^*(Z_t)$ is the implied worst-case drift. The local compensation vector ω^* for being exposed to aggregate shocks equals minus the local exposure to the increment dW_t:

$$\omega^*(Z_t) = (.01)\sigma_c - \eta^*(Z_t).$$

Thus, the instantaneous expected return compensation is the dot product of ω^* with the return exposure.[20] The components of ω^* are usually interpreted as local "risk prices," but the decomposition

$$\textbf{local price vector} \quad = \quad \underbrace{.01\sigma_c}_{\textbf{risk price}} \quad \underbrace{-U_t^*}_{\textbf{uncertainty price}}, \tag{27}$$

motivates us to think of $.01\sigma_c$ as "risk prices" that are induced by the curvature of log utility and $-U_t^*$ as "uncertainty prices" that are induced by the representative investor's aversion to model ambiguity and misspecification.[21] Here the uncertainty price vector $-U_t^*$ is state dependent. Appendix D describes in detail how we use competitive markets to implement the allocation chosen by a robust planner. While local prices $-U_t^*$ are large in both good and bad macroeconomic growth states, the prices of uncertainty at longer horizons display more complicated responses to shocks to the macro growth state that we now proceed to characterize.

7.2 Uncertainty prices across investment horizons

In this subsection, we deduce horizon-dependent counterparts to the local uncertainty price vector. We do so by studying how an expected return as of today varies as we alter exposures to shocks $\tau > 0$ periods into the future. Formally, we construct "shock price elasticities" that tell how state dependencies in exposures to future shocks affect expected returns today of payoffs that materialize at different different investment horizons t. We shall show that in addition to being intrinsically interesting, "shock price elasticities" defined in this way link uncertainty prices to relative entropy.

We start by considering a hypothetical payoff that equals future equilibrium consumption. The logarithm of the expected return from a consumption payoff at date t has two components, namely,

$$\log E\left(\frac{C_t}{C_0} \mid Z_0 = z\right) - \log E\left[Sdf_t\left(\frac{C_t}{C_0}\right) \mid Z_0 = z\right]. \tag{28}$$

The first component is an expected payoff and the second is the cost of purchasing that

[20]Please see equation (D.8) for derivation of this formula for $\omega^*(z)$.

[21]In this decompostion, 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.

payoff. The unitary elasticity of substitution for preferences that are logarithmic in consumption implies that

$$Sdf_t \left(\frac{C_t}{C_0} \right) = M_t^{U*}.$$

Thus, the second term in (28) is the date t component of the martingale M^{U^*} that is shaped by the representative investor's concern that he does not know which member of his set of structured models is correct and also by his concern that each of his structured models is misspecified.

To deduce prices, we change localized exposures to the Brownian increment dW_t at date t in a way that lets us define and compute a "shock-price elasticity". Formally, a shock-price elasticity quantifies the local change in an expected return that results from a local change in the exposure of consumption to the underlying Brownian motion. We use Malliavin derivatives to make a local martingale perturbation at date t in continuous time. Changing an exposure to the payoff alters both the expected payoff and its price. We analyze small changes by computing derivatives.[22] The derivative of the logarithm of the expected return given in (28) is

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

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 dW_t for consumption and the worst-case martingale, respectively.

A formula familiar from other notions of differentiation implies

$$\begin{aligned} \mathcal{D}_t C_t &= C_t \left(\mathcal{D}_t \log C_t\right) \\ \mathcal{D}_t M_t^{U*} &= M_t^{U*} \left(\mathcal{D}_t \log M_t^{U*}\right). \end{aligned}$$

The Malliavin derivative of $\log C_t$ is the vector $.01\sigma_c$, which equals both the exposure vector of $\log C_t$ to the increment dW_t and also the local risk price vector. Similarly, the Malliavin derivative of $\log M_t^{U*}$ is the vector U_t^*, which equals the exposure of $\log M_t^{U*}$ to

[22]See Borovička et al. (2014) for more about how Malliavin derivatives can be used in some related computations.

the increment dW_t and also the negative of the local uncertainty price vector. Thus,

$$\frac{E\left(\mathcal{D}_t C_t \mid \mathcal{F}_0\right)}{E\left(C_t | \mathcal{F}_0\right)} = .01\sigma_c$$
$$\frac{E\left(\mathcal{D}_t M_t^{U*} | \mathcal{F}_0\right)}{E\left(C_t \mid \mathcal{F}_0\right)} = E\left(M_t^{U*} U_t^* \mid \mathcal{F}_0\right).$$

We summarize these computations in the following proposition.

Proposition 7.1. *Shock price elasticity vectors equal*

$$\upsilon^t(z) \doteq .01\sigma_c - E\left(M_t^{U*} U_t^* \mid Z_0 = z\right)$$

and depend on both the horizon t and the initial state z.

This proposition provides a term-structure counterpart to the local price decomposition (27). The first component $.01\sigma_c$ of the price vector $\upsilon^t(z)$ is invariant over time. As in formula (27), we interpret this as the risk price component. It is very small for our example economy. The second component reflects uncertainty depends on the state and the investment horizon t and takes center stage in the computations that follow. The shock price elasticity vector $\upsilon^t(z)$ has a well defined limit as $t \to +\infty$. Specifically, $E\left(M_t^{U*} U_t^* \mid Z_0 = z\right)$ converges to the unconditional expectation of U_t^* under the limiting distorted distribution induced by U^*.

Because of how ambiguity about the structured models and concerns about misspecification of all of the structured models contribute to the U^* process, we find a further decomposition to be enlightening:

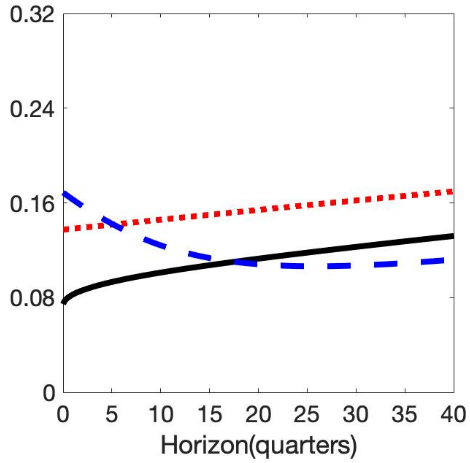
$$\underbrace{-E\left(M_t^{U*} U_t^* \mid Z_0 = z\right)}_{\textbf{uncertainty price elasticity}} = \underbrace{-E\left(M_t^{U*} S_t^* | Z_0 = z\right)}_{\textbf{ambiguity price elasticity}} \underbrace{-E\left[M_t^{U*}\left(U_t^* - S_t^*\right) \mid Z_0 = z\right]}_{\textbf{misspecification price elasticity}}. \tag{29}$$

We interpret the first term on the right side as coming from not knowing the best structured model and the second term as coming from concerns that each of the structured models might be misspecified.

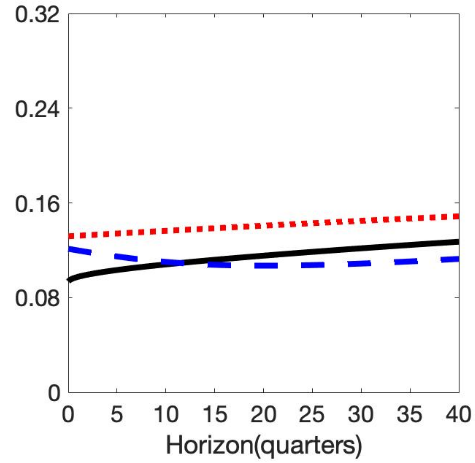
Remark 7.2. *These shock elasticities fit within a framework proposed by Borovička et al. (2011) and are related to distinct objects computed by Borovička et al. (2014). Borovička et al. (2014) use a typical impulse response timing convention by reporting elasticities that* *tell how changing exposures to a shock next period affects the expected return today of an asset that pays off t periods into the future. In contrast, here we shift the date of an asset's exposure to a shock t time periods into the future, which is also when the asset will pays out.*

7.3 Shock price elasticities for our example economy

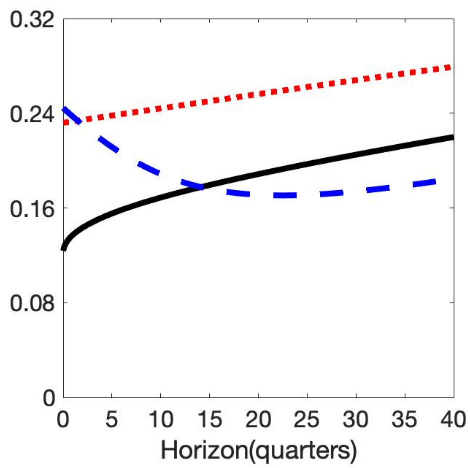
Dependency of the term structure of shock elasticities on the economic growth state encodes an interesting asymmetry in valuations and their sources. We use two figures to bring out how nonlinearities in valuation dynamics play out across investment horizons. Figure 6 shows shock price elasticities for our section 6 economy, while figure 7 plots the separate components of these elasticities defined by the right-hand side of equation (29). We feature the case in which $\mathsf{q}_{u,s} = .2$. Each of the two Brownian increments in our baseline model has its own term structure of elasticities. To generate these figures, we used the model specification from section 6, including the specific structure described in section 6.2.1.



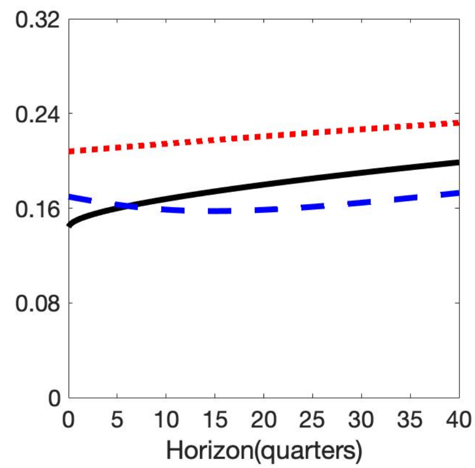
(a) first shock with $\mathbf{q}_{s,0} = .1$



(b) first shock with $\mathbf{q}_{s,0} = .05$

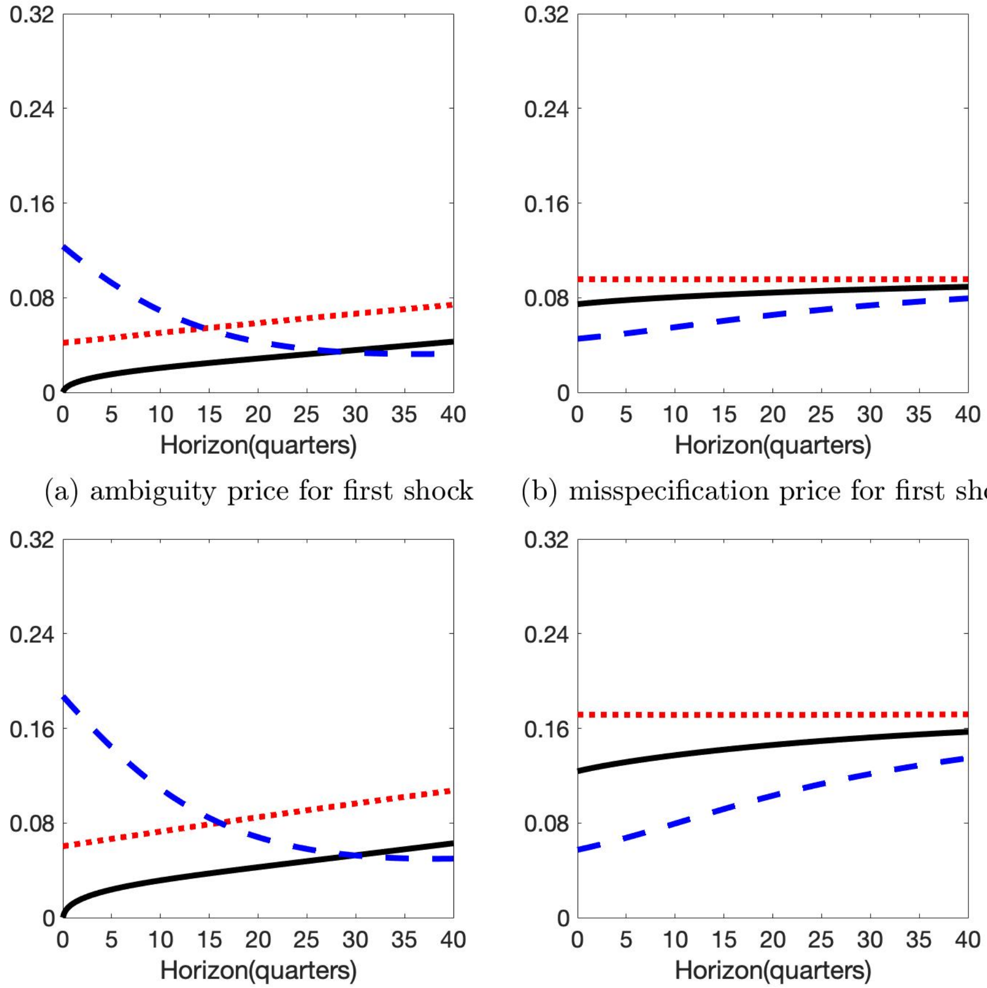


(c) second shock with $\mathbf{q}_{s,0} = .1$



(d) second shock with $\mathbf{q}_{s,0} = .05$

Figure 6: Uncertainty components of the shock price elasticities as measured by $-E\left(M_t^{U^*} U_t^* \mid Z_0 = z\right)$. The figures reports the median and deciles for the section 6 specification with (β_c, β_z) structured uncertainty. **Black** solid: median of the Z stationary distribution **red** dotted: .1 decile; and **blue** dashed: .9 decile.



(a) ambiguity price for first shock (b) misspecification price for first shock

(c) ambiguity price for second shock (d) misspecification price for second shock

Figure 7: Contributions to shock price elasticities across investment horizons. Panels in the left-hand side column plot the ambiguity component $-E\left(M_t^{U^*} S_t^* | Z_0 = z\right)$. Panels in the right-hand side column plot the misspecification component $-E\left[M_t^{U^*}\left(U_t^* - S_t^*\right) \mid Z_0 = z\right]$. In both panels $\mathsf{q}_{s,0} = .1$. The figure reports the median and deciles for the section 6 specification with (β_c, β_z) structured uncertainty. **Black** solid: median of the Z stationary distribution **red** dotted: .1 decile; and **blue** dashed: .9 decile.

Notice from Figure 6 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 ambiguity 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; but 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 making 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. Contributions from misspecification are thus important for understanding the magnitudes of the uncertainty price elasticities as well as their initial orderings and their subsequent reversals.

The model ambiguity components of the elasticities and hence the elasticities themselves diminish with horizon because the probability measure implied by the martingale M^{U^*} has reduced persistence for positive growth states. Under the M^{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 that is induced by the representative agent's min-max preferences makes uncertainty prices larger for negative than for positive values of z. In summary, horizon dependence of shock price elasticities is an important avenue through which aversion to ambiguity over the structured models and concerns about their misspecifications influence valuations of assets.

7.3.1 Uncertainty prices and relative entropy

There is a fascinating connection between relative entropy and the long-horizon limit of the uncertainty component of the elasticities defined in proposition 7.1. 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.[23] As we reported earlier, these limits equal $-E\left(M_t^{U^*}U_t^*\right)$, i.e., the unconditional expectation of the corresponding drift distortion vector computed under the worst-case stationary probability measure. Table 2 also reports $-E\left[M_t^{U^*}\left(U_t^*-S_t^*\right)\right]$, which we interpret as the misspecification component of the price vector. Uncertainty prices of the second shock are about fifty percent larger than those of the first. Misspecification concerns contribute to both of these prices, roughly one

[23]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.

third when $\mathsf{q}_{s,0} = .10$ and roughly one half when $\mathsf{q}_{s,0} = .05$. Using Table 2 information, we compare these limit prices to the relative entropy divergences, $\mathsf{q}_{u,s}$ and $\mathsf{q}_{u,0}$ that are square roots of twice the expected square of the absolute value of the vector of the uncertainty price vectors computed under worst-case stationary probability measures. That mean contributions as reflected in uncertainty prices account for most of the relative entropy measures can be seen from comparing the squares of the number in the second 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 good approximation to the magnitude of long-run uncertainty prices.

$\mathsf{q}_{s,0}$	$\mathsf{q}_{u,s}$	$\mathsf{q}_{u,0}$	shock 1 price	shock 2 price
.10	.10	.33	.17 (.05)	.26 (.09)
.10	.20	.62	.33 (.10)	.51 (.17)
.05	.10	.19	.10 (.05)	.15 (.08)
.05	.20	.36	.20 (.11)	.30 (.17)

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

We offer a final word about how we constructed our quantitative example. We designed it to illustrate a particular mechanism that causes statistically plausible amounts of uncertainty to generate fluctuations in uncertainty prices. We intentionally inferred parameters of the baseline model for these examples solely from time series of macroeconomic quantities and in doing so completely ignored asset prices and returns and so did not impose the cross-equation and cross-frequency restrictions on the consumption process that our asset pricing theory implies. We proceeded in this way in order to 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.[24]

[24]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?

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 ambiguity and potential misspecification in terms of the consequences of alternative statistically plausible models for discounted expected utilities. We characterize the effects of such concerns on the part of individual investors as shadow prices for a planner's problem.

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). Further applications seem natural. For example, the tools developed here could shed light on a recent public debate between two groups of macroeconomists and economic historians, one prophesying secular stagnation because of technology growth slowdowns, the other discounting those pessimistic forecasts.[25] 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 how our robust investor responds to the presence of 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 under the rational expectations assumption of a single-known-probability-model and so exclude ambiguity about alternative models and fears of their misspecification from the mind of their representative investor. Campbell and Cochrane construct a utility function in which the history of consumption expresses an externality that makes the investor's local risk aversion respond in a countercyclical way to the economy's growth state. Ang and Piazzesi (2003) use an exponential-quadratic stochastic discount factor in a no-arbitrage statistical model as a vehicle for exploring 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 all components of a general equilibrium model. A third approach introduces stochastic volatility into the macroecon-

[25]See Gordon and Mokyr (2016).

omy 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), we used a representative investor's robust model averaging to drive countercyclical uncertainty prices. The investor carries along two difficult-to-distinguish models of consumption growth, one with substantial growth rate persistence, 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 and exponentially twisting that posterior distribution. 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. Collin-Dufresne et al. (2016) and Andrei et al. (2019) showed that a comparable outcome emerges with a risk-based, recursive utility formulation with Bayesian learning in which investors have full confidence in the subjective probabilities and are not concerned about model misspecification. Learning occurs in the settings of these papers because the parameterized structured models are time invariant, which makes historical evidence valuable to the decision maker.

In this paper, we propose a different way to make uncertainty prices vary. We exclude learning by including alternative models with parameters whose prospective 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, a robust planner and a representative investor both worry about many other specifications. A 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 that captures concerns about the baseline model's specification of persistence in consumption growth.

Our continuous-time formulation (17) exploits mathematically convenient properties of a Brownian information structure. There exists a discrete-time version of our formulation that starts from a baseline model cast in terms of a nonlinear stochastic difference equation and in which counterparts to structured and unstructured models play the roles that they do 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. 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. We conclude by discussing these in turn.

A robust Bayesian 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 could be characterized and computed date-by-date through the use of conjugate priors. The resulting family of probabilities would not be rectangular. But if we were to augment the set of probabilities to make it rectangular as recommended by Epstein and Schneider (2003), the worst-case structured model coming from the rectangular set may not come from applying Bayes' rule to a single prior. That outcome would disable Good's recommendation for assessing the plausibility of max-min choice theory.

Another promising approach would be to 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. Peng (2004) uses a theory of stochastic differential equations under a broad notion of model ambiguity that is rich enough to allow uncertainty about the conditional volatility of Brownian increments. In Peng's setting, alternative probability specifications fail to be absolutely continuous, so standard likelihood ratio analysis doesn't apply. If we could construct bounds on uncertainty under a nondegenerate rectangular embedding, we could then modify how we would construct worst-case structured models and still restrain relative entropy as a way to limit a decision maker's set of unstructured models.[26]

In this paper, we have chosen to apply the tractable Hansen and Sargent (2020) decision

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

theoretic way of representing and distinguishing structured and unstructered uncertainties to a quantitative version of a stochastic investment model in the spirit of Eberly and Wang (2011). Replacing the single baseline model typically used in robust control theory with a set of structured models has the attractive feature that it allows us naturally to articulate the decision maker's ambiguity aversion over a targeted set of alternative models while retaining relative entropy neighborhoods to express less completely articulated misspecification concerns. Although the Hansen and Sargent (2020) decision theory adds some computational complexity, it adds no state variables. Additional computational complexity comes only in the minimization step. We extended the Eberly and Wang framework to include a richer stochastic structure and used it as our quantitative laboratory because we view the resulting model as an excellent environment for endogenizing and analyzing sources of shocks with long-run consequences that complicate the choices facing investors. In our application, this attractive bundle of "behavioral assumptions" turns out to yield sharp and quantitatively plausible implications about state-dependent countercyclical contributions to market prices of uncertainty. But perhaps more important, we believe that the way of modeling and quantifying two distinct sources of model uncertainty that we use here is a promising way to improve a broad range of quantitative dynamic stochastic models in macroeconomics, public finance, and applied welfare economics. In each of these fields, quantitative fits, operating characteristics, and responses to hypothetical alterations in government policies and regulations all depend on versions of the same inter-temporal margins that are at the heart of our analysis. The approach we have illustrated here provides a tractable set of tools for thinking creatively about how to improve quantitative models along those margins.

Acknowledgment

We thank Lloyd Han and John Wilson for help with estimation, Victor Zhorin for help calculating Chernoff entropy, Yiran Fan for help in formulating and solving ode's for robust decision problems, and Jiaming Wang for additional research assistance. Fernando Alvarez, Jarda Borovicka, Benjamin Brooks, John Campbell, Vera Chau, Timothy Christensen, Stavros Panageas, Doron Ravid, Bálint Szőke, and two anonymous referees provided insightful comments on earlier drafts. We thank the Alfred P. Sloan Foundation Grant [G-2018-11113] for the financial support. Computational code in MATLAB and Python can be found at `https://github.com/lphansen/TenuousBeliefs`.

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

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

A.1 Chernoff entropy

The Markov structure 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 $t = 0$:

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

where $[\mathbb{G}(\gamma)g](z)$ 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

$$\frac{\mathsf{q}^2}{2} - \frac{d\rho}{dz}(z)[\widehat{\alpha}_z - \widehat{\beta}_z z + \sigma_z \cdot \eta_u(z)] - \frac{|\sigma_z|^2}{2}\frac{d^2\rho}{dz^2}(z) = \frac{|\eta_u(z) - \eta_s(z)|^2}{2} \tag{A.1}$$

for q numerically using a finite difference approach like that described in section A.1. Notice that the left-hand side of (A.1) is linear in $(\rho, \mathsf{q}^2/2)$. We evaluate equation (A.1) 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 $\mathsf{q}^2/2$ 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

B.1 Calibrating the baseline model

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

i) Let

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

$$\widehat{Y}_t = \begin{bmatrix} \log G_t - \log C_t \\ \log D_t - \log C_t \end{bmatrix}$$

where as described in the body of this paper, C_t is consumption, G_t is business income, and D_t is personal dividend income. 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.

ii) Let $X_t = \begin{bmatrix} Y_t \\ Y_{t-1} \\ Y_{t-2} \\ Y_{t-3} \\ \widehat{Y}_{t-4} \end{bmatrix}$. Express a vector autoregression in the stacked form

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

where A is a stable matrix (i.e., its eigenvalues are all bounded in modulus below unity) and BB' is the innovation covariance matrix. Let selector matrix J verify $Y_{t+1} = JX_{t+1}$. The level variables $\log C_t, \log G_t, \log D_t$ are cointegrated. Each of $\log C_t, \log G_t, \log D_t$ is an additive functional in the sense of Hansen (2012). Each has an additive decomposition into trend, martingale, and stationary components that can be constructed using a method described in Hansen (2012). The martingale components of the three series are identical. The innovation to this martingale process is identified as the only shock having long-term consequences. We identify B by assuming that the square matrix JB is lower triangular.

iii) Compute the implied mean μ of the stationary distribution for X from

$$\mu = (I - A)^{-1}H$$

and the associated covariance matrix Σ that solves a discrete Lyapunov equation

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

that can be solved by a doubling algorithm.

iv) 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}.

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

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

implied by the VAR estimates. The analogue to Z_t^p for the continuous time version of the model is

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

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

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

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

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

We identify 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 VAR

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

where A is a stable matrix, 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 appendix B.1.

i) Construct a "posterior" for the coefficients of the VAR by using a 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 be uncorrelated. We impose 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 use this simulation to form a synthetic empirical distribution that assigns probability $\frac{1}{N}$ 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 weighting scheme adjusts the empirical distribution for the contribution to the likelihood function from the random initial state X_0. We 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 and minimizing worst-case r are approximately linear in the state variable. The linear approximations differ depending on whether z is greater or less than $\bar{z}$. The linear approximations provide good Neumann boundary conditions to use in an approximation that restricts z to be in a compact interval that includes $z = \bar{z}$.

Recall the constraint:

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

where $\Lambda = \left(\sigma^{-1}\right)' \sigma^{-1}$. Let d denote a vector of approximate slopes for a minimizing r. Since the quadratic terms in z dominate the constraint, impose the following restriction on d:

$$\frac{1}{2} d' \Lambda d - \rho_2 \widehat{\beta}_z + \rho_2 d_2 = 0 \tag{C.2}$$

where d_2 is the second coordinate of d. From the HJB equation:

$$\begin{aligned} (-\delta - \widehat{\beta}_z + d_2)\psi + .01(\widehat{\beta}_k + d_1) &= 0 \\ \Lambda d + \begin{bmatrix} 0 \\ \rho_2 \end{bmatrix} &\propto \begin{bmatrix} .01 \\ \psi \end{bmatrix} \end{aligned} \tag{C.3}$$

where ψ is the approximate slope of value function. The first equation in equation (C.3) is the derivative of the value function for constant coefficients, putting minimization aside. The second block in (C.3) consists of two equations derived as the large z approximation to the first-order conditions implied by (25). After taking ratios of these two latter equations we can cancel the constant of proportionality (the multiplier on the constraint) leaving us with one equation that emerges from the second block. Thus we are left solving three equations in the three unknowns d and ψ.

Depending on which boundary we target, minimization will result in different choices of d. We let d^- be the approximate solution for the left boundary with a corresponding value function derivative ψ^-. We define d^+ and ψ^+ analogously. Combining equation (C.2) and the two equations that emerge from (C.3), we are left with three equations that determine (d_1^-, d_2^-, ψ^-) and (d_1^+, d_2^+, ψ^+), 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

$$dA_t = -C_t dt + A_t \iota(Z_t) dt + A_t F_t \cdot dW_t + A_t \omega(Z_t) \cdot D_t dt, \tag{D.4}$$

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 (D.4), the vector of risk prices $\omega(z)$, and the quadratic function ρ and relative entropy $\mathsf{q}^2/2$ 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

$$\begin{aligned} 0 = \max_{c,f} \min_{u,s} & -\delta \widetilde{\Psi}(z) - \delta \log a - \delta \log \delta + \delta \log c - \frac{c}{k} + \iota(z) \\ & + \omega(z) \cdot f + f \cdot u - \frac{|f|^2}{2} + \frac{d\widetilde{\Psi}}{dz}(z) \left[-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot u \right] \\ & + \frac{1}{2} |\sigma_z|^2 \frac{d^2 \widetilde{\Psi}}{dz^2}(z) + \frac{\theta}{2} |u - s|^2 \end{aligned} \tag{D.5}$$

where extremization is subject to

$$\frac{|s|^2}{2} + \frac{d\rho}{dz}(z)[-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot s] + \frac{|\sigma_z|^2}{2} \frac{d^2 \rho}{dz^2}(z) - \frac{\mathsf{q}^2}{2} = 0. \tag{D.6}$$

The first-order condition for consumption is

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

which implies 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

$$\omega(z) + u^* - f^* = 0 \tag{D.7a}$$

$$f^* + \theta(u^* - s^*) + \frac{d\widetilde{\Psi}}{dz}(z)\sigma_z = 0. \tag{D.7b}$$

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 (D.6). 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.[27]

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.* 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 (D.7b) and the solution for s^*

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

[27] 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).

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 (D.7a), $\omega = \omega^*$, where

$$\omega^*(z) = (.01)\sigma_c - \eta^*(z). \tag{D.8}$$

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

$$\iota^*(z) = \delta + .01(\widehat{\alpha}_c + \widehat{\beta}_c z) + .01\sigma_y \cdot \eta^*(z) - \frac{.0001}{2}\sigma_c \cdot \sigma_c. \tag{D.9}$$

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 Peter 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.

Andrei, Daniel, Michael Hasler, and Alexandre Jeanneret. 2019. Asset Pricing with Persistence Risk. *The Review of Economic Studies* 32 (7):2809–2849.

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, Mark Hendricks, and José A. Scheinkman. 2011. Risk-Price Dynamics. *Journal of Financial Econometrics* 9 (1):3–65.

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.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016. Parameter Learning in General Equilibrium: The Asset Pricing Implications. *American Economic Review* 106 (3):664–98.

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.

Gilboa, Itzhak and David Schmeidler. 1989. Maxmin Expected Utility with Non-unique Prior. *Journal of Mathematical Economics* 18 (2):141–153.

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

Gordon, Robert J. and Joel Mokyr. 2016. *Boom or Doom: Debating the future of the U.S. economy.* /https://news.northwestern.edu/stories/2016/11/podcast-boom-vs-doom.

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

———. 2012. Dynamic Valuation Decomposition Within Stochastic Economies. *Econometrica* 80 (3):911–967.

Hansen, Lars Peter and Thomas J. Sargent. 1995. Discounted Linear Exponential Gaussian Control. *IEEE Transactions on Automatic Control* 40 (5):968–971.

———. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.

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

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

———. 2020. Structured Ambiguity and Model Misspecification. https://www.ssrn.com/abstract=3280597. SSRN 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.

Jacobson, David H. 1973. Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games. *IEEE Transactions for Automatic Control* AC-18:1124–1131.

James, Matthew R. 1992. Asymptotic Analysis of Nonlinear Stochastic Risk-Sensitive Control and Differential Games. *Mathematics of Control, Signals and Systems* 5 (4):401–417.

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.

Whittle, Peter. 1981. Risk Sensitive Linear Quadratic Gaussian Control. *Advances in Applied Probability* 13 (4):764–777.

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