Twisted Probabilities, Uncertainty, and Prices[*]

Lars Peter Hansen[†]   Bálint Szőke[‡]   Lloyd S. Han[§]

Thomas J. Sargent[¶]

May 21, 2019

Abstract

A decision maker constructs a convex set of nonnegative martingales to use as likelihood ratios that represent alternatives that are statistically close to a decision maker's baseline model. The set is *twisted* to include some specific models of interest. Max-min expected utility over that set gives rise to equilibrium prices of model uncertainty expressed as worst-case distortions to drifts in a representative investor's baseline model. Three quantitative illustrations start with baseline models having exogenous long-run risks in technology shocks. These put endogenous long-run risks into consumption dynamics that differ in details that depend on how shocks affect returns to capital stocks. We describe sets of alternatives to a baseline model that generate countercyclical prices of uncertainty.

[†]University of Chicago, E-mail: lhansen@uchicago.edu.
[‡]New York University, E-mail: balint.szoke@nyu.edu
[§]University of Chicago, E-mail: lloydhan@uchicago.edu
[¶]New York University, E-mail: thomas.sargent@nyu.edu.

1 Introduction

Our paper is related to one of George Tiao's many fruitful lines of work, namely, the Box and Tiao (1977) "canonical correlation" approach to linear time series analysis that anticipated later work on co-integration by time series econometricians and macroeconomists.[1] Especially relevant for our paper is Box and Tiao's use of eigenfunction methods to identify the most persistent component of a linear vector time series, techniques that Chen et al. (2009) extended to nonlinear Markov settings. These methods provide ways to identify "long-run risks," the topic that we take up in this paper.

Long-run risks are difficult to estimate well. The preferences over intertemporal outcomes that macro and finance economists typically attribute to decision makers make long-run risks play an especially important role. We explore settings in which consumers and investors acknowledge possible misspecifications of their models and adjust their decisions accordingly. Decision makers' continuation values are especially sensitive to misspecifications of those very persistent components, a sensitivity that plays a central role when we assume that decision makers use robust control theory to cope with their specification doubts.

Acknowledging that a good model is an approximation means conceding that another statistically similar alternative models might be better. This paper proposes a new way to imagine how a decision maker forms that set of alternative models and then uses it to construct equilibrium asset prices for a class of growth models that feature long-run risks.[2] We extend work by Hansen and Sargent (2001) and Hansen et al. (2006) that described a decision maker who expresses distrust of a single baseline probability model having a finite number of parameters by surrounding it with an infinite dimensional family of difficult-to-learn-about alternative models. The Hansen and Sargent (2001) decision maker represents these alternative models by multiplying baseline probabilities with likelihood ratios whose discounted entropies relative to the baseline model are less than a bound that makes alternative models stay statistically close to the baseline model. He wants to evaluate outcomes under these alternative models.[3]

---

[1] Box and Tiao also connects to rho-mixing measures of temporal dependence.

[2] Tractable ways to specify priors and compute posteriors facilitated a revolution in applied Bayesian statistics. We require an analogous practical science if the max-min expected utility decision theory elegantly axiomatized by Gilboa and Schmeidler (1989), Maccheroni et al. (2006a,b), and Strzalecki (2011) is to enlist a community of applied researchers. Viewing a set of models as a decision maker's way of coping with approximation issues is a perspective that complements theoretical work about axioms.

[3] Applications of what Hansen and Sargent (2001) and Maccheroni et al. (2006a,b) call multiplier preferences to macroeconomic policy design and dynamic incentive problems include Karantounias (2013) and

By building on insights from the the robust control theory contribution of Petersen et al. (2000), this paper differs from Hansen et al. (2006) by refining how a decision maker forms a set of models surrounding a baseline model. A new object appears here: a quadratic function ξ of a Markov state constructed from alternative parametric models that the decision maker uses to "twist" discrepancy measures to make statistical neighborhoods include these models. The decision maker wants valuations that are robust to these models in addition to unspecified models expressed as before by multiplying the baseline model by likelihood ratios. The quadratic function can be constructed to include alternatives with either fixed or time-varying parameters, and also alternative statistically similar probability models inside a convex set $\mathcal{M}(\xi)$ of martingales that we use to pose a robust decision problem. We offer a quantitative example that illustrates how the set $\mathcal{M}(\xi)$ more concisely expresses concerns about particular parametric alternatives than does the special case of that set, one associated with a particular ξ function, that was in effect used by Hansen et al. (2006).

We apply our approach to an investor who represents "the market" and whose specification uncertainty affects prices of exposures to underlying economic shocks. We describe how our twisted discrepancy method for constructing the set of probability models $\mathcal{M}(\xi)$ affects uncertainty components of these shock exposures. Our continuous-time specification simplifies asset pricing. A key object is an endogenously determined vector of worst-case drift distortions to a baseline model.[4] The negative of the drift distortion vector equals the vector of market prices of model uncertainty that compensate investors for facing ambiguity about probabilities that describe random fluctuations.

A new mechanism amplifies and makes uncertainty prices fluctuate. We introduce no new risks associated with stochastic volatility.[5] Instead, we amplify the prices of exposures to the "original" shocks. Fluctuations in those prices reflect investors' struggles to confront doubts about the baseline model. We study how uncertainty prices vary across investment horizons.

Our quantitative illustrations feature three model economies characterized by (i) AK production technologies, (ii) quadratic capital adjustment costs, and (iii) investment re-

---

Bhandari (2014).

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

[5]By way of contrast, models in which a representative investor's consumption process has innovations with stochastic volatility introduce new risk exposures in the form of the shocks to volatilities. Their presence induces time variation in equilibrium compensations for exposures to shocks that include both the stochastic volatility shocks as well as the "original" shocks whose volatilities now move.

turns subject to "long-run risk". The economies differ either in the number of investment opportunities or in those investment opportunities' relative shock exposures. The underlying differences in the three environments affect worst-case models in non trivial ways. In particular, when a representative investor can choose between two investment opportunities with asymmetric exposures to "long-run risk", interactions between the capital distribution and the persistent growth process can render the latter, which is exogenous under the baseline model, endogenous under the investor's worst-case model.

Our quantitative examples illustrate the following three phenomenon induced by the alternative probability models that we put on a representative investor's radar together with her aversion to model uncertainty:

a. An investor particularly fears decisions and acts of nature that put persistence into her consumption process.

b. When offered opportunities to invest in multiple types of capital that are characterized by different exposures to random shocks having very persistent effects, an investor tends to move her portfolio toward capital stocks that are less exposed to those shocks.

c. If she has opportunities to invest in multiple types of capital that are characterized quantitatively similar exposures to random shocks having very persistent effects but at the same time also face costs of adjusting their portfolios, an investor diversifies even more across capital stocks.

Result **a** shows how concerns about model misspecification manifest themselves in a particular direction. Results **b** and **c** operate in qualitatively similar ways to risk aversion, but we show that quantitatively model uncertainty matters more.

Section 2 specifies an investor's baseline probability model and martingale perturbations to it. Section 3 describes a function ξ that we use to characterize a set of parametric alternatives to a decision maker's baseline statistical model. It then describes relative entropy, a statistical divergence gauge of discrepancies between martingales, and uses it to define a convex set of probability measures that interest the decision maker. We express the set of probabilities in terms of a convex set $\mathcal{M}(\xi)$ of martingales that alter a baseline model. The martingale representation provides a tractable way for us to formulate a robust decision problem as a zero-sum two-player game in section 4.

By invoking a Bellman-Isaacs condition, we construct a unique worst-case model that renders the maxmin decision rule optimal against one of the probability models in the

convex set. As a consequence, the max-min decision rule is admissible.[6] That in turn allows is to implement a recommendation by Good (1952) to assess the plausibility of a worst-case model when using a max-min decision theory like that of section 4. Nevertheless, we acknowledge that we purchase admissibility at a cost because, as discussed in Hansen and Sargent (2019), there is a tension between dynamic consistency and admissibility. In this paper, we accept dynamic inconsistency by using a solution of the max-min decision problem that requires that the minimizing player who choose probabilities commits to its choice once and for all at time 0. Our formulation can be viewed as a continuous-time counterpart to the robust control theory formulation of Petersen et al. (2000). In our dynamic setting, we reassess robustness in subsequent time periods by constructing robustness bounds and associated sets of probability models for each date and for which the decision rule remains robust.

As an alternative way to confront the tension between admissibility and dynamic consistency, section 5 studies a way of attaining dynamic consistency by formulating a non-zero-sum game between a statistician and a decision maker. While we do not feature that alternative model in the quantitative illustrations provided in this paper, we think it is of interest in its own right and use it for sake of comparison.

By using estimates from Hansen and Sargent (2018) that extend and alter the empirical findings of Hansen et al. (2008), section 6 calculates key objects for a quantitative version of a baseline model together with convex sets of alternative models that concern a robust investor and a robust planner in three versions of a stochastic growth model. Section 7 constructs a recursive representation of a competitive equilibrium of an economy with a representative investor. Then it links the worst-case model that emerges from a robust planning problem to equilibrium compensations that the representative investor earns for bearing model uncertainty. Our equilibrium features an exponential quadratic stochastic discount factor whose mathematical form closely resembles one that Ang and Piazzesi (2003) used in a no-arbitrage statistical model of asset prices and macroeconomic variables. Section 8 offers concluding remarks. Two technical appendices include formulas that we use to create our quantitative illustrations .

---

[6]That it is admissible means that the robust decision rule cannot be weakly dominated for all probability models in the set and strictly dominated for some.

2 Models

This section describes how to form a set of probability models by multiplying a probability density associated with a baseline model by a set of nonnegative martingales .[7] Section 3 uses a family of parametric alternatives to a baseline model to form convex sets of martingales. Section 4 uses that set to pose a robust decision problem.

We start with a stochastic process $X \doteq \{X_t : t \geqslant 0\}$ described by a baseline model[8]

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

where W is a multivariate Brownian motion.[9] Because a decision maker does not fully trust baseline model (1), he constructs nearby probability models by multiplying probabilities associated with (1) by likelihood ratios. We represent the likelihood ratio of an alternative model by a positive martingale M^H with respect to the baseline model (1) that satisfies[10]

$$dM_t^H = M_t^H H_t \cdot dW_t \tag{2}$$

or

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

where H is progressively measurable 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. When H_t satisfies

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

with probability one, the stochastic integral $\int_0^t H_\tau \cdot dW_\tau$ is well defined[11] and is a local martingale.[12] Imposing initial condition $M_0^H = 1$, we can express a solution of stochastic

---

[7]Earlier papers sometimes referred to what we now call the baseline model as the decision maker's approximating or benchmark model.

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

[9]It is possible to generalize things to allow non Markov stochastic processes that can be constructed as functions of a Brownian motion information structure. Applications typically use Markov specifications.

[10]James (1992), Chen and Epstein (2002), and Hansen et al. (2006) used this representation.

[11]It is a limit (in probability) of the sequence of Ito integrals $\int_0^t \phi_\tau^n dW_\tau$, where $\{\phi^n\}_{n \geqslant 1}$ are simple functions approximating H on $[0,t]$ so that $\int_0^t |\phi_\tau^n - H_\tau|^2 d\tau \to 0$ in probability as $n \to \infty$.

[12]When H_t is such that $\int_0^t |H_\tau|^2 d\tau$ is infinite with positive probability, we adopt the convention that M_t^H is zero when $\int_0^t |H_\tau|^2 d\tau$ is infinite.

differential equation (2) as the stochastic exponential

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

M_t^H is a local martingale, but not necessarily a martingale.[13]

Definition 2.1. $\mathcal{M}$ *denotes the set of all local martingales* M^H *constructed via a stochastic exponential* (5) *with an H that satisfies* (4) *and is progressively measurable with respect to* $\mathcal{F}$.

To construct an alternative to the probability measure associated with baseline model (1), take any bounded $\mathcal{F}_t$-measurable random variable B_t and multiply it by M_t^H before computing a mathematical expectation conditioned on $\mathcal{F}_0$. Associated with H are probabilities defined by

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

for any $t \geqslant 0$ and any bounded $\mathcal{F}_t$-measurable random variable B_t. Thus, the positive random variable M_t^H acts as a Radon-Nikodym derivative for the conditional expectation operator $E^H [\cdot | \mathcal{F}_0]$ to be applied to date t random variables. The martingale property of the process M^H ensures that a Law of Iterated Expectations connects a family of conditional expectations operators.

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

$$dW_t = H_t dt + dW_t^H, \tag{6}$$

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

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

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

$$dX_t = \left[\hat{\mu}(X_t) + \sigma(X_t) \cdot H_t \right] dt + \sigma(X_t) dW_t^H. \tag{8}$$

---

[13]It is inconvenient here to impose either Kazamaki's or Novikov's sufficient conditions for the stochastic exponential to be a martingale. Instead we will verify that minimizers of pertinent problems do indeed result in martingales.

3 Sets of models

We use entropy relative to the probability implied by baseline model (1) to delineate a set of statistically nearby models. From (2) and (3) it follows that $M^H \log M^H$ evolves as an Ito process with drift

$$\nu_t \doteq \frac{1}{2} M_t^H |H_t|^2,$$

so we can write its conditional mean in terms of a history of local means[14]

$$E\left[M_t^H \log M_t^H \mid \mathcal{F}_0\right] = E\left(\int_0^t \nu_\tau d\tau \mid \mathcal{F}_0\right) = \frac{1}{2} E\left(\int_0^t M_\tau^H |H_\tau|^2 d\tau \mid \mathcal{F}_0\right). \qquad (9)$$

To obtain a notion of *relative entropy* of a martingale process M^H with respect to baseline model (1), we divide (9) by t and let t go to infinity, which leads to

$$\begin{aligned}
\Delta^*\left(M^H\right) &\doteq \lim_{t\to\infty} \frac{1}{2t} E\left(\int_0^t M_\tau^H |H_\tau|^2 d\tau \mid \mathcal{F}_0\right) \\
&= \lim_{\delta\downarrow 0} \frac{\delta}{2} E\left(\int_0^\infty \exp(-\delta\tau) M_\tau^H |H_\tau|^2 d\tau \mid \mathcal{F}_0\right).
\end{aligned} \qquad (10)$$

The second equality expresses relative entropy as an exponentially discounted mean, where scaling by δ makes the weights integrate to one. This equivalence motivates us to define *discounted relative entropy*:

$$\Delta\left(M^H \mid \mathcal{F}_0\right) \doteq \frac{\delta}{2} E\left(\int_0^\infty \exp(-\delta\tau) M_\tau^H |H_\tau|^2 d\tau \mid \mathcal{F}_0\right).$$

We shall set δ equal to the subjective rate that the decision maker uses to discount expected utility flows. Discounting makes the discrepancy measure $\Delta\left(M^H \mid \mathcal{F}_0\right)$ depend on the information set $\mathcal{F}_0$; but its $\delta \downarrow 0$ limit $\Delta^*\left(M^H\right)$ – the entropy concept commonly used in applied probability theory – is independent of the initial state X_0.

---

[14]In this paper, we simply impose the first equality. Various sufficient conditions justify this equality. When choosing probabilities that minimize expected utilities, we will use the representation after the second equalityin (10) without first imposing that M^H is a martingale, but then shall verify that the minimizing choice is a martingale. Claims 6.1 and 6.2 of Hansen et al. (2006) justify this approach.

3.1 Twisted relative entropy

For a constant $\zeta \in \mathbb{R}_+$, Hansen and Sargent (2001) used the inequality

$$\Delta(M^H \mid \mathcal{F}_0) \leqslant \zeta \tag{11}$$

to construct a set of alternative models across which a decision maker seeks robustness. A vast number of models satisfy constraint (11). In this paper, we describe how the decision maker can insist that a set include specific worrisome models. To articulate specification doubts about particular parametric alternatives in this way, we use a continuous time analogue to a discrete time method of Petersen et al. (2000). We use a non-negative function of the state variable $\xi(X)$ – call it a *twisting function* – to define *twisted relative entropy*:

$$\varrho\left(M^H; \xi \mid \mathcal{F}_0\right) \doteq \frac{\delta}{2} E\left(\int \exp(-\delta\tau) M_\tau^H [|H_\tau|^2 - \xi(X_\tau)] \mid \mathcal{F}_0\right) d\tau. \tag{12}$$

Properties that facilitate our analysis are: (i) ϱ is convex in M^H, (ii) ϱ can readily be computed. As with discounted relative entropy in (11), we bound the discrepancy measure ϱ to define a convex set of martingale processes:

$$\mathcal{M}(\xi) \doteq \left\{ M^H \in \mathcal{M} : \varrho\left(M^H; \xi \mid \mathcal{F}_0\right) \leqslant 0 \right\} \tag{13}$$

The special case $\xi(x) = 2\zeta$ implies a constraint on martingales M^H equivalent with (11), so $\mathcal{M}(\xi)$ with a constant twisting function can be viewed as a Hansen-Sargent discounted relative entropy ball.

From the viewpoint of robust control, a particularly important subset of $\mathcal{M}(\xi)$ comprises those models that satisfy the constraint on the right side of (13) with equality and thereby trace out an iso-ϱ curve of all local martingales in the set $\mathcal{M}$ with twisted relative entropy value equal to 0. Importantly, state dependence in ξ alters the shape of this curve and thereby the location of models on the boundary of $\mathcal{M}(\xi)$. This happens because a state-dependent ξ twists the shock process relative to which discounted entropy penalizes deviations. It thereby affects reallocations of probability distortions across states: when $\xi(X_t)$ is high, distorting the mean of the baseline Brownian motion W is less expensive in terms of ϱ than it is when $\xi(X_t)$ is low. In contrast, while the constant ζ of Hansen and Sargent (2001) restricts intertemporal reallocations of instantaneous distortions H_t to W_t,

it does not affect reallocations of instantaneous distortions across states.

Because it contains the martingale $M^0 \equiv 1$ associated with the baseline model (1), the set $\mathcal{M}(\xi)$ is non empty. Furthermore, unlike the set of Hansen and Sargent (2001), so long as ξ is state-dependent, the twisted set $\mathcal{M}(\xi)$ is non-degenerate even if its constant term is zero.

3.2 Focusing uncertainty

A state-dependent twisting function brings an interesting interpretation. Suppose that the decision maker has a particular parametric alternative in mind – call it a worrisome model – that she includes in the set because she wants an evaluation under this model. To highlight the special role of this worrisome alternative, we reserve the notation $\widetilde{S}$ to denote its martingale process $M^{\widetilde{S}}$. When this special worrisome model is Markov, $|\widetilde{S}_t|^2$ can be written as $\xi(X_t)$ and its discounted relative entropy is

$$\Delta\left(M^{\widetilde{S}} \mid \mathcal{F}_0\right) = \frac{\delta}{2} \int \exp(-\delta\tau) E\left(M_\tau^{\widetilde{S}} \xi(X_\tau) \mid \mathcal{F}_0\right) d\tau.$$

Consequently, by restricting the decision maker's set of models with the inequality

$$\frac{\delta}{2} \int \exp(-\delta\tau) E\left(M_\tau^H |H_\tau|^2 \mid \mathcal{F}_0\right) d\tau \leqslant \frac{\delta}{2} \int \exp(-\delta\tau) E\left(M_\tau^H \xi(X_\tau) \mid \mathcal{F}_0\right) d\tau, \qquad (14)$$

we ensure that $M^{\widetilde{S}}$ is included because $H = \widetilde{S}$ satisfies with equality the constraint that appears on the right side of equation (13). Moving the expression on the right side to the left leads to the inequality constraint of (13). In this way, we can make the set $\mathcal{M}(\xi)$ express concerns about a particular parametric model $M^{\widetilde{S}}$ embodied in the twisting function ξ. If there are multiple models $M^{\widetilde{S}}$ whose drift distortion satisfy $\xi(X_t) = |\widetilde{S}_t|^2$, by using (14) we include all of these models in the set. Thus, the twisting function provides a tractable way to supplement the baseline model with other models that are of particular interest to a decision maker while exploiting the tractability of relative entropy in dynamic settings. By casting concerns for misspecification in terms of a Lagrange multiplier on the constraint of models that are under consideration, we are able to formulate dynamic decision problems that have a convenient recursive structure while expressing concerns about model misspecifications. See Petersen et al. (2000) for further discussion of this type of formulation.

There are other ways of constructing sets of models for use in dynamic decision problems.

For instance, Hansen and Sargent (2019) propose a different way of incorporating multiple models that interest a decision maker, but they also use relative entropy to represent model misspecification concerns. Furthermore, we note that there are extensive literatures in operations research, statistics, and control theory that build ambiguity sets of probabilities by using measures of discrepancy other than relative entropy. Not surprisingly, details of the set construction matter significantly and have interesting decision theoretic and substantive implications.[15]

The following example illustrates that by placing the worrisome alternative $M^{\widetilde{S}}$ in $\mathcal{M}(\xi)$, we also include all models that are in between $M^{\widetilde{S}}$ and the baseline (1) as measured by relative entropy.

Example 3.1. *Consider a Markov alternative to baseline model (1) of the form*

$$dX_t = \widetilde{\mu}(X_t)dt + \sigma(X_t)dW_t^{\widetilde{S}}, \tag{15}$$

where $\widetilde{S} = \widetilde{\eta}(X)$ and

$$\widetilde{\mu}(x) = \widehat{\mu}(x) + \sigma(x)\widetilde{\eta}(x).$$

Set $0 \leqslant \lambda \leqslant 1$ and form

$$\xi(x) = \widetilde{\eta}(x) \cdot \widetilde{\eta}(x)$$

and

$$\mu(x,\lambda) = \widehat{\mu}(x) + \lambda\sigma(x)\widetilde{\eta}(x) = (1 - \lambda)\widehat{\mu}(x) + \lambda\widetilde{\mu}(x).$$

Martingales M^S with $S = \lambda\widetilde{\eta}(X)$ parameterized in this way delineate a family of diffusions with a common Brownian exposure $\sigma(x)$ and drifts $\mu(x,\lambda)$ that are convex combinations of $\widehat{\mu}$ and $\widetilde{\mu}$. All such martingales are included in $\mathcal{M}(\xi)$. Allowing λ to be time varying would extend the parameterized family.

While the set $\mathcal{M}(\xi)$ contains a myriad of alternative models, the worrisome model is special in the sense that it specifies a direction of deviations from the baseline model $\mathcal{M}$ about which the decision maker is especially concerned. To emphasize this, in the following subsection we define a subset of $\mathcal{M}(\xi)$ that are models that come from the parametric class of the worrisome model.

---

[15]A referee informed us of two recent contributions to the operations research literature, Jiang and Guan (2018) and Liu et al. (2018). The Jiang and Guan paper in particular studies a static decision problem with ambiguity over probability measures restrained by an L^1 norm over the densities. The authors construct an outcome that is a weighted average of a conditional value of risk and a maximum potential loss.

3.2.1 Parameter uncertainty

The twisting function can express misspecification doubts about models with parametric forms that differ from that of the baseline model. However, in our examples we shall focus on an important special case that we call "parameter uncertainty" in which both the baseline model and the worrisome model shaping the twisting function ξ belong to the same parametric class.

Thus, consider the following example. Let the baseline model have an affine drift $\widehat{\mu}(x) = \widehat{\alpha} - \widehat{\kappa}x$ and a constant volatility $\sigma(x) = \sigma$. Consider models with the same functional form as the baseline model but different (α, κ) values. For a given worrisome model $(\widetilde{\alpha}, \widetilde{\kappa})$ from this class, set

$$\widetilde{\eta}(x) = \frac{\widetilde{\alpha} - \widehat{\alpha}}{\sigma} + \frac{\widehat{\kappa} - \widetilde{\kappa}}{\sigma}x$$

and form the twisting function $\xi(x) = \widetilde{\eta}(x)^2$, namely,

$$\xi(x) = \xi_0 + 2\xi_1 x + \xi_2 x^2 \doteq \left(\frac{\widetilde{\alpha} - \widehat{\alpha}}{\sigma}\right)^2 + 2\left(\frac{\widetilde{\alpha} - \widehat{\alpha}}{\sigma}\right)\left(\frac{\widehat{\kappa} - \widetilde{\kappa}}{\sigma}\right)x + \left(\frac{\widehat{\kappa} - \widetilde{\kappa}}{\sigma}\right)^2 x^2. \quad (16)$$

By construction, the set $\mathcal{M}(\xi)$ includes the $(\widetilde{\alpha}, \widetilde{\kappa})$ model among many other models whose ϱ discrepancies from the baseline model are smaller than that of $(\widetilde{\alpha}, \widetilde{\kappa})$.

Projecting the set $\mathcal{M}(\xi)$ defined in (13) onto the parametric model class indexed by (α, κ) tells the sense in which the function ξ expresses a decision maker's concern about particular parametric alternatives. Figure 1 represents such projections of $\mathcal{M}(\xi)$ for alternative parameterizations of the quadratic ξ function that we construct in equation (16).[16] Each shaded region, induced by a different ξ, denotes those elements of the parametric class that are included in $\mathcal{M}(\xi)$. The solid lines are iso-ϱ curves associated with $\varrho = 0$. These regions are convex sets and thus include the λ-weighted drifts described in Example 3.1.

The quadratic twisting function (16) can be decomposed into a constant and a state-dependent part: $\xi(X_t) = \xi_0 + \widetilde{\xi}(X_t)$. The constant ξ_0 affects every alternative model's ϱ measure in the same way, which invites us to reinterpret the untwisted relative entropy ball of Hansen and Sargent (2001) as a set that includes a particular worrisome model that

---

[16]We evaluate $\varrho\left(M^H; \xi \mid \mathcal{F}_0\right)$ using $X_0 = 0$ (the unconditional mean) as the initial state. Under the indicated parameterization (and assuming reasonable discounting) the quantitative impact of dependence on X_0 is marginal.
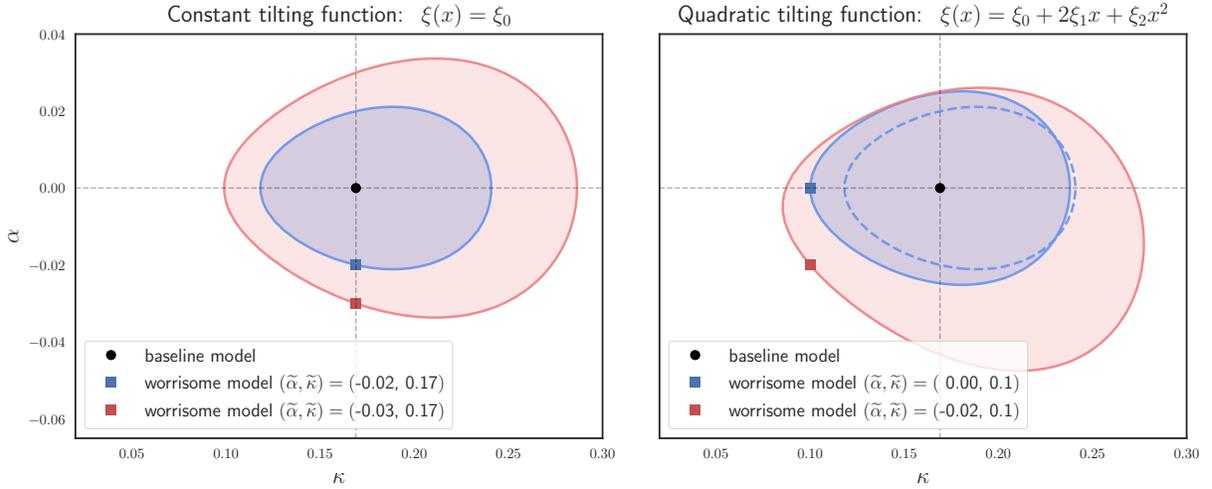
Figure 1: Projections of $\mathcal{M}(\xi)$ onto the parametric class indexed by (α, κ) for alternative ξ functions. The baseline parameters are $(\widehat{\alpha}, \widehat{\kappa}, \sigma) = (0, 0.169, 0.195)$. twisted relative entropy is evaluated using $X_0 = 0$ as the initial state. For comparison, the dashed contour on the right replicates the smaller set on the left panel associated with $\xi_0 = 0.01$.

differs from the baseline only in its intercept term. The left panel of Figure 1 illustrates this case by displaying the iso-ϱ curves with $\xi(x) = \xi_0$ for two ξ_0 values. Inspection of (16) reveals that by setting ξ_0, we include parametric models with specific α values within the set $\mathcal{M}(\xi)$. The blue and red squares depict two such models having the same persistence parameter κ as the baseline. Although higher ξ_0 values tend to enlarge the $\varrho = 0$ level sets, the general shapes of these sets do not change.

The right panel shows how a state-dependent ξ changes this situation. The blue square represents a worrisome model with a more persistent state than the baseline. The implied set is twisted toward this persistent model relative to the (smaller) untwisted set in the left panel (the dashed ellipse). Thus, by setting $\widetilde{\kappa}$ and constructing the implied $\mathcal{M}(\xi)$, we can endow a decision maker with especial fears of specific amounts of persistence. Similarly, fear of a combination of high persistence and low mean can be modeled by setting the pair $(\widetilde{\alpha}, \widetilde{\kappa})$ and constructing ξ by using (16); an example is depicted by the red square on the right panel.

13

3.3   Relative entropy neighborhoods

The twisting function ξ can guarantee that particular worrisome parametric alternatives are included in the set $\mathcal{M}(\xi)$. But $\mathcal{M}(\xi)$ also contains a vast number of less structured models that are statistically close to the parametric alternatives represented by the twisting function. To illustrate this point, we introduce notation that allows us to express a notion of entropy relative to a model $M^{\widehat{H}}$ other than the baseline model (1). We form the log likelihood ratio process $\log M^H - \log M^{\widehat{H}}$ and calculate its conditional expectation under model M^H

$$E\left[M_t^H\left(\log M_t^H - \log M_t^{\widehat{H}}\right)|\mathcal{F}_0\right] = \frac{1}{2}E\left(\int_0^t M_\tau^H |H_\tau - \widehat{H}_\tau|^2 d\tau|\mathcal{F}_0\right).$$

A corresponding notion of discounted relative entropy is

$$\widehat{\Delta}\left(M^H, M^{\widehat{H}}|\mathcal{F}_0\right) \doteq \frac{\delta}{2}E\left(\int_0^\infty \exp(-\delta\tau)M_\tau^H |H_\tau - \widehat{H}_\tau|^2 d\tau|\mathcal{F}_0\right).$$

Evidently, by setting the second argument of $\widehat{\Delta}$ to $M^{\widehat{H}} \equiv 1$, we obtain $\Delta\left(M^H \mid \mathcal{F}_0\right)$.

Consider now a drift distortion process $\widehat{S}$ that for fixed $0 < \lambda < 1$ satisfies the instant-by-instant constraint

$$|\widehat{S}_t|^2 \leqslant \lambda^2 \xi(X_t).$$

Define

$$\tau \doteq \min_H \frac{\delta}{2}\int_0^\infty \exp(-\delta\tau)E\left[M_\tau^H \xi(X_\tau)\ \Big|\ \mathcal{F}_0\right]d\tau.$$

Proposition 3.2. *Suppose that $\tau > 0$ and $0 < \lambda < 1$. If*

$$\widehat{\Delta}\left(M^H, M^{\widehat{S}}\ \mid\ \mathcal{F}_0\right) \leqslant \tau\left(1 - \lambda\right)^2,$$

then $\varrho\left(M^H; \xi\ \mid\ \mathcal{F}_0\right) \leqslant 0$ and so $M^H \in \mathcal{M}(\xi)$.

Proof. See Appendix A. □

Proposition 3.2 asserts that restricting a martingale M^H to reside in a small enough relative entropy neighborhood of $M^{\widehat{S}}$ guarantees that M^H is in $\mathcal{M}(\xi)$.

4 Robust decision problem

This section describes a way to make evaluations and decisions that are robust to worrisome misspecifications of a baseline model that we describe by a twisting function $\xi(x)$. We cast the robust decision problem as a two-player zero-sum game between a utility-minimizing agent who chooses probability models from the set $\mathcal{M}(\xi)$ and a utility-maximizing decision maker who evaluates plans and possibly chooses actions. Both agents use δ to discount the future. The minimizing player chooses a worst-case model that teaches a maximizing player aspects of alternative specifications to which evaluations of decisions are most fragile.

A feature of our formulation is that a key robustness penalty parameter can depend on the initial Markov state, a tell-tale sign that we are using a timing that assumes permanent "commitment" to a worst-case model chosen at time 0, a common feature of robust control models. Our formulation of a robust control problem is an infinite horizon, continuous time discounted version of one proposed by Petersen et al. (2000). Szőke (2018) utilized this formulation to estimate the twisting function ξ in a consumption-based asset pricing model.

4.1 A two-player zero-sum game

We now assume that the baseline model is

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

where $D \doteq \{D_t : t \geqslant 0\}$ is a decision maker's control process that is required to be progressively measurable with respect to the filtration $\mathcal{F}$. We have modified the notation in baseline model (1) to allow D to influence the local dynamics of X. We assume an instantaneous utility process $\upsilon(X, D)$.

We pose a robust decision problem as a two-player zero-sum differential game with a single value function. One player chooses decision D to maximize expected discounted utility flows, while the other player chooses drift distortion H to minimize them. Drift distortion H must be such that the associated martingale M^H satisfies

$$\varrho(M^H; \xi) \leqslant 0. \tag{17}$$

We solve this zero-sum game by first conditioning on a Lagrange multiplier ℓ attached to

constraint (17). We then compute an ℓ^* that makes (17) satisfied at equality by solving a maximization problem from Lagrange multiplier theory. This timing protocol imposes commitment to the H model chosen from the set $\mathcal{M}(\xi)$, a feature reflected in the dependence of the multiplier ℓ^* on the initial Markov state. The value function V for the differential game satisfies HJB equation

$$0 = \max_{d \in \mathcal{D}} \min_{h} - \delta V(x) + \upsilon(x, d) + \frac{\partial V}{\partial x}(x) \cdot [\hat{\mu}(x, d) + \sigma(x, d)h]$$
$$+ \frac{1}{2}\text{trace}\left[\frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x, d)\sigma(x, d)'\right] + \frac{\ell}{2}[h \cdot h - \xi(x, d)]. \qquad (18)$$

We assume a Bellman-Isaacs condition that justifies exchanging the order of maximization and minimization. The minimizing H^* takes a form encountered often in robust control theory, for example Anderson et al. (2003), namely,

$$H^*(x, \ell) = -\frac{1}{\ell}\sigma(x, d)'\frac{\partial V}{\partial x}(x),$$

which depends on the Lagrange multiplier ℓ through the common value function V. Substituting the minimizing drift distortion into HJB equation (18) gives:

$$0 = \max_{d \in \mathcal{D}} - \delta V(x) + \upsilon(x, d) + \frac{\partial V}{\partial x}(x) \cdot \hat{\mu}(x, d) + \frac{1}{2}\text{trace}\left[\frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x, d)\sigma(x, d)'\right]$$
$$- \frac{\ell}{2}\xi(x, d) - \frac{1}{2\ell}\text{trace}\left(\left[\frac{\partial V}{\partial x}(x)\right]' \sigma(x, d)\sigma(x, d)' \left[\frac{\partial V}{\partial x}(x)\right]\right).$$

The value function V depends on ℓ. To impose constraint (17), compute

$$\ell^* = \underset{\ell}{\text{argmax}} \int V(x, \ell)dQ(x),$$

where Q is a measure over initial states that we can allow to assign probability one to a particular initial state $X_0 = x$. Alternatively, we can use the implied stationary distribution for X as Q, in which case Q depends on ℓ. Dependence on the initial state or the measure Q emerges because we discount future utilities and entropies when constructing a set of models.

4.2 Worst-case models

We illustrate how features of the worst-case H^* hinge on the twisting function ξ by extending the section 3.2.1 example. As before, suppose that the baseline for the state X is affine and is hit by a single Brownian shock W

$$dX_t = (\widehat{\alpha} - \widehat{\kappa}X_t)\,dt + \sigma dW_t. \tag{19}$$

The decision maker cares about her consumption process C with period utility $v(C) = \delta \log C$. Suppose that she contemplates a particular policy indexed by $(\bar{\alpha}_c, \bar{\beta})$ that, for a given X process, implies the following consumption dynamics[17]

$$d\log C_t = \left(\bar{\alpha}_c + \bar{\beta}X_t\right)dt.$$

The agent takes $(\bar{\alpha}_c, \bar{\beta})$ as given but wants to investigate the policy's fragility to misspecifications of the state dynamics (19) by computing a worst-case model associated with consumption policy $d = (\bar{\alpha}_c, \bar{\beta})$. For a fixed policy d, the HJB equation (18) boils down to a minimization problem over H.

As in section 3.2.1, we use a quadratic ξ to focus the agent's uncertainty about parameters $(\widehat{\alpha}, \widehat{\kappa})$. Appealing to formulas from Appendix C, we find that H^* is affine in X, so the worst-case model belongs to the parametric family of the baseline indexed by (α, κ). This allows us to use our projection exercise from section 3.2.1 to produce Figure 2 which is designed to highlight the effect of ξ on the worst-case model associated with the proposed policy.

To appreciate our characterization of a worst-case model, recall that the set of alternative models includes ones with shifts in means of shocks that can occur at any future time. Included in such specifications are adverse mean shifts that are the same at all future dates. Such a permanent change in a shock mean can imitate a shift in the constant term α of the state evolution equation. Similarly, a permanent mean shift in a shock that is linear in the Markov state can imitate a change in the persistence coefficient κ. While the worst case probability can look like a change in both coefficients in the absence of state dependence in ξ, in some of our examples a worst case model imitates a shift in α only. As we will see, by including a quadratic specification of ξ that is motivated by coefficient uncertainty, changes

---

[17]We can view this as a special case of the single capital model in section 6.1, when $\sigma_k = 0$ and W is univariate. In that model, by choosing d^*, the decision maker can affect $\bar{\alpha}_c$ but not $\bar{\beta}$.

in the persistence coefficient κ emerge as part of a worst-case probability specification and include interesting interactions between α and κ.
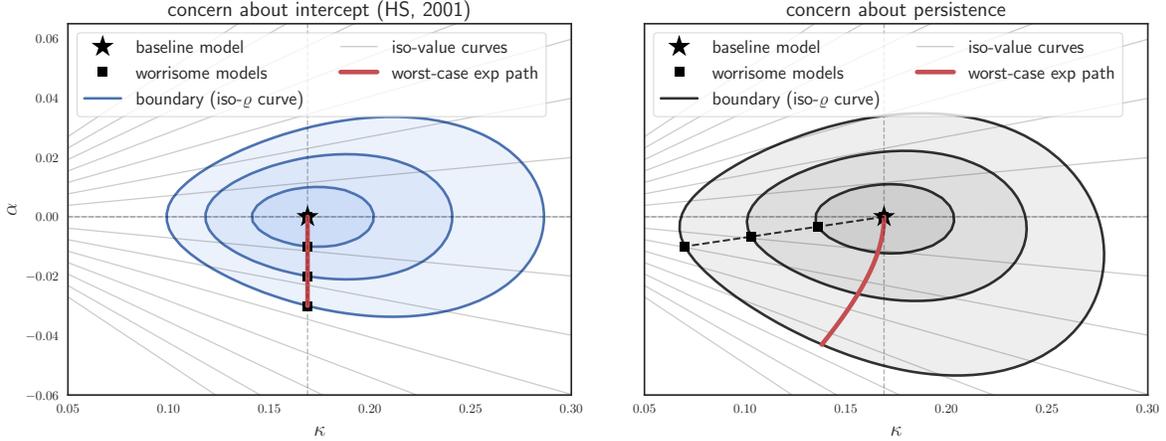

Figure 2: Iso-ϱ and iso-value curves in the parametric class indexed by (α, κ) for different twisting functions. The locus of points of tangency between these curves shows worst-case models for a sequence of constant ξ (left) and for a sequence of state-dependent ξ (right). Baseline model parameters are $(\widehat{\alpha}, \widehat{\kappa}, \sigma) = (0, 0.169, 0.195)$, the policy is $(\bar{\alpha}_c, \bar{\beta}) = (0.499, 1)$. We evaluate the iso-$\varrho$ and iso-value curves using $X_0 = 0$ as the initial state. The dash-dotted line denotes worrisome models as convex combinations of the baseline and $(\widetilde{\alpha}, \widetilde{\kappa}) = (-0.01, 0.07)$ models.

While the iso-ϱ curves circumscribe choices available to the minimizing player within the parametric class, the thin grey lines capture how the decision maker values these models when he uses a particular $(\bar{\alpha}_c, \bar{\beta})$ policy. These thin lines are iso-value curves representing those (α, κ) pairs to which the decision maker is indifferent at time 0 when she follows policy $(\bar{\alpha}_c, \bar{\beta})$. Utility increases as we move north. Evidently, the decision maker views persistence of X asymmetrically: when $\alpha < 0$, increasing persistence (lowering κ) decreases utility, but the opposite is true when $\alpha > 0$.

Figure 2 shows forcefully why it matters that a state-dependent twisting function ξ alters shapes of iso-ϱ curves and thereby positions of models relative to the baseline. The minimizing agent exhausts his "entropy budget", so a worst-case model lies on the boundary of the set: it is determined by points of tangency between iso-value curves and the boundary of $\mathcal{M}(\xi)$. The left panel shows the expansion path with respect to entropy as defined by Hansen and Sargent (2001) when the decision maker considers unstructured alternatives to her baseline model. The resulting worst-case models impart negative shifts to the intercept of X, but leave the persistence parameter κ at its baseline level. On the other hand,

the right panel shows the expansion path when the decision maker includes particular worrisome models that focus misspecification doubts on persistence. This makes worst-case state dynamics exhibit enhanced persistence.

4.3 Recursive representation of preferences

A decision maker ranks alternative consumption plans with a scalar continuation value stochastic process. 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 plans, a Hamilton-Jacobi-Bellman (HJB) equation restricts the evolution of continuation values. In particular, for a plan $\{C_t\}$, a continuation value process $\{V_t\}_{t=0}^{\infty}$ is defined by

$$V_t = \min_{\{H_\tau : t \leqslant \tau < \infty\}} E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^H}{M_t^H} \right) \left[\psi(C_{t+\tau}) + \left(\frac{\theta\delta}{2} \right) \left[|H_{t+\tau}|^2 - \xi(X_t) \right] \right] d\tau \mid \mathcal{F}_t \right)$$
(20)

where ψ is an instantaneous utility function. We can use (20) to derive an inequality that describes a sense in which a minimizing process $\{H_\tau : t \leqslant \tau < \infty\}$ isolates a statistical model that is robust. After first deriving and discussing this inequality and the associated robustness bound, we shall use (20) to present a recursive representation of preferences.

Turning to the derived bound, we proceed by applying an inequality familiar from optimization problems subject to penalties. Let H^o be the minimizer for problem (20). The process affiliated with the pair H^o gives a lower bound on discounted expected utility that can be represented in the following way.

Remark 4.1. *If H satisfies:*

$$\frac{\delta}{2} E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^H}{M_t^H} \right) \left[|H_{t+\tau}|^2 - \xi(X_t) \right] d\tau \mid \mathcal{F}_t \right)$$
$$\leqslant \frac{\delta}{2} E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^{H^o}}{M_t^{H^o}} \right) \left[|H_{t+\tau}^o|^2 - \xi(X_t) \right] d\tau \mid \mathcal{F}_t \right)$$
(21)

then

$$E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^H}{M_t^H} \right) \psi(C_{t+\tau}) d\tau \mid \mathcal{F}_t \right)$$
$$\geqslant E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^{H^o}}{M_t^{H^o}} \right) \psi(C_{t+\tau}) d\tau \mid \mathcal{F}_t \right).$$
(22)

for all $t \geqslant 0$.

Inequality (22) is a direct implication of minimization problem (20). It gives probability specifications that have date t discounted expected utilities that are at least as large as the one parameterized by H^o. Thus, inequality (21) defines the models of concern from the perspective of date t implicit in our date zero commitment formulation. This representation uses a twisted version of "continuation entropy" to express robustness concerns, namely:

$$\frac{\delta}{2} E \left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^{H^o}}{M_t^{H^o}} \right) \left[|H_{t+\tau}^o|^2 - \xi(X_t) \right] d\tau \mid \mathcal{F}_t \right).$$

But please note that we don't have to compute continuation entropy in order to construct the robust decision rules presented in sections 6 and 7.

In the next section, we describe an alternative specification of preferences that attains dynamic consistency by modifying the dynamic game.

5   Statistician and decision maker

Our formulation of robust decision making is a continuous time extension of a max-min expected utility decision problem axiomitized by Gilboa and Schmeidler (1989). Epstein and Schneider (2003) note that such preferences need not be dynamically consistent. Chen and Epstein (2002) provide continuous-time restrictions that are sufficient for dynamic consistency to prevail. But our max-min utility specification fails to satisfy those restrictions and are dynamically inconsistent in the sense that they violate the dynamic consistency axiom of Epstein and Schneider (2003). Epstein and Schneider (2003) suggest a way that we restore dynamic inconsistency by expanding the set of models that concern the decision maker. As shown by Hansen and Sargent (2019), however, their suggestion leads to a degenerate decision theory when used in conjunction with relative entropy neighborhoods of the type we want to use.

This section suggests a different way to attain dynamic consistency by modifying the section 4 formulation of the robust decision problem. This alternative formulation achieves dynamic consistency by having the minimizing player and the maximizing policy maker play a non-zero-sum game.[18] Instead of modeling robust decision making as a simultaneous zero-sum game with a single value function as we do in the remaining sections of this paper, the

---

[18]The non-zero-sum game has some flavor of a dynamic version of an adversarial generative network. See Goodfellow et al. (2014).

alternative approach temporarily under discussion in this section separates the decision-making process into two parts. In a first step, a utility-minimizing player who can be called a "robust statistician" chooses a model using an intertemporal criterion that does *not* discount the maximizing player's utility process through time. In a second step, the utility-maximizing player decision maker, who *does* discount the utility process through time, chooses actions, taking the statistician's worst-case model as given.[19]

The minimizing player's objective serves two useful purposes. The first is that by changing the criterion that the minimizing player cares about, we subtly realign payoffs so that the two players no longer play a zero-sum game. We do this in a way that renders the minimizing player's optimal choice time consistent. The second benefit from reordering payoffs is that it makes the minimizing player's decision problem align even better with techniques used in the statistical literature on model discrimination. Thus, that the robust statistician does not discount future utilities renders the multiplier on the constraint that restricts his minimizing choice of probabilities independent of the Markov state. That means that the statistician's views about what constitute a worst-case model do not change over time, delivering time consistency in his choices. We proceed to describe details next.

5.1   Two Markov decision problems

A robust statistician takes as given an instantaneous utility function $\upsilon(X, D)$ and minimizes undiscounted utility over martingales M^H that satisfy

$$\limsup_{\delta \downarrow 0} \varrho \left(M^H; \xi \right) \leqslant 0. \tag{23}$$

Replacing (17) with constraint (23) is the key difference from the section 4 formulation. This change allows us to use a relative entropy concept commonly used in applied probability theory, the second analytic benefit mentioned in the previous subsection. For our purposes, it suffices to let the statistician assume that $D_t = d(X_t)$ for a Borel measurable control law d.[20] The resulting minimization problem is a constraint preference counterpart of an infinite-horizon risk-sensitive control problem posed by Fleming and McEneaney (1995). The statistician temporarily chooses an $\mathsf{r} > 0$ and constructs a function $R(x)$ for the

---

[19]This setup is reminiscent of Brunnermeier and Parker's (2005) formulation in which one agent chooses beliefs using an undiscounted utility function while the other agent takes those beliefs as fixed when evaluating alternative plans. Brunnermeier and Parker's model was not about a decision maker's concerns about robustness to misspecification.

[20]See Hansen and Sargent (2001) for definitions of constraint and multiplier preferences.

minimizing choice of martingale M^H that makes the pair (r, R) satisfy the HJB equation[21]

$$0 = \min_h - \mathsf{r} + \upsilon[x, d(x)] + \frac{\partial R}{\partial x}(x) \cdot (\widehat{\mu}[x, d(x)] + \sigma[x, d(x)]h)$$
$$+ \frac{1}{2}\text{trace}\left(\frac{\partial^2 R}{\partial x \partial x'}(x)\sigma[x, d(x)]\sigma[x, d(x)]'\right) + \frac{\ell}{2}(h \cdot h - \xi[x, d(x)]), \quad (24)$$

where ℓ is a Lagrange multiplier on constraint (23). Given ℓ, the minimizing h satisfies:

$$H^*(x, \ell) = -\frac{1}{\ell}\sigma[x, d(x)]'\frac{\partial R}{\partial x}(x). \quad (25)$$

Substituting H^* into the HJB equation (24) gives:

$$\mathsf{r} = \upsilon[x, d(x)] + \frac{\partial R}{\partial x}(x) \cdot \widehat{\mu}[x, d(x)] + \frac{1}{2}\text{trace}\left[\frac{\partial^2 R}{\partial x \partial x'}(x)\sigma[x, d(x)]\sigma[x, d(x)]'\right]$$
$$- \frac{1}{2\ell}\text{trace}\left(\left[\frac{\partial R}{\partial x}(x)\right]'\sigma[x, d(x)]\sigma[x, d(x)]'\left[\frac{\partial R}{\partial x}(x)\right]\right) - \frac{\ell}{2}\xi(x, d(x)).$$

Here r depends implicitly on the Lagrange multiplier ℓ and is a concave function ℓ. We impose constraint (23) by computing the constant:

$$\ell^* = \underset{\ell}{\text{argmax}} \ \ \mathsf{r}(\ell).$$

That the multiplier ℓ^* is constant follows from the fact that only the time-averaged asymptotic behavior of X and H matters in determining an optimal value r. Influence of the initial state vanishes in the limit, making the solution be independent of the physical date chosen for time 0.

The decision maker takes the statistician's worst-case model $H^*(x, \ell^*)$ as given and chooses action d that solves the following discounted HJB equation:

$$0 = \max_{d \in \mathcal{D}} - \delta V(x) + \upsilon(x, d) + \frac{\partial V}{\partial x}(x) \cdot [\widehat{\mu}(x, d) + \sigma(x, d)H^*(x, \ell^*)]$$
$$+ \frac{1}{2}\text{trace}\left[\frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x, d)\sigma(x, d)'\right].$$

---

[21]Technically, this ergodic control problem requires a slightly different treatment than a problem with discounting. Finding a solution to the ergodic HJB involves finding a function $R(x)$ and a constant r, where r represents the "value function" of the control problem that does not depend on x, while $R(x)$ is used to determine the optimal action for a given state x. The function R is defined only up to an additive constant.

5.2 A caveat

As mentioned above, the game between the statistician and the decision maker presented in this section is not a zero-sum game because the two decision have different discount rates. Their differing discount factors alter the players' incentives, making them only imperfectly misaligned in just the right way to render both players' decisions time consistent.

If we were to specify preferences in this non-zero sum fashion of this section but use a timing protocol that allows the maximizing decision player to take account of the impact of its decisions on the statistician, it might lead to a different equilibrium. This is in sharp contrast with many zero-sum games, like the one in section 4, in which a Bellman-Isaacs condition allows exchanging the order of extremization.[22]

While we find the formulation of this section to be substantively interesting, in the qualitative illustrations to be described in sections 6 and 7, we nevertheless use the original max-min expected utility formulation of section 4, comforted by our recursive representation of the implied robustness bounds. In addition, we provide an example in section 6.1 in which moving to the formulation of this section would has very modest quantitative consequences. One can see this by comparing Table 1 with Table 3 in Appendix B.

6   Capital accumulation with robustness concerns

We illustrate effects of concerns about robustness in three environments using a model of capital accumulation with adjustment costs proposed by Eberly and Wang (2011). We modify their model to expose investment returns to long-run risks and make investors concerned about misspecifications of them. Three distinct example economies environments feature:

  i) a single capital stock

 ii) two capital stocks with a common exposure to long-run risk

iii) two capital stocks with only one being exposed to long-run risk

While economy i) has capital accumulation with adjustment costs, by design the implied consumption evolution conforms to a so called "long-run risk" model. Robustness concerns

---

[22]Hansen et al. (2006) describe various sufficient conditions for verifying a Bellman-Isaacs condition. See Hansen and Sargent (2008, ch. 5) for equivalent outcomes in several zero-sum games distinguished only by their timing protocols.

affect uncertainty prices but not consumption dynamics. Equilibrium uncertainty prices have an affine representation that exhibit endogenous state dependence. We use economies ii) and iii) to study how robustness concerns alter allocations between two types of capital as well as risks and their prices. By comparing economies ii) and iii), we study how exposure to growth rate uncertainty affects allocations and prices. Because they don't have affine representations, we approximate equilibrium objects for economies ii) and iii) numerically.

6.1 Single capital stock

Aggregate output is proportional to a single capital stock with a constant productivity parameter $\mathcal{A}$. A representative household cares about consumption C with instantaneous utility $\delta \log C$. Under the baseline model, investment I affects capital K according to

$$dK_t = \left[\frac{I_t}{K_t} - \frac{\phi}{2} \left(\frac{I_t}{K_t} \right)^2 + (.01) \left(\widehat{\alpha}_k + \widehat{\beta} Z_t \right) \right] K_t dt + (.01) \sigma_k K_t \cdot dW_t \qquad (26)$$

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

with adjustment cost parameter ϕ. With zero investment, the rate of change of capital is $(.01) \left(\widehat{\alpha}_k + \widehat{\beta} Z_t \right)$, where Z is an exogenously specified continuous-time scalar autoregressive process that puts long-term risks into returns on capital. Under baseline model (26), the mean of Z is $\overline{z} = \frac{\widehat{\alpha}_z}{\widehat{\kappa}}$.

The aggregate resource constraint is $C + I = \mathcal{A}K$. It is convenient to pose a robust planner's problem in terms of an investment-capital ratio $D_t \doteq \frac{I_t}{K_t}$ and $\log K_t$ that satisfy

$$d \log K_t = \left[D_t - \frac{\phi}{2} \left(D_t \right)^2 + (.01) \left(\widehat{\alpha}_k + \widehat{\beta} Z_t \right) - \frac{(.01)^2 |\sigma_k|^2}{2} \right] dt + (.01) \sigma_k \cdot dW_t.$$

Using notation from section 4, we have $X_t = \left[\log K_t, \; Z_t \right]'$ and instantaneous utility function

$$v(X_t, D_t) = \delta \log \left(\mathcal{A} - D_t \right) + \delta \log K_t.$$

24

6.1.1 Parametric alternatives

A quartet of baseline parameters $(\widehat{\alpha}_k, \widehat{\beta}, \widehat{\alpha}_z, \widehat{\kappa})$ appears in (26). A planner is concerned about alternative models for capital evolution in the following parametric class

$$d \log K_t = \left[D_t - \frac{\phi}{2} (D_t)^2 + (.01) (\alpha_k + \beta Z_t) - \frac{(.01)^2 |\sigma_k|^2}{2} \right] dt + (.01)\sigma_k \cdot dW_t^H \qquad (27)$$

$$dZ_t = (\alpha_z - \kappa Z_t) \, dt + \sigma_z \cdot dW_t^H$$

where W^H denotes a distorted shock process of the form (6). Uncertainty about (α_k, β) describes unknown components in returns to investment while uncertainty about (α_z, κ) captures an unknown growth evolution process. We represent a model of the form (27) by restricting the drift distortion H for W to satisfy

$$H_t = \eta(X_t) \doteq \eta_0 + \eta_1 Z_t \qquad (28)$$

and using (26) and (27) to deduce the following restrictions on η_0 and η_1 as functions of $(\alpha_k, \beta, \alpha_z, \kappa)$:

$$\sigma \eta_0 = \begin{bmatrix} \alpha_k - \widehat{\alpha}_k \\ \alpha_z - \widehat{\alpha}_z \end{bmatrix}, \qquad\qquad \sigma \eta_1 = \begin{bmatrix} \beta - \widehat{\beta} \\ \widehat{\kappa} - \kappa \end{bmatrix} \qquad (29)$$

where

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

Pairs (η_0, η_1) that satisfy restrictions (29) represent models having parametric form (27).

We restrict the set of alternative models by using a quadratic twisting function $\xi(z)$. Following the method of section 3.2.1 we ensure that the set $\mathcal{M}(\xi)$ includes a prespecified quartet $(\widetilde{\alpha}_k, \widetilde{\beta}, \widetilde{\alpha}_z, \widetilde{\kappa})$ by solving an instance of (29) for $\widetilde{\eta}_0$ and $\widetilde{\eta}_1$, then setting

$$\xi(x) = |\widetilde{\eta}_0 + \widetilde{\eta}_1 z|^2. \qquad (30)$$

For the quantitative examples below, either we twist the pair $(\widetilde{\alpha}_z, \widetilde{\kappa})$ and form $\xi^{[\kappa]}$, or we twist the pair $(\widetilde{\alpha}_z, \widetilde{\beta})$ and form $\xi^{[\beta]}$.

To set worrisome parameters $(\widetilde{\alpha}_z, \widetilde{\kappa})$ or $(\widetilde{\alpha}_z, \widetilde{\beta})$, we target the entropy of the implied worst-case model M^{H^*} relative to the baseline model (26). As discussed in section 3,

relative entropy $\Delta^*\left(M^H\right)$ of martingale process M^H equals one half the expectation of $|H_t|^2$ under the H induced probability model. As a result, we can parameterize relative entropy of the worst-case model using the scalar q:

$$\Delta^*\left(M^{H*}\right) = \frac{\mathsf{q}^2}{2}$$

One can think of q as a measure of the magnitude of a probability distortion. We calibrate worrisome models to induce specific q values.

To see how state-dependence of a quadratic twisting function influences this calibration, consider again the decomposition $\xi(Z_t) = \xi_0 + \widetilde{\xi}(Z_t)$ and suppose that constraint (17) is satisfied with equality to arrive at

$$\frac{\mathsf{q}^2}{2} = \Delta^*\left(M^{H*}\right) = \frac{\xi_0}{2} + \lim_{\delta\downarrow 0}\frac{\delta}{2}\int \exp(-\delta\tau)E\left(M_\tau^{H*}\widetilde{\xi}(Z_\tau)\mid\mathcal{F}_0\right)d\tau \doteq \xi_0 + \Omega(\xi). \quad (31)$$

The last equality defines $\Omega(\xi)$ as the part of relative entropy that arises from state-dependence in ξ. Indeed, when (30) holds, worrisome parameters $\widetilde{\kappa}$ and $\widetilde{\beta}$ influence the implied worst-case model only through $\Omega(\xi)$. For this reason, when we contrast the effects of $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ on the worst-case model in section 6.1.4, we will fix q and ξ_0 to make sure that the contributions from state-dependence are the same, i.e., $\Omega\left(\xi^{[\kappa]}\right) = \Omega\left(\xi^{[\beta]}\right)$.

6.1.2 Planner's problem

To capture a desire for robustness to parametric misspecifications represented by the class (27), we suppose that the twisting function has the form (30) and seek a function $V(z, k; \ell) = \log k + \nu(z; \ell)$ that satisfies the robust planner's HJB equation

$$0 = \max_d \min_h \quad -\delta\nu(z; \ell) + \delta\log(\mathcal{A} - d) + \left[d - \frac{\phi}{2}d^2 + (.01)\left(\widehat{\alpha}_k + \widehat{\beta}z - \frac{(.01)|\sigma_k^2|}{2}\right)\right]$$

$$+ \frac{\partial\nu(z; \ell)}{\partial z}\left(\widehat{\alpha}_z - \widehat{\kappa}z\right) + \frac{1}{2}\frac{\partial^2\nu(z; \ell)}{\partial z^2}|\sigma_z|^2$$

$$+ (.01)\left[\sigma_k \cdot h\right] + \frac{\partial\nu(z; \ell)}{\partial z}\left[\sigma_z \cdot h\right] + \frac{\ell}{2}\left[|h|^2 - \xi(z)\right]. \quad (32)$$

The last line of (32) depicts how concerns about robustness enter the HJB equation via h. The minimizing choice of h depends on the state z through the marginal value function:

$$H^*(z;\ell) \doteq \begin{bmatrix} h_k^* \\ h_z^* \end{bmatrix} = -\frac{1}{\ell} \left[(.01)\sigma_k + \frac{\partial \nu(z;\ell)}{\partial z}\sigma_z \right]. \tag{33}$$

The maximizing choice of d solves:

$$1 - \phi d^* = \frac{\delta}{\mathcal{A} - d^*} \tag{34}$$

where d^* denotes the chosen investment-capital ratio. Notice that d^* is constant and does not depend on the underlying state variables, so the optimal consumption-capital ratio $c^* = \mathcal{A} - d^*$ is also constant. Because neither c^* nor d^* is affected by robustness concerns, under the baseline model capital and consumption dynamics are not altered by robustness concerns. We write the implied dynamics for the logarithm of consumption as:

$$d\log C_t = (.01)\left(\widehat{\alpha}_c + \widehat{\beta}Z_t\right)dt + (.01)\sigma_c \cdot dW_t \tag{35}$$

where

$$\widehat{\alpha}_c = 100\left(d^* - \frac{\phi}{2}(d^*)^2\right) + \widehat{\alpha}_k - \frac{(.01)|\sigma_k|^2}{2} \qquad \text{and} \qquad \sigma_c = \sigma_k. \tag{36}$$

As asserted, these are independent of any robustness parameters that affect h^*. In contrast to quantity dynamics, shadow prices for exposure to uncertainty are modified by robustness concerns as intermediated through the minimizing choice of h^*. This illustrates again a finding of Hansen et al. (1999) and an approximation to it by Tallarini (2000) that robustness concerns affect prices but not quantities in models with a single capital stock.

In Appendix C.1, we show that the value function $\nu(z;\ell)$ is quadratic in z. As a consequence, h^* is affine in z. The multiplier ℓ^* is then determined by solving

$$\ell^*(z_0) = \arg\max_{\ell>0} \nu(z_0;\ell)$$

where z_0 is an initial value of Z. We can find the worst-case model by substituting $\ell^*(z_0)$ in equation (33).

6.1.3 Baseline model

The basis of our quantitative analysis is an empirical model of aggregate consumption dynamics. We follow Hansen et al. (2008) by fitting a trivariate VAR for macroeconomic time series that contain information about long-term consumption growth, namely, log consumption growth, the difference between logs of business income and consumption, and the difference between logs of personal dividend income and consumption.[23] We restrict all three time series to have a common martingale component by imposing a known cointegration relation among them.

We convert the discrete time VAR estimates to baseline parameters $(\widehat{\alpha}_k, \widehat{\beta}, \widehat{\alpha}_z, \widehat{\kappa})$ and (σ_c, σ_z) by setting $\widehat{\alpha}_z = 0$ and $\widehat{\beta} = 1$ and matching the dynamics of the VAR implied long-term consumption growth forecast with those of Z.[24] As a result, we obtain the following parameters for the baseline model of consumption (35):

$$
\begin{aligned}
\widehat{\alpha}_c &= .484 & \widehat{\beta} &= 1 \\
\widehat{\alpha}_z &= 0 & \widehat{\kappa} &= .014 \\
\sigma = \begin{bmatrix} (\sigma_c)' \\ (\sigma_z)' \end{bmatrix} &= \begin{bmatrix} .477 & 0 \\ .011 & .025 \end{bmatrix}. & &
\end{aligned}
\tag{37}
$$

We set the household's subjective discount rate equal to $\delta = .002$.

So long as condition (36) is satisfied, the particular values of the technology parameters $(\phi, \mathcal{A}, \widehat{\alpha}_k)$ are inconsequential in the single-capital case, but they will be important in section 6.2 when we study economies with two capital stocks. We now use the single capital stock economy to illustrate the impact of parameter choice on the steady state consumption/capital ratio, the investment/capital ratio and the implied rate of return. To be clear, we are not formally calibrating coefficients to match these steady states, be we are choosing them in part to illustrate the consequences of uncertainty. Formal calibration is

---

[23]The time series are quarterly data from 1948 Q1 to 2018 Q2. Our consumption measure is per capita consumption of non-durables and services from NIPA Table 1.1.5. Business income is measured as proprietor's income plus corporate profits per capita and the series are from NIPA Table 1.12. Personal dividend income is from NIPA Table 2.1. 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. Moreover, Hansen et al. (2008) did not include personal dividends in their VAR analysis.

[24]In more detail, we choose $\widehat{\alpha}_c$ and $(1-\widehat{\kappa})$ to match the VAR implied unconditional mean of consumption growth and autoregressive coefficient of expected long-term consumption growth, respectively. In addition, we set (σ_c, σ_z) such that $(1, \widehat{\kappa}^{-1})\sigma\sigma'(1, \widehat{\kappa}^{-1})'$ is equal to the VAR implied covariance matrix of the one-step-ahead forecast error for consumption growth and expected long-term consumption growth. We achieve identification by imposing a lower triangular structure with positive diagonal elements on σ.

hard to defend because of the stylized nature of the economic model.[25] We set the capital productivity parameter $\mathcal{A} = .05$. For the adjustment cost parameter, ϕ, we use an annual value of 7, By converting the annual quadratic cost parameter to quarterly units, we obtain $\phi = 28$.[26] We then use (36) to compute $\widehat{\alpha}_k$ so that $\alpha_c = .484$ is satisfied. As a result, we obtain the following values

$$\mathcal{A} = .05 \qquad\qquad \phi = 28 \qquad\qquad \widehat{\alpha}_k = -1.279 \qquad\qquad (38)$$

The implied steady states are:

$$c^* = .0182 \qquad\qquad d^* = .0318 \qquad\qquad \text{rate of return (quarterly)} = .006$$

6.1.4 Quantitative results

Table 1 reports worst-case models for three specifications of the quadratic twisting function formed according to (30). As discussed in section 6.1.1, we choose the underlying worrisome models to attain specific levels of relative entropy of the implied worst-case martingale M^{H*}, in particular, for $\mathsf{q} = .1, .2$. The top panel shows the case of a constant twisting function, which has the property that the worrisome model changes the value of $\widetilde{\alpha}_z$ relative to α_z only. This example implements the untwisted set used by Hansen and Sargent (2001).

The middle panel shows the case of $\xi^{[\kappa]}$, the twisting function that arises from altering the value of $\widetilde{\kappa}$ thereby expressing fears of specific degrees of persistence. We report results for two $\widetilde{\kappa}$ values and choose $\widetilde{\alpha}_z \leqslant \widehat{\alpha}_z$ so that the implied worst-case models achieve the prespecified q targets. The effect of concerns about exposure of consumption growth to the long-run risk state, i.e., the case with $\widetilde{\beta} \neq \widehat{\beta}$, is displayed in the bottom panel of Table 1. We make the twisting functions, $\xi^{[\kappa]}$ and $\xi^{[\beta]}$, comparable by guaranteeing that the fraction of q that comes from the state-dependent twisting, $\Omega(\xi)$, is the same. This involves using the same $\widetilde{\alpha}_z$ as for $\xi^{[\kappa]}$ and choosing $\widetilde{\beta}$ so that a prespecified q value is achieved. Accordingly, one can compare the middle and bottom panels line by line.

In all cases, worst-case models reduce α_c and impart a negative mean $\frac{\alpha_z}{\kappa}$ to the long-run risk process that describes persistence in consumption growth. Worst-case models

---

[25]For instance, capital should presumably be broadly conceived to include human capital. Moreover, we are abstracting from the durable goods component to consumption. Presumably steady state return might be better aligned with a riskless return and not the return to "risky capital" as is typically done in the real business cycle literature.

[26]The empirical evidence for this parameter is notoriously weak. For instance, Cooper and Haltiwanger (2006) note a range 3–20 range from empirical evidence.

q	$\widetilde{\alpha}_z$	$\widetilde{\kappa}$	$\widetilde{\beta}$	α_c	β	α_z	κ	$\Delta\bar{c}$	m_z	s_z
Baseline										
0.000	0.000	0.014	1.000	0.484	1.000	0.000	0.014	0.000	0.000	0.163
Constant ξ										
0.100	-0.002	0.014	1.000	0.455	1.000	-0.003	0.014	-0.219	-0.190	0.163
0.200	-0.005	0.014	1.000	0.427	1.000	-0.005	0.014	-0.437	-0.380	0.163
State-dependent $\xi^{[\kappa]}$										
0.100	-0.002	0.010	1.000	0.459	1.005	-0.003	0.013	-0.218	-0.192	0.167
0.200	-0.004	0.010	1.000	0.434	1.005	-0.005	0.013	-0.437	-0.385	0.167
0.100	-0.000	0.005	1.000	0.466	1.026	-0.002	0.010	-0.211	-0.187	0.190
0.200	-0.002	0.005	1.000	0.448	1.026	-0.004	0.010	-0.431	-0.386	0.190
State-dependent $\xi^{[\beta]}$										
0.100	-0.002	0.014	1.156	0.463	1.025	-0.002	0.010	-0.212	-0.186	0.190
0.200	-0.004	0.014	1.196	0.453	1.045	-0.003	0.008	-0.421	-0.374	0.221
0.100	-0.000	0.014	1.172	0.467	1.032	-0.002	0.009	-0.206	-0.183	0.199
0.200	-0.002	0.014	1.215	0.464	1.059	-0.002	0.006	-0.399	-0.358	0.257

Table 1: Worst-case parameter values when ξ is defined by (30). We evaluate $\ell^*(z_0)$ using $\bar{z}$ as the initial state for the calculations. $\Delta\bar{c} \doteq \left(\alpha_c + \frac{\beta\alpha_z}{\kappa}\right) - \left(\widehat{\alpha}_c + \frac{\widehat{\beta}\widehat{\alpha}_z}{\widehat{\kappa}}\right)$ denotes the change in the long run consumption growth expectation. Note that $\left(\widehat{\alpha}_c + \frac{\widehat{\beta}\widehat{\alpha}_z}{\widehat{\kappa}}\right) = .484$. m_z and s_z denote the unconditional mean and standard deviation of Z under the worst-case model.

implied by the state-dependent twisting functions also feature increased persistence of Z and enhanced exposure β of consumption growth to Z. In both cases, although the minimizing agent could choose to increase persistence and exposure further, he instead chooses to allocate some of his entropy budget to append adverse constant shifts to the Brownian increments. This trade-off differs across the two state-dependent twisting functions. Evidently, worst-case models implied by $\xi^{[\beta]}$ tend to distort (β, κ) more and (α_c, α_z) less relative to those implied by $\xi^{[\kappa]}$. The difference originates in the shapes of $\xi^{[\kappa]}$ and $\xi^{[\beta]}$.

Figure 3 displays instances of two twisting functions that correspond to given q and ξ_0 levels. As we saw in section 3.1, the relative values that ξ assigns to the different z's indicate the extent to which the minimizing agent is constrained, thereby shaping the worst-case model. More specifically, in order to reach a given level of $\Omega(\xi) > 0$, defined as the unconditional mean of $\xi(z) - \xi_0$ under model M^{H^*}, the minimizing agent avoids regions with $\xi(z) < \xi_0$ and focuses instead on z-values with $\xi(z) > \xi_0$. Given that, $\xi^{[\kappa]}$ clearly discourages positive values of z and achieves $\Omega\left(\xi^{[\kappa]}\right)$ by distorting the *mean* of Z.
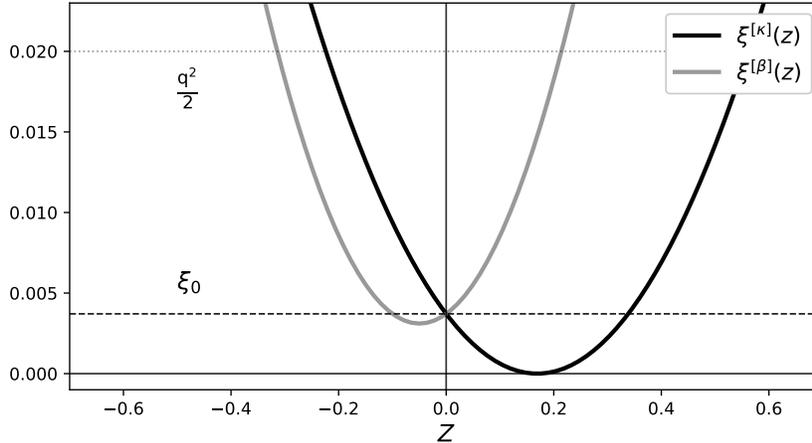
Figure 3: Twisting functions $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ associated with worrisome $(\tilde{\alpha}_z, \tilde{\kappa}) = (-.002, .005)$ and $(\tilde{\alpha}_z, \tilde{\beta}) = (-.002, 1.215)$, respectively. Dashed horizontal line shows the state-independent part of the twisting function, while the dotted horizontal line denotes relative entropy of the induce worst-case models with $q = .2$.

On the other hand, by encouraging z values with large absolute values, $\xi^{[\beta]}$ implies an H^* that distorts the *variance* of Z relatively more. These forces produce the last two columns of Table 1.

The adverse worst-case shifts have long-lasting effects on the growth rate of consumption. These effects are reflected in the quantity $\Delta \bar{c}$ that represents the difference between the unconditional consumption growth means in the worst-case and baseline models. $\Delta \bar{c}$ traces out long-term consequences of an adverse shift in the drift vector of the Brownian motion at a given date. Because an adverse shift could occur at any moment, the implied worst-case parameters reflect what would happen if an adverse drift shift occurred at *every* future moment.

The middle panel of Table 1 also shows that for any fixed $\tilde{\kappa}$, expanding the set of models by altering the worrisome $\tilde{\alpha}_z$ leaves the worst-case persistence parameter κ and exposure parameter β intact, inducing extra distortions only in the intercept terms α_c and α_z. This is the same effect that we can see in the top panel showing the case without twisting.

In Table 2 we compare relative entropy to Chernoff entropy and the implied half lives. We use half-lives defined by $HL \doteq \frac{\log 2}{c}$, where c is Chernoff entropy of the worst-case model. Chernoff entropy is a sharp upper bound on the asymptotic decay rate in type

$\widetilde{\kappa}$	$\widetilde{\beta}$	$\widetilde{\alpha}_z$	q	Δ^*	c	HL
0.014	1.000	-0.002	0.100	0.005	0.001	555
0.014	1.000	-0.005	0.200	0.020	0.005	139
0.010	1.000	-0.002	0.100	0.005	0.001	581
0.010	1.000	-0.004	0.200	0.020	0.005	145
0.005	1.000	-0.000	0.100	0.005	0.001	758
0.005	1.000	-0.002	0.200	0.020	0.004	191

Table 2: Entropies of worst-case models relative to the baseline model. The column Δ^* is relative entropy between the baseline and the worst-case model. The column "HL" is the half-life computed from Chernoff entropy c between the baseline and the worst-case model.

I and type II error probabilities. We use it to indicate statistical distance between the baseline and the worst-case models, with low values of c (high values of HL) indicating that the two models are hard to distinguish.[27]

6.2   Two capital stocks

We now extend investment opportunities by adding a second productive capital that can be used to produce the common consumption good with constant productivity. Under the baseline model, investment $I^{(1)}$ and $I^{(2)}$ affect two capital stocks $K^{(1)}$ and $K^{(2)}$ according to

$$dK_t^{(1)} = K_t^{(1)} \left[\left(\left(\frac{I_t^{(1)}}{K_t^{(1)}} - \frac{\phi_1}{2} \left(\frac{I_t^{(1)}}{K_t^{(1)}} \right)^2 \right) + (.01) \left(\widehat{\alpha}_1 + \widehat{\beta}_1 Z_t \right) \right) dt + (.01)\sigma_1 \cdot dW_t \right]$$

$$dK_t^{(2)} = K_t^{(2)} \left[\left(\left(\frac{I_t^{(2)}}{K_t^{(2)}} - \frac{\phi_2}{2} \left(\frac{I_t^{(2)}}{K_t^{(2)}} \right)^2 \right) + (.01) \left(\widehat{\alpha}_2 + \widehat{\beta}_2 Z_t \right) \right) dt + (.01)\sigma_2 \cdot dW_t \right]$$

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

subject to the aggregate resource constraint

$$C_t + I_t^{(1)} + I_t^{(2)} = \mathcal{A}_1 K_t^{(1)} + \mathcal{A}_2 K_t^{(2)}.$$

---

[27]See Hansen and Sargent (2018) more details about the computation and the interpretation of Chernoff entropy.

The two sectors are identical in their technology parameters $(\mathcal{A}_1, \phi_1, \widehat{\alpha}_1) = (\mathcal{A}_2, \phi_2, \widehat{\alpha}_2)$ and exposures to the Brownian shocks, $\sigma_1 = \sigma_2$, but can differ in exposures to long-run risk, so that $\widehat{\beta}_1 \neq \widehat{\beta}_2$. To study how multiple capital stocks and heterogeneity in their exposures to long-run risk affect decision rules and the worst-case model, we consider two cases:

***ex post* heterogeneity:** two capital stocks possess identical evolution equations but are exposed to idiosyncratic shocks that give rise to a non degenerate distribution of capital *ex post*. This case features a trade-off between diversification and adjustment costs similar to studied by Eberly and Wang (2011). We study how concerns about model misspecification affect this trade-off.

***ex ante* heterogeneity:** two capital stocks differ in their evolution equations, the first capital stock is immune to long-run risk becaase $\widehat{\beta}_1 = 0$, while the second capital stock is exposed to it because $\widehat{\beta}_2 > 0$.

Like the case with a single capital stock, it is convenient to use the investment ratios $D_t^{(1)} \doteq \frac{I_t^{(1)}}{K_t^{(1)}}$ and $D_t^{(2)} \doteq \frac{I_t^{(2)}}{K_t^{(2)}}$ as controls, and the logarithm of aggregate capital, $\log K_t \doteq \log \left(K^{(1)} + K^{(2)} \right)$, and the long-run risk component, Z, as state variables. However, with costly reallocation between the two capital stocks, the distribution of capital becomes an additional endogenous state variable that we express with the ratio[28]

$$R_t \doteq \frac{K_t^{(2)}}{K_t^{(1)} + K_t^{(2)}} = \frac{K_t^{(2)}}{K_t} \in [0, 1].$$

Adjustment costs prevent the household from setting R_t ideally at every instant; instead the household influences its motion at each instant by setting the investment ratios $D_t^{(1)}$ and $D_t^{(2)}$.

The state vector is then $X_t \doteq [\log K_t, \ R_t, \ Z_t]'$. Importantly, the first two of these variables are heteroskedastic with volatilities being[29]

$$\sigma_K(X_t) = (.01)\left(\sigma_1(1 - R_t) + \sigma_2 R_t\right) \quad \text{and} \quad \sigma_R(X_t) = R_t(1 - R_t)(.01)\left(\sigma_2 - \sigma_1\right). \quad (39)$$

In Appendix C, we show that the value function of the robust planner is additively separable

---

[28]Because we solve the problem numerically on a grid, we use a numerically more stable choice for the endogenous state, namely, $L_t \doteq \log K_t^{(2)} - \log K_t^{(1)}$. One can get R_t from L_t using the one-to-one mapping $R_t = \exp(L_t)/(1 + \exp(L_t))$. Using a logarithmic transform allows us to place relatively more grid points near the boundaries of R.

[29]For explicit formulas of the drift and volatility terms of the state vector X, see Appendix C.

in $\log K$, making the optimal investment ratios d_1^* and d_2^* become functions of (R, Z) only. Implied equilibrium consumption is

$$\log C_t = \log K_t + \log \Big([\mathcal{A}_1 - d_1^*(R_t, Z_t)](1 - R_t) + [\mathcal{A}_2 - d_2^*(R_t, Z_t)]R_t \Big). \qquad (40)$$

The worst-case model ceases to share the parametric form of the baseline model, but the distortion h^* remains Markov and depends only on the contemporaneous capital distribution, R_t, and the long-run risk state Z_t.

To induce independent variation in the two capital return processes, we presume that there are three underlying Brownian motions with volatility vectors

$$\sigma_1 = \mathsf{s} \begin{bmatrix} .477 \\ 0 \\ 0 \end{bmatrix}, \quad \sigma_2 = \mathsf{s} \begin{bmatrix} 0 \\ .477 \\ 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} .011\sqrt{.5} \\ .011\sqrt{.5} \\ .025 \end{bmatrix},$$

where s is a scaling parameter that we introduce to control effects of diversification on aggregate consumption volatility. Multiplying the first two entries of σ_z by $\sqrt{.5}$ ensures that, despite having an extra shock, the local volatility of the long-run risk state, $|\sigma_z|$, is unchanged relative to the single capital economy.

To ensure comparability across different settings, in each case we recalibrate the depreciation rate $\widehat{\alpha}_1 = \widehat{\alpha}_2$ and exposure scale parameters s so that the following are true:

1. The unconditional expectation of the drift and volatility terms of equilibrium consumption growth, evaluated under the stationary distribution of the baseline model with robust controls d_1^* and d_2^*, are equal to $\widehat{\alpha}_c$ and $|\sigma_c|$ from (37), respectively.

2. Relative entropies between baseline and worst-case models are kept constant across the different settings.

Regarding the last point, the baseline parameters $(\widehat{\alpha}_z, \widehat{\kappa})$ are always set to the values in (37). We will use different types of worrisome models calibrated to induce worst-case models with a specific level of relative entropy. In this section, we use one that differs only in the intercept term $\widetilde{\alpha}_z$, that is, we keep $\widetilde{\kappa} = \widehat{\kappa}$, $\widetilde{\beta}_1 = \widehat{\beta}_1$, and $\widetilde{\beta}_2 = \widehat{\beta}_2$, and choose $\widetilde{\alpha}_z$ to get $\mathsf{q} = .2$. In section 7 we will use worrisome models that imply state-dependent twisting functions and calibrate them to set $\mathsf{q} = .2$.

Directing the decision maker's concerns toward particular parametric misspecification has relatively small effects on quantities, but has substantial effects on prices. Studying the

two cases demonstrates that the forces in this section are also present in the robust control model of Hansen and Sargent (2001). In section 7, we turn to pricing and investigate effects of a state-dependent ξ.

6.2.1 Symmetric returns to two capital shocks

Consider the case with *ex post* heterogeneity only, that is, set $\widehat{\beta}_1 = \widehat{\beta}_2 = .5$ and $(\mathcal{A}_1, \phi_1, \widehat{\alpha}_1) = (\mathcal{A}_2, \phi_2, \widehat{\alpha}_2) = (\mathcal{A}, \phi, \widehat{\alpha}_k)$ using values from section 6.1.3. Because diversification can reduce aggregate consumption volatility, we increase the exposures $\sigma_1 = \sigma_2$ relative to the single-capital economy by setting $\mathsf{s} = 1.32$, so that average volatility of equilibrium consumption growth equal that estimated from U.S. data.
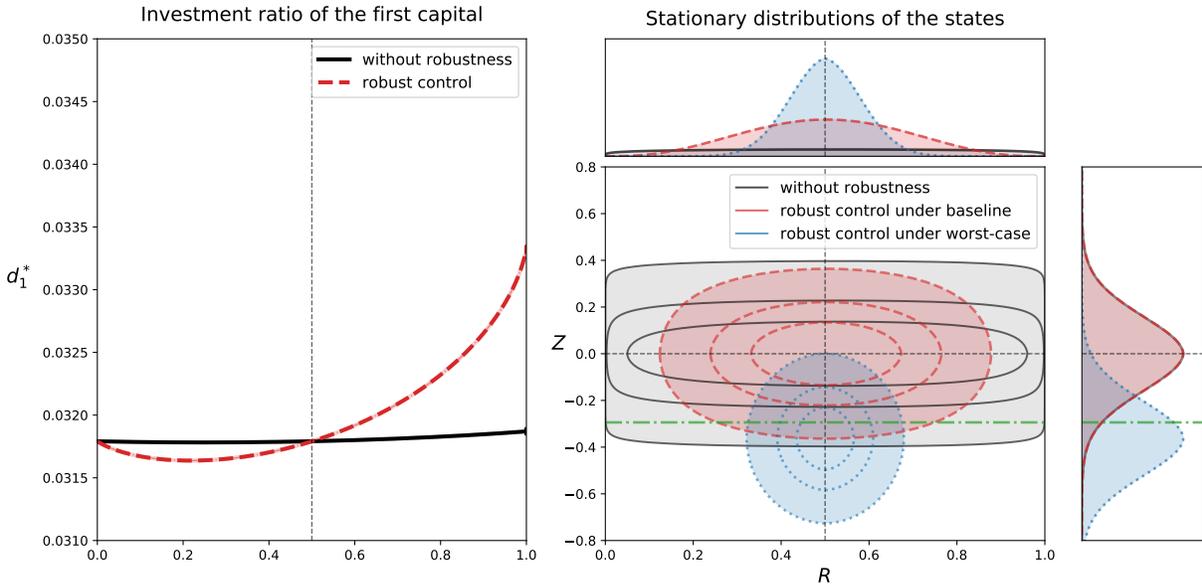

Figure 4: Features of robust controls and the worst-case model with *ex post* heterogeneity. *Left panel:* Optimal investment ratio as a function of the allocation of capital. *Right panel:* stationary distributions (marginals and 90% "confidence sets") for the states (R, Z) under three scenarios: **Solid:** with non-robust controls using the baseline state evolution, **Dashed:** with robust controls using the baseline state evolution, **Dotted**: with robust controls using the worst-case state evolution. Green dash-dotted line denotes the worst-case mean of Z in the single capital economy. Thin dashed lines show median values of the states under the baseline model with robust controls.

Figure 4 summarizes properties of the optimal controls and worst-case model with and without robustness concerns. With *ex ante* identical capital stocks, the risk-averse

35

household seeks a balanced capital distribution with desired level $R^* = .5$ in order to minimize the volatility of aggregate consumption. However, an exactly balanced capital distribution is accompanied by the highest volatility of R, so maintaining this state is costly due to the adjustment costs. Thus, this setting presents a tension between diversification and adjustment costs. The black color in Figure 4 shows the case without robustness, i.e., the investor with expected logarithmic utility analyzed by Eberly and Wang (2011). As the share of the first capital stock becomes low as R becomes high, there is a mild increase in the corresponding investment ratio. Due to symmetry, the analogous effect prevails for the second capital stock, leading to a slight mean reversion in the dynamics of R.

The red color represents the case with robustness. Evidently, a preference for robustness makes the agent's diversification concern become stronger. The robust investment ratio reacts more aggressively to deviations from a balanced capital distribution, so mean reversion of R under the baseline model becomes stronger than without robustness concerns. Because it fears negative mean shifts in the Brownian shocks to Z that reduce perceived expected returns, the household is willing to pay more adjustment costs. With symmetric exposures, $\widehat{\beta}_1 = \widehat{\beta}_2$, long run risk is not diversifiable, so the lower expected returns push the household toward a more diversified portfolio with relatively low variance.

Long run effects of these forces are present on the right panel of Figure 4 depicting the joint and marginal stationary distributions of the states (R, Z). The difference between the black and red distributions illustrates the shrinking effect of robustness concerns on the distribution of capital. The blue color represents the worst-case probability model that supports these "precautious" decisions in an *ex post* Bayesian sense. As mentioned before, the crucial feature of this worst-case model is that it imparts a negative unconditional mean to the long-run risk process.[30]

To study the speed of convergence of the capital distribution, Figure 5 plots the "generalized" impulse response functions (IRFs) proposed by Gallant et al. (1993) and Koop et al. (1996). The impulse response of R for the linear combination of shocks chosen by the vector w is defined as

$$\Phi_R(t, x, \mathsf{w}) \doteq E\left[R_t \mid dW_0 = \mathsf{w}, X_0 = x\right] - E\left[R_t \mid dW_0 = 0, X_0 = x\right].$$

---

[30]The stationary *variance* of R under the worst-case model appears to be smaller than it is under the baseline dynamics, suggesting a slight R-dependence in distortions to the two idiosyncratic shocks that emerges because volatility of $\log K_t$ (hence $\log C_t$) is a function of R (see the expressions in (39)). Due to symmetry, however, these distortions do not affect the stationary *mean* of the ratio R.
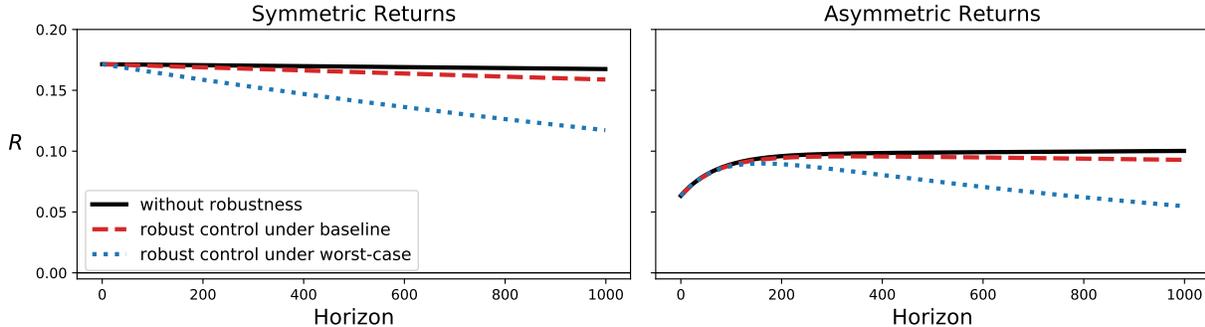
Figure 5: Impulse response functions of R to the second idiosyncratic shock with symmetric (left) and asymmetric returns (right). The size of the shock is $\mathsf{w} = [0, 1, 0]'$. The initial R_0 is set at the median of the stationary distributions of R under the baseline model. In all cases, $Z_0 = \bar{z}$. For explanation of the different colors see the caption of Figure 4

Because our model is nonlinear, the generalized impulse responses depend on the initial state x and shock vector w. We focus on effects of a positive one standard deviation shock to the second capital stock[31] and evaluate Φ_R at the median of the stationary distribution of R. As illustrated by the left panel of Figure 5, idiosyncratic capital shocks have small but very persistent effects on the capital distribution. Since the robust investment rates are more sensitive to R than those without robustness (see left panel of Figure 4), the persistence of R is slightly reduced when the household is concerned about model misspecification. The dependence of Φ_R on the initial capital distribution is very small.[32] This follows from the fact that with homogeneous exposure to long run risk, the dynamics of R and Z are independent of each other. Among other things, this implies that (1) shock to the long run risk state does not affect Φ_R, and (2) the IRFs on the left panel of Figure 5 do not depend on Z_0.

6.2.2 Asymmetric returns to two capital stocks

We now turn to *ex ante* heterogeneity by studying an economy in which the two capital stocks are exposed to long run uncertainty asymmetrically. In particular, we set $\hat{\beta}_1 = 0$ and $\hat{\beta}_2 = .5$ and choose $\mathsf{s} = 1.14$ so that unconditional means of the drift and volatility terms of equilibrium consumption growth are equal to those estimated from U.S. data.

---

[31]Changing the magnitude (or sign) of the shock does not affect the qualitative properties of our results.

[32]Setting R_0 equal to the .1 and .9 deciles of the stationary distribution of R leads to impulse response functions that are indistinguishable from those depicted by the lines in Figure 5.

Studying the drift of R sheds light on how heterogeneous exposure to Z changes outcomes relative to section 6.2.1:

$$\mu_R(X_t) = \left[\left(d_2^* - \frac{\phi_2(d_2^*)^2}{2} + .01\widehat{\alpha}_2 \right) - \left(d_1^* - \frac{\phi_1(d_1^*)^2}{2} + .01\widehat{\alpha}_1 \right) + .01 \left(\widehat{\beta}_2 - \widehat{\beta}_1 \right) Z_t + \right.$$
$$\left. (.01)^2 \left(|\sigma_1|^2(1 - R_t) - |\sigma_2|^2 R_t + (\sigma_1 \cdot \sigma_2)(2R_t - 1) \right) \right] R_t(1 - R_t).$$

In the special case $\widehat{\beta}_1 = \widehat{\beta}_2$, the baseline evolution of R is independent of Z and, as we saw before, the desired capital share is $R^* = .5$ irrespective of Z. With heterogeneous exposure, this is no longer true: movements in the long-run risk state induce fluctuations in the relative attractiveness of the two assets thereby rendering the desired capital share R^* dependent on Z. Effectively, Z becomes an exogenous aggregate shock with the property that negative values ("recession") push the household towards the less exposed first capital, while positive values ("boom") shift the desired portfolio towards the more shock-exposed second capital. Since Z is persistent, the household wants to reallocate capital in order to get closer to the Z-dependent desired capital share. Figure 6 shows that this force increases the unconditional variation of R even without robustness. Nevertheless, because the baseline distribution of Z is symmetric around $\bar{z} = 0$, the unconditional mean of R is still close to .5.

In contrast, because the worrisome model assigns a negative unconditional mean to Z, from the worst-case point of view the less exposed first capital stock appears relatively more attractive on average. Indeed, as can be seen in Figure 6, the robust investment rates induce a stationary capital distribution with mean $\approx .1$. Moreover, just as in section 6.2.1, concerns about misspecifications increase the household's desire to stabilize fluctuations of R around the optimal capital share. This induces strong positive correlation between R and Z that also shapes the IRFs in the right panel of Figure 5. With heterogeneous exposures to Z, the idiosyncratic capital shock has an amplification effect on R relative to the response occurring with symmetric returns. Amplification occurs because the idiosyncratic shock hits Z as well so that and the induced positive response *temporarily* increases the second capital stock's expected excess return; this leads to hump-shaped impulse responses of R.

The worst-case model makes the drift of R become Z dependent. In addition, under the household's *ex post* Bayesian belief, the long-run risk state is no longer exogenous. Instead, as depicted by the left panel of Figure 7, the robust household expects larger falls in Z when holdings more of the Z-exposed asset. The right panel of Figure 7 (accompanied with the right panel of Figure 5) illustrates how such beliefs affect the perceived impulse
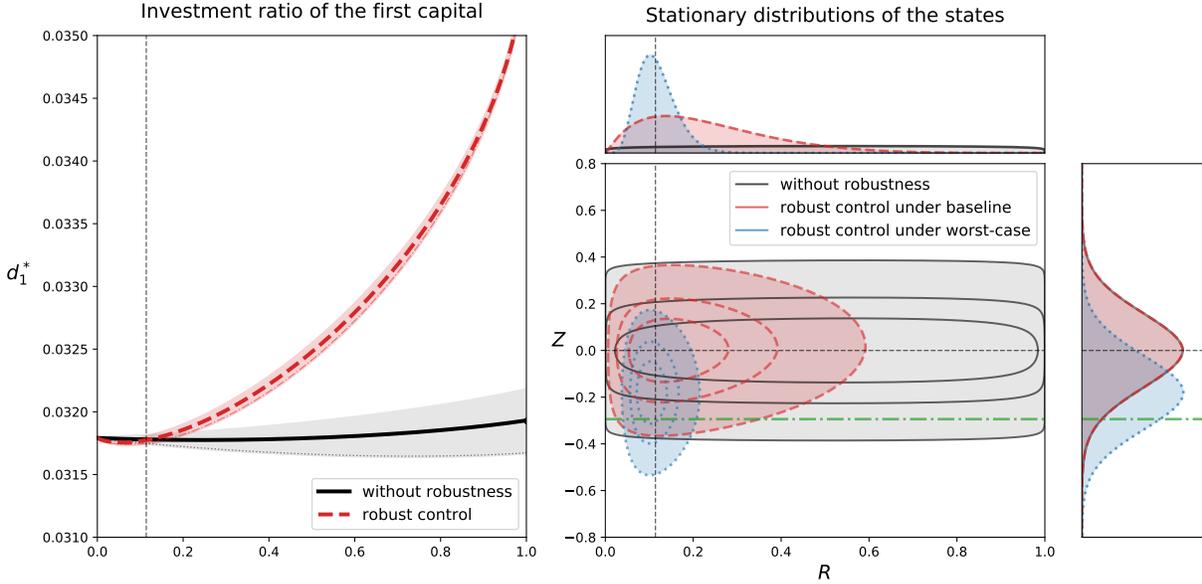
Figure 6: Features of robust controls and the worst-case model with *ex ante* heterogeneity. For explanation of the different panels and colors see the caption of Figure 4. Shaded areas on the left panel show inter-decile ranges for the stationary distributions of Z under the baseline model with the thin dash-dotted line denoting the .9 decile.

response of the long-run risk state to an idiosyncratic capital shock: (1) for short horizons, the positive response appears *less* persistent than it is under the baseline model; and (2) for longer horizons, the effect becomes negative temporarily.

7   Prices of model uncertainty

In this section, we study asset pricing implications of the two state-dependent twisting functions, $\xi^{[\kappa]}$ and $\xi^{[\beta]}$, in the environments of section 6. In the single-capital economy, we saw that the two twisting functions induce noticeably different worst-case models: while $\xi^{[\kappa]}$ tends to impart adverse shifts in the mean of Z, $\xi^{[\beta]}$ focuses relatively more attention on an increased variance of Z. The pricing implications of these worst-case models are substantial in an economy with two capital stocks, where we observed an intricate relationship between the long-run risk state Z and capital distribution R.

First, we consider pricing in the single-capital economy and show that in this special case the representative investor's stochastic discount factor takes a tractable affine form.
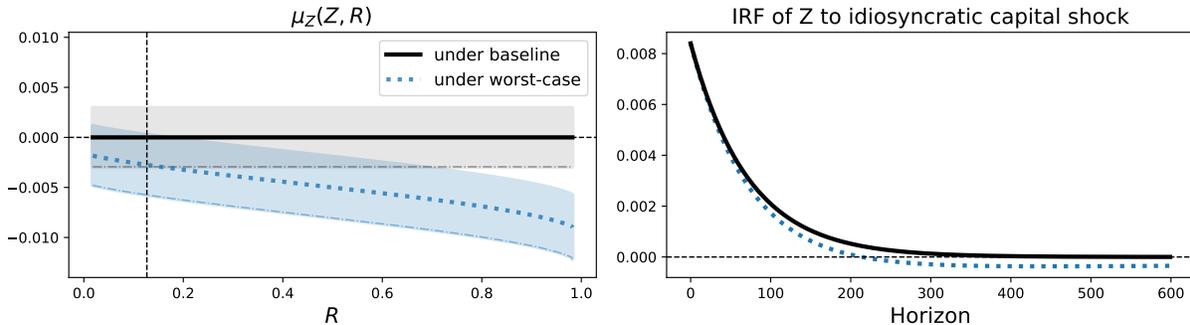
Figure 7: Drift of the long-run risk state (left) and IRFs of Z to the second idiosyncratic shock (right). *Left:* Shaded areas show inter-decile ranges for the stationary distribution of Z under the baseline model with the thin dash-dotted lines denoting the .9 deciles. Vertical dashed line shows the median value of R. *Right:* The size of the shock is $\mathsf{w} = [0, 1, 0]'$. Initial states (R_0, Z_0) are the median values of the stationary distribution under the baseline model with robust controls.

Then we turn to economies with two capital stocks and study how alternative assumptions about the exposure of consumption to long-run risk affect prices of model uncertainty.

7.1 An affine stochastic discount factor

Consider first the economy with a single capital stock. We construct market prices of risk and uncertainty from shadow prices for the planning problem from section 6.1.2. By evaluating the household's marginal rate of substitution at the proposed equilibrium consumption process, we deduce a stochastic discount factor process Λ that obeys

$$
d\Lambda_t = -\underbrace{(.01)\left[100\delta + \widehat{\alpha}_c + \widehat{\beta}Z_t + \sigma_c'H_t^* - \frac{(.01)|\sigma_c|^2}{2}\right]}_{\text{risk free rate}}dt - \left[\underbrace{(.01)\sigma_c}_{\substack{\text{risk price}}} + \underbrace{-H_t^*}_{\substack{\text{uncertainty} \\ \text{price}}}\right]' dW_t.
$$

(41)

The log stochastic discount factor has a linear-quadratic local mean and local variance. The instantaneous interest rate is affine in the state variable as are uncertainty prices. Functional forms of this type have been used extensively in empirical asset pricing applications. A prominent example is Ang and Piazzesi (2003) who estimated a term structure model with an exponential quadratic stochastic discount factor process driven by macroe-

40

conomic state variables. More recently, Piazzesi et al. (2015) assumed an affine form for the likelihood ratio between a well-fitting statistical model of zero-coupon interest rates and a potentially different model assumed to be used by financial experts' to make forecasts about those interest rates. Using survey data on expert's forecasts, they found large and systematic differences between the two models.

Szőke (2018) used our Section 4 formulation to broaden findings of Piazzesi et al. (2015). He posited a representative investor with a log-linear baseline model for consumption growth and inflation and a set $\mathcal{M}(\xi)$ induced by a quadratic twisting function. Exploiting the tractable stochastic discount factor (41), he estimated the parameters of ξ with maximum likelihood using data on zero-coupon yields and macro aggregates. Then, appealing to the *ex post* Bayesian interpretation of the worst-case model, he compared interest rate forecasts under the estimated worst-case model with survey data on experts' forecasts. He found evidence that the discrepancy found by Piazzesi et al. (2015) can partly be attributed to pessimistic adjustments consistent with the tight cross-equation restrictions implied by our refinement of robust control theory.

Minus the local exposures to the Brownian shocks are usually interpreted as local "risk prices", but we reinterpret them as follows. We think of $(.01)\sigma_c$ as risk prices induced by the curvature of log utility and $-H_t^*$ as "uncertainty prices" induced by a representative investor's doubts about different aspects of the baseline model. For the single-capital economy, the uncertainty price takes the form

$$U_t \doteq -H_t^* = -\eta_0^* - \eta_1^* Z_t. \tag{42}$$

In the constant ξ case, $\eta_1^* = 0$, hence U_t is constant; but when ξ is state dependent, η_1^* differs from zero, so the uncertainty prices are time varying and depend linearly on the growth state Z_t. When U depends positively on Z, uncertainty prices are higher in bad than in good times. It is noteworthy that countercyclical uncertainty prices emerge endogenously from a baseline model that excludes stochastic volatility in the underlying consumption risk as an exogenous input under the baseline model.[33]

Figure 8 depicts estimates of the local uncertainty prices associated with the two shocks using the three specifications of ξ presented in Table 1. These estimates were obtained

---

[33]Stochastic volatility models introduce new risks to be priced while also inducing fluctuations in the prices of the "original" risks. The mechanism in this paper simultaneously enhances and induces fluctuations in the uncertainty prices, but it introduces no new sources of risk. Instead, the mechanism features investors' responses to their uncertainty about those risks.
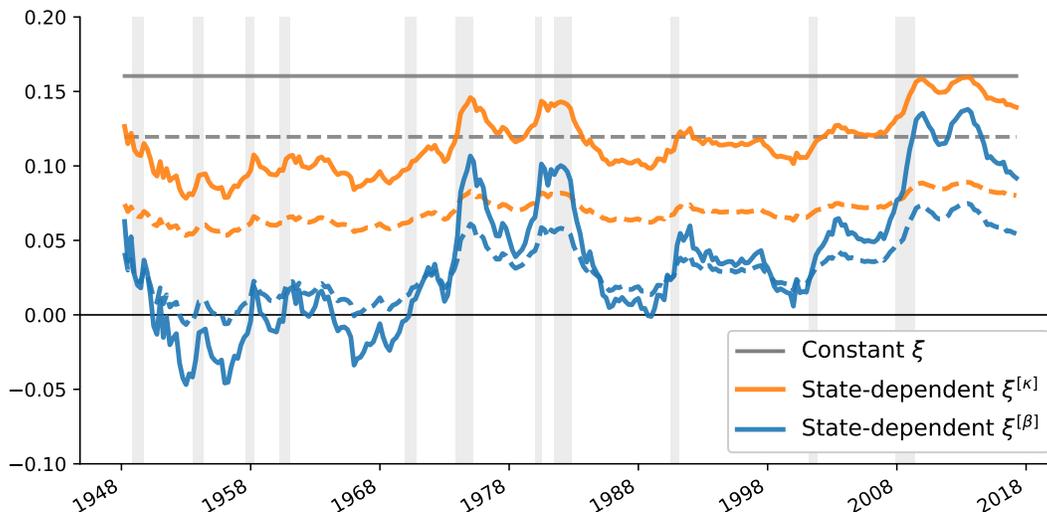
Figure 8: Local uncertainty prices of the direct capital shock (dashed) and the long-run risk shock (solid) for the economy with a single capital stock. The state-dependent twisting functions $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ are induced by worrisome models $(\widetilde{\alpha}_z, \widetilde{\kappa}) = (-.002, .005)$ and $(\widetilde{\alpha}_z, \widetilde{\beta}) = (-.002, 1.215)$, respectively. The constant twisting function is associated with $\widetilde{\alpha}_z = -.005$. In all cases, $\mathsf{q} = .2$. We evaluate $\ell^*(z_0, \theta)$ using $\bar{z}$ as the initial state for the calculations. NBER recessions are shaded.

by first using the Kalman filter to estimate $\widehat{Z}_t = E[Z_t | \Delta \log C_t, \ldots, \Delta \log C_1]$ under the baseline model of section 6.1.4.[34] For each twisting function, we then constructed a bivariate time series model (42) for the two Brownian shocks using the implied worst-case distortion. These series are depicted in Figure 8. Evidently, the estimated uncertainty prices from the state-dependent twisting functions all fluctuate over time in a *countercyclical* manner. These fluctuations arise because both shocks hit the long-run risk state and so affect the "size of the entropy ball" through $\xi(z)$. This makes uncertainty prices increase during times of diminished expected consumption growth. Fluctuations are proportional since they are constructed as an affine function of the single state variable Z.

Consistent with our findings in section 6.1.4, both $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ induce uncertainty prices that are on average lower than those arising from a constant ξ. This is due to the fact that the minimizing agent chooses to distort the baseline intercepts relatively less in order also to distort κ and β. As we saw before, this trade-off is most pronounced for $\xi^{[\beta]}$. Concerns about the exposure parameter β imply uncertainty prices that are relatively low on average but exhibit large fluctuations. In fact, as a result of the exceptionally high

---

[34]We initialized the Kalman filter at the stationary distribution of Z under the baseline model.

long-run consumption growth expectations, the uncertainty price of the growth rate shock becomes slightly negative in the 1950's, indicating that during this period the investor feared positive shocks to Z because of how they relax the minimizing agent's constraint.

7.2 Economies with two capital stocks

Although uncertainty prices do not have an affine representation in our settings with two capital stocks, they still equal $-H^*(R, Z)$, where H^* is a vector of worst-case distortions to the Brownian shocks. In these settings, U_t also depends on the distribution of capital R. To use twisting functions that are comparable with our previous results, we follow the calibration strategy discussed in the beginning of section 6.2.[35] Moreover, to make settings with three shocks comparable with the single capital economy, we construct exposure to an equally weighted sum of the first two shocks scaled to have a unit variance. We compare the uncertainty price of the shock to capital to that for the long-run shock to the growth rate in the section 6.2.1 (symmetric returns) and the section 6.2.2 (asymmetric returns) economies.

Figure 9 displays stationary distributions of the two components of local uncertainty prices under the baseline state dynamics with robust controls. Regarding the twisting functions, we see similar effects as before: in both economies, uncertainty prices implied by $\xi^{[\beta]}$ exhibit lower means and much larger variances than those implied by $\xi^{[\kappa]}$.

In more detail, in the economy with symmetric returns to capital depicted by solid lines, uncertainty prices appear to have very similar properties to those for the single-capital case. This because these prices are independent of the capital distribution R:[36] the dispersion in these prices arises solely from their dependence on Z introduced by the twisting functions. In particular, as we saw in the previous subsection, $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ imply local uncertainty prices that are *decreasing* in Z.

On the other hand, the endogenous positive correlation between the two state variables in the economy with asymmetric capital stocks (dotted lines) makes uncertainty prices depend both the long-run risk state and on the capital distribution. In particular, the

---

[35]We use parameter values that ensure that in all cases $\mathsf{q} = .2$ and that the unconditional means of the drift and volatility of equilibrium consumption growth equal those estimated from U.S. data. With symmetric returns to capital, $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ are constructed from $(\widetilde{\alpha}_z, \widetilde{\kappa}) = (-.0015, .005)$ and $(\widetilde{\alpha}_z, \widetilde{\beta}_1, \widetilde{\beta}_2) = (-.0015, 1.194, 1.184)$, respectively. With asymmetric returns to capital, $\xi^{[\kappa]}$ and $\xi^{[\beta]}$ are constructed from $(\widetilde{\alpha}_z, \widetilde{\kappa}) = (-.0023, .005)$ and $(\widetilde{\alpha}_z, \widetilde{\beta}_1, \widetilde{\beta}_2) = (-.0023, 1.194, 1.194)$, respectively.

[36]More precisely, although the individual idiosyncratic shock prices depend on R due to these shocks' R-dependent contributions to the volatility of aggregate capital σ_K (see (39)), symmetry and our way of considering a composite shock with equal weights that offsets this R-dependence.
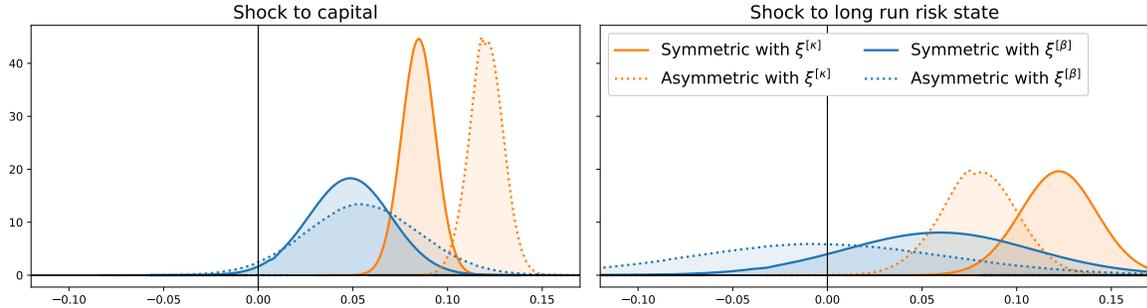
Figure 9: Stationary distributions of local uncertainty prices induced by the two state-dependent twisting functions in economies with symmetric (solid) and asymmetric (dotted) returns to capital. Prices are plotted under the stationary distribution of the baseline model with robust controls. The left panel shows the uncertainty price of an equally weighted sum of the first two idiosyncratic capital shocks scaled to have a unit variance.

amplification effect of Z on shocks to R o observed earlier increases the uncertainty price of "shock to capital" relative to the symmetric economy. In contrast, the price of the growth rate shock is much lower. This follows from the fact that with $\widehat{\beta}_1 \neq \widehat{\beta}_2$, the investor can affect the exposure of her consumption to Z, thereby moderating her fears of negative growth rate shocks. This channel makes negative uncertainty prices of the long-run risk shock much more frequent than what we saw in the economy with a single capital stock or in the economy with symmetric returns.

8   Concluding remarks

As we noted at the outset, Box and Tiao (1977) proposed a statistical procedure for identifying components of the time series that are most persistent. We view our paper as complementary to theirs. Since their procedure is statistical in nature, in practice it results in an imperfect extraction. By posing a representative investor's decision problem and deducing associated shadow prices of uncertainty, we identify components of the dynamic evolution of the macroeconomy associated with the most relevant uncertainties. Our approach highlights how long-term growth uncertainty and its persistence activate a representative investor's aversions to statistical ambiguity and potential model misspecifications. Our analysis thus connects statistical challenges analyzed by Box and Tiao (1977) to decision problems in modern models of the macroeconomy and financial markets.

We close by mentioning again that the dynamic version of the max-min preferences axiomatized by Gilboa and Schmeidler (1989) fails to be dynamically consistent in a sense formalized by Epstein and Schneider (2003). Nevertheless, our section 4 decision problem can be computed recursively using an HJB equation, and the associated preferences satisfy recursively constructed robustness bounds at all subsequent dates along a decision tree. As an alternative, section 5 described a two-player game that differs in subtle ways from the section 4 game that we use in all of the quantitative illustrations in this paper. The computations reported in appendix B confirm that the section 4 and 5 games lead to decision rules that are quantitatively very similar.

Hansen and Sargent (2019, 2018) describe a conceptually and substantively different way to attain dynamic consistency that does not introduce the distinct statistician of section 5; but they do so at the cost of possibly losing admissibility. While the section 4 formulation used throughout this paper is within the class of max-min utility axiomatized by Gilboa and Schmeidler (1989), Hansen and Sargent (2018) instead used the broader class of preferences rationalized by Maccheroni et al. (2006a). In contrast to the findings for the single-capital-stock economy of section 6, Hansen and Sargent (2018) find that a representative investor's concerns about persistence depend substantially in both sign and magnitude on the macroeconomy's growth state.

Appendices

A   Relative entropy neighborhoods reconsidered

We prove the inequality stated in Proposition 3.2. Define the probability measure conditioned on $X_0 = x$ implied by the martingale M^H and construct the product probability measure that includes the time dimension by using the density $\delta \exp(-\delta t)$ over $t \geqslant 0$ for any $\delta > 0$. Call the expectation operator (conditioned on $X_0 = x$) associated with this measure E^H and use it to define the norm

$$\|H\|_H \doteq \left(E^H|H|^2\right)^{1/2}$$

For notational convenience leave the conditioning implicit. Notice that we can express

$$\widehat{\Delta}\left(M^H, 1\right) = \frac{1}{2}\|H\|_H^2 \qquad \text{and} \qquad \widehat{\Delta}\left(M^H, M^{\widehat{S}}\right) = \frac{1}{2}\|H - \widehat{S}\|_H^2$$

Define

$$\tau \doteq \min_H \ \delta \int_0^\infty \exp(-\delta\tau) E\left[M_\tau^H \xi(X_\tau) \mid \mathcal{F}_0\right] d\tau = \min_H \ E^H[\xi(X)].$$

Suppose now that

$$\|\widehat{S}\|_H^2 \leqslant \lambda^2 E^H[\xi(X)] \qquad \text{and} \qquad \|H - \widehat{S}\|_H^2 \leqslant (1-\lambda)^2\tau.$$

By the Triangle Inequality and the concavity of the square root function

$$\|H\|_H \leqslant \|H - \widehat{S}\|_H + \|\widehat{S}\|_H \leqslant \lambda\left(E^H[\xi(X)]\right)^{1/2} + (1-\lambda)\tau^{1/2}$$
$$< \left(\lambda E^H[\xi(X)] + (1-\lambda)\tau\right)^{1/2} \leqslant \left(E^H[\xi(X)]\right)^{1/2}$$

Consequently, $\varrho\left(M^H; \xi\right) < 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

B   Results for the game with a statistician

The following table is a counterpart of Table 1 in section 6.1.4. The only difference is the way the worst-case models are computed. While for Table 1 we use the zero-sum game

formulation in section 4, for Table 3 we use the statistician's game discussed in section 5.

q	$\widetilde{\alpha}_z$	$\widetilde{\kappa}$	$\widetilde{\beta}$	α_c	β	α_z	κ	$\Delta\bar{c}$	m_z	s_z
Baseline										
0.000	0.000	0.014	1.000	0.484	1.000	0.000	0.014	0.000	0.000	0.163
State-dependent $\xi^{[\kappa]}$										
0.100	-0.002	0.010	1.000	0.460	1.005	-0.003	0.013	-0.218	-0.193	0.167
0.200	-0.003	0.010	1.000	0.436	1.005	-0.005	0.013	-0.436	-0.386	0.167
0.100	-0.000	0.005	1.000	0.467	1.028	-0.002	0.010	-0.208	-0.186	0.193
0.200	-0.001	0.005	1.000	0.451	1.028	-0.004	0.010	-0.429	-0.385	0.193
State-dependent $\xi^{[\beta]}$										
0.100	-0.002	0.014	1.157	0.465	1.028	-0.002	0.010	-0.209	-0.185	0.193
0.200	-0.003	0.014	1.199	0.459	1.053	-0.002	0.006	-0.410	-0.366	0.240
0.100	-0.000	0.014	1.164	0.468	1.031	-0.002	0.010	-0.206	-0.184	0.198
0.200	-0.001	0.014	1.208	0.467	1.062	-0.002	0.005	-0.390	-0.352	0.267

Table 3: Worst-case parameter values implied by the section 5 formulation when ξ is defined by (30). The change in the long run consumption growth expectation is denoted by $\Delta\bar{c} \doteq \left(\alpha_c + \frac{\beta\alpha_z}{\kappa}\right) - \left(\widehat{\alpha}_c + \frac{\widehat{\beta}\widehat{\alpha}_z}{\widehat{\kappa}}\right)$. Note that $\left(\widehat{\alpha}_c + \frac{\widehat{\beta}\widehat{\alpha}_z}{\widehat{\kappa}}\right) = .484$. m_z and s_z denote the unconditional mean and standard deviation of Z under the worst-case model.

C   Robust value functions

We provide formulas and discuss methods to compute the value function for the robust control problem in section 6. The state vector is

$$X_t \doteq [\log K_t, \ L_t, \ Z_t - \bar{z}]' \qquad \log K_t \doteq \log\left(K_t^{(1)} + K_t^{(2)}\right) \qquad L_t \doteq \log K_t^{(2)} - \log K_t^{(1)}$$

Define the ratio

$$R_t \doteq \frac{K_t^{(2)}}{K_t^{(1)} + K_t^{(2)}} = \frac{\exp(L_t)}{1 + \exp(L_t)}.$$

The period utility function is

$$\upsilon(X, D) = \delta \log\left((1 - R)\left(\mathcal{A}_1 - D^{(1)}\right) + R\left(\mathcal{A}_2 - D^{(2)}\right)\right) + \delta \log K$$

where we used the resource constraint

$$C_t = \left[(1 - R_t) \left(\mathcal{A}_1 - D_t^{(1)} \right) + R_t \left(\mathcal{A}_2 - D_t^{(2)} \right) \right] K_t.$$

Denote expected capital growth $E_t \left[dK_t^{(i)} / K_t^{(i)} \right]$ for $i = 1, 2$ as

$$\varphi_i \left(D_t^{(i)}, Z_t \right) \doteq D_t^{(i)} - \frac{\phi_i}{2} \left(D_t^{(i)} \right)^2 + (.01) \left(\hat{\alpha}_z + \hat{\beta} Z_t \right)$$

State variables then follow

$$d \log K_t = \left[\varphi_1 (1 - R_t) + \varphi_2 R_t - \frac{(.01)^2 \, |\sigma_1 (1 - R_t) + \sigma_2 R_t|^2}{2} \right] dt + (.01) \left[\sigma_1 (1 - R_t) + \sigma_2 R_t \right] \cdot dW_t$$

$$dL_t = \left[\varphi_2 - \varphi_1 - \frac{(.01)^2}{2} \left(|\sigma_2|^2 - |\sigma_1|^2 \right) \right] dt + (.01) \left[\sigma_2 - \sigma_1 \right] \cdot dW_t$$

$$dZ_t = -\hat{\kappa} \left(Z_t - \bar{z} \right) dt + \sigma_z \cdot dW_t$$

Using Ito's lemma, we can derive the following dynamics for R_t:

$$dR_t = R_t (1 - R_t) \left[\varphi_2 - \varphi_1 + (.01)^2 \left(|\sigma_1|^2 (1 - R_t) - |\sigma_2|^2 R_t + \sigma_1' \sigma_2 (2R_t - 1) \right) \right] dt +$$
$$+ R_t (1 - R_t)(.01) \left[\sigma_2 - \sigma_1 \right] \cdot dW_t.$$

Let σ denote the stacked volatility matrix

$$\sigma(X_t) \doteq \begin{bmatrix} (.01) \left(\sigma_1' (1 - R_t) + \sigma_2' R_t \right) \\ (.01) [\sigma_2 - \sigma_1]' \\ \sigma_z' \end{bmatrix}.$$

We seek a value function $V(X) = \log K + \nu(L, Z)$ that solves the HJB equation

$$0 = \max_{d^{(1)}, d^{(2)}} \min_h \; \delta \log \left((1 - r) \left(\mathcal{A}_1 - d^{(1)} \right) + r \left(\mathcal{A}_2 - d^{(2)} \right) \right) - \delta \nu(l, z) + \frac{\ell}{2} \left[|h|^2 - \xi(z) \right]$$

$$+ \left[\varphi_1 (1 - r) + \varphi_2 r - \frac{(.01)^2 \left[\sigma_1 (1 - r) + \sigma_2 r \right]^2}{2} + (.01) [(1 - r) \sigma_1 + r \sigma_2] \cdot h \right]$$

$$+ \nu_l(l, z) \left[\varphi_2 - \varphi_1 - \frac{(.01)^2}{2} \left(|\sigma_2|^2 - |\sigma_1|^2 \right) + (.01) [\sigma_2 - \sigma_1] \cdot h \right]$$

$$+ \nu_z(l,z) \left[-\widehat{\kappa}(z - \bar{z}) + \sigma_z \cdot h \right] + \frac{1}{2} \mathrm{tr}\left(V_{xx} \sigma \sigma' \right) \tag{43}$$

where

$$\mathrm{tr}\left(V_{xx} \sigma \sigma' \right) = (.01)^2 |\sigma_2 - \sigma_1|^2 \nu_{ll}(l,z) + 2(.01) \left([\sigma_2 - \sigma_1] \cdot \sigma_z \right) \nu_{lz}(l,z) + |\sigma_z|^2 \nu_{zz}(l,z).$$

We assume that a Bellman-Isaacs condition holds so that first-order conditions can be stacked

$$\frac{\delta(1 - r)}{(1 - r)(\mathcal{A}_1 - d^{(1)}(l,z)) + r(\mathcal{A}_2 - d^{(2)}(l,z))} = \left(1 - \phi_1 d^{(1)}(l,z) \right) \left[1 - r - \nu_l(l,z) \right] \tag{44}$$

$$\frac{\delta r}{(1 - r)(\mathcal{A}_1 - d^{(1)}(l,z)) + r(\mathcal{A}_2 - d^{(2)}(l,z))} = \left(1 - \phi_2 d^{(2)}(l,z) \right) \left[r + \nu_l(l,z) \right] \tag{45}$$

$$h(l, z, \ell^*) = -\frac{1}{\ell^*} \sigma'(r) \begin{bmatrix} 1 \\ \nu_l(l,z) \\ \nu_z(l,z) \end{bmatrix}. \tag{46}$$

These equations determine optimal investment-capital ratios $d^{(1)}(l,z)$ and $d^{(2)}(l,z)$, and also the worst-case drift distortion $h(l,z)$. Here ℓ^* is the multiplier that makes the minimizing agent's constraint bind for a given initial (l_0, z_0):

$$\ell^*(l_0, z_0) = \arg\max_\ell \ \nu(l_0, z_0, \ell).$$

C.1 Single capital stock

The "boundaries" $r = 0$ and $r = 1$ can be described in terms of two single-capital economies. The HJB equation becomes

$$0 = \max_{d^{(i)}} \min_{h} \ \delta \log(\mathcal{A}_i - d^{(i)}) - \delta\nu(z) + \frac{\ell}{2} \left[|h|^2 - \xi(z) \right] + \tag{47}$$

$$+ \left[d^{(i)} - \frac{\phi_i}{2} [d^{(i)}]^2 + (.01) \left(\widehat{\alpha}_i + \widehat{\beta}_i \bar{z} + \widehat{\beta}_i(z - \bar{z}) \right) - \frac{(.01)^2 |\sigma_i|^2}{2} + (.01)\sigma_i \cdot h \right] +$$

$$+ \nu_z(z) \left[-\widehat{\kappa}(z - \bar{z}) + \sigma_z \cdot h \right] + \frac{1}{2} \mathrm{tr}\left(V_{xx} \sigma \sigma' \right)$$

with $i = 1$ when $r = 0$ and $i = 2$ when $r = 1$. The optimal choice $d^{(i)}$ is given by (44), namely,

$$d^* = \frac{1}{2}\left[\mathcal{A}_i + \frac{1}{\phi_i} - \sqrt{\left(\frac{1}{\phi_i} - \mathcal{A}_i\right)^2 + \frac{4\delta}{\phi_i}}\right]$$

With $\xi(z) = \xi_0 + 2\xi_1(z - \bar{z}) + \xi_2(z - \bar{z})^2$, the value function $\nu(z)$ is quadratic

$$\nu(z, \ell) = \frac{1}{2}\left[\nu_0(\ell) + 2\nu_1(\ell)(z - \bar{z}) + \nu_2(\ell)(z - \bar{z})^2\right];$$

$\nu_0(\ell)$, $\nu_1(\ell)$, $\nu_2(\ell)$ can be obtained by plugging optimal policies into HJB equation (47) and matching coefficients

$$\nu_2(\ell) = -\ell\left[\frac{\delta + 2\widehat{\kappa} - \sqrt{(\delta + 2\widehat{\kappa})^2 - 4|\sigma_z|^2\xi_2}}{2|\sigma_z|^2}\right] \doteq -\ell\omega_2$$

$$\nu_1(\ell) = \frac{-\ell\xi_1 + (.01)\widehat{\beta}_i + (.01)\omega_2(\sigma_i \cdot \sigma_z)}{\delta + \widehat{\kappa} - \omega_2|\sigma_z|^2}$$

$$\delta\nu_0(\ell) = \frac{1}{\delta}\left\{2\left[\delta\log(\mathcal{A}_i - d^*) + d^* - \frac{\phi_i}{2}(d^*)^2 + (.01)\left(\widehat{\alpha}_i + \widehat{\beta}_i\bar{z}\right)\right]\right.$$
$$\left. - (.01)^2|\sigma_i|^2 - \ell\xi_0 + \nu_2(\ell)|\sigma_z|^2 - \frac{1}{\ell}|(.01)\sigma_i + \sigma_z\nu_1(\ell)|^2\right\}.$$

For a given initial z_0, the multiplier ℓ^* satisfies

$$\ell^*(z_0) = \arg\max_{\ell} \frac{1}{2}\left[\nu_0(\ell) + 2\nu_1(\ell)(z_0 - \bar{z}) + \nu_2(\ell)(z_0 - \bar{z})^2\right].$$

C.2   Numerical Method

For the value function $\nu(l, z)$ in the two-capital-stock problem, we solve HJB equation (43) numerically using the finite difference method with implicit upwind scheme described by Candler (2001) and the Online Appendix of Achdou et al. (2017). We construct a two-dimensional grid for l and z, so that $(l, z) \in [-l^\star, l^\star] \times [-z^\star, z^\star]$. We set $l^\star = 20$ and $z^\star = 1.2$. For a given ℓ, we use the finite difference method to derive the function $\nu(l_i, z_j; \ell)$ on the grid. To initialize iterations, we exploit that the value functions at the $r = 0$ ($l = -\infty$)

and $r = 1$ ($l = \infty$) boundaries are known (with z as their only argument); we extend these function to the whole grid by linearly interpolating between them (over l). Using formulas from Appendix C.1, we can also derive $\ell^*(-\infty, z_0)$ and $\ell^*(\infty, z_0)$. We find the optimal multiplier $\ell^*(l_0, z_0)$ *at a given* (l_0, z_0) by maximizing $\nu(l_0, z_0; \ell)$ with respect to ℓ. A good initial guess for the optimizer is $\ell^0 = \frac{\ell^*(-\infty, z_0) + \ell^*(\infty, z_0)}{2}$.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2017. Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. Working Paper 23732, National Bureau of Economic Research.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1 (1):68–123.

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

Bhandari, Anmol. 2014. Doubts, Asymmetries, and Insurance. University of Minnesota.

Box, George E. P. and George C. Tiao. 1977. A Canonical Analysis of Multiple Time Series. *Biometrika* 355–365.

Brunnermeier, Markus K. and Jonathan A. Parker. 2005. Optimal Expectations. *American Economic Review* 95 (4):1092–1118.

Candler, Graham. 2001. Finite-Difference Methods for Continuous-Time Dynamic Programming. In *Computational Methods for the Study of Dynamic Economies*, edited by Ramon Marimon and Andrew Scott. Cambridge, England: Cambridge University Press.

Chen, Xiahong, Lars Peter Hansen, and Jose Scheinkman. 2009. Nonlinear Principal Components and Long-run Implications of Multivariate Diffusions. *The Annals of Statistics* 37 (6B):4279–4312.

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

Cooper, Russell W. and John C. Haltiwanger. 2006. On the Nature of Capital Adjustment Costs. *The Review of Economic Studies* 73 (3):611–633.

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

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

Fleming, Wendell H. and William M. McEneaney. 1995. Risk-Sensitive Control on an Infinite Time Horizon. *SIAM Journal on Control and Optimization* 33 (6):1881–1915.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen. 1993. Nonlinear Dynamic Structures. *Econometrica* 61 (4):871–907.

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.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*, edited by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, 2672–2680. Curran Associates, Inc.

Hansen, Lars Peter and Thomas J. Sargent. 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.

———. 2018. Macroeconomic Uncertainty Prices when Beliefs are Tenuous. University of Chicago.

———. 2019. Structured Uncertainty and Model Misspecification. Tech. Rep. 3280597, SSRN Working Paper.

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* .

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.

Jiang, Ruiwei and Yongpei Guan. 2018. Risk-Averse Two-Stage Stochastic Program with Distributional Ambiguity. *Operations Research* .

Karantounias, Anastasios G. 2013. Managing Pessimistic Expectations and Fiscal Policy. *Theoretical Economics* 8 (1).

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74 (1):119 – 147.

Liu, Yongchao, Alois Pichler, and Huifu Xu. 2018. Discrete Approximation and Quantification in Distributionally Robust Optimization. *Mathematics of Operations Research* forthcomin.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006a. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica* 74 (6):1447–1498.

———. 2006b. Dynamic Variational Preferences. *Journal of Economic Theory* 128:4–44.

Petersen, Ian R., Matthew R. James, and Paul Dupuis. 2000. Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints. *Automatic Control, IEEE Transactions on* 45 (3):398–412.

Piazzesi, Monika, Juliana Salomao, and Martin Schneider. 2015. Trend and Cycle in Bond Premia. Tech. rep., Stanford University.

Strzalecki, Tomasz. 2011. Axiomatic Foundations of Multiplier Preferences. *Econometrica* 79 (1):47–73.

Szőke, Bálint. 2018. Estimating Robustness. New York University.

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