Generalized Entropy and Model Uncertainty*

Alexander Meyer-Gohde[†][§]

This Version: May 9, 2019

Abstract

I provide a model uncertainty foundation to the power certainty equivalent of Epstein-Zin-Weil risk sensitive preferences (EZ), enabling the analysis of these preferences using detection probabilities (DEPs) and worst case models. This completes the connection between these preferences and the model uncertainty of Hansen and Sargent (2007) (HS) that was previously limited to the special case of unit elasticity of intertemporal substitution. The connection between EZ and HS rests on a powerlike extension of entropy and its associated statistics from Tsallis (1988) and I show that the same additional margin of pessimism that implies this connection can close the gap to the empirical Sharpe ratio in a more general specification. For the specific cases of EZ and HS preferences, I find that calibrations that match detection error probabilities yield comparable asset pricing implications across models. Surprisingly, I find that the low levels of risk aversion with EZ preferences that match asset pricing facts are associated with a high level of model uncertainty in the long run risk environment of Bansal and Yaron (2004).

JEL classification: C61, D80, E03, E44, G12

Keywords: model uncertainty; robust control; recursive preferences; equity premium puzzle; Tsallis entropy

---

*I am thankful to Rhys Bidder and Michael Burda as well as participants of the 2015 CFE and of seminars at the Federal Reserve Bank of San Francisco, the DIW, and the HU Berlin for useful comments, discussions, and suggestions; and am grateful to Nawid Hoshmand, Maximilian Mayer, and Noa Tamir for research assistance. This research was supported by the DFG through the SFB 649 "Economic Risk". Any and all errors are entirely my own.

[†]Goethe-Universität Frankfurt, Chair of Financial Markets and Macroeconomics, Theodor-W.-Adorno-Platz 3, 60629 Frankfurt am Main, Germany; Tel.: +49-69 798 34501; E-Mail: meyer-gohde@econ.uni-frankfurt.de

[§]Institute for Monetary and Financial Stability, Theodor-W.-Adorno-Platz 3, 60629 Frankfurt am Main, Germany

1 Introduction

Tallarini (2000), Barillas, Hansen, and Sargent (2009), Ju and Miao (2012) and others have emphasized the close relationship between model uncertainty preferences[1] and the risk-sensitive preferences of Epstein and Zin (1989) and Weil (1990) (henceforth EZ). This relationship, however, has only been shown formally under a Hansen, Heaton, Lee, and Roussanov's (2007, p. 3975) "special set of assumptions" that align EZ recursive preferences with Hansen and Sargent's (2007) multiplier model uncertainty (henceforth HS) through a logarithmic transformation and a unit elasticity of intertemporal substitution.[2]

Model uncertainty in macroeconomics (see Hansen and Sargent (2001, 2010) and the detailed treatment in the monograph Hansen and Sargent (2007)) places agents in a decision environment riddled with unstructured, Knightian uncertainty that leads to agents forming their decision rules to be robust to a worst case (i.e., welfare minimizing) model. Barillas, Hansen, and Sargent (2009) evoke a critique from Robert E. Lucas, Jr., in their epigraph that although it would be nice to resolve the equity premium puzzle, we need to look past high values of risk aversion to do so and, as an alternative, offer a small amount of model uncertainty as a possible contributor to the resolution of this conflict. This is especially promising as Swanson (2016) shows that a large amount of risk aversion summarizes many of the puzzles in macro-finance literature. Yet Bansal and Yaron's (2004) popular long-run risk resolution of the equity premium, for example, requires that the intertemporal elasticity of substitution differ from one, exactly when the relationship between HS and EZ breaks down.

Specifically, this paper addresses the open question of Backus, Routledge, and Zin (2005, p. 361) as to "whether there's a similar relationship between Kreps-Porteus preferences [(EZ)] with (say) a power certainty equivalent and a powerlike alternative to the entropy constraint [in HS]" by presenting ambiguity averse preferences that contain EZ preferences for arbitrary felicity functions. Furthermore is makes a first step towards assessing the plausibility of EZ preferences as interpreted from an ambiguity perspective using detection error probabilities (henceforth DEP) as proposed by Anderson,

[1] Hansen and Marinacci (2016) summarize the connection between Hansen and Sargent's (2007) multiplier preference approach and other "variational preferences" (Maccheroni, Marinacci, and Rustichini 2006) such as the multiple priors of Gilboa and Schmeidler (1989) and smooth ambiguity of Klibanoff, Marinacci, and Mukerji (2005). Hansen and Sargent (2010) provide a discussion of the link between them. Ju and Miao's (2012) generalized smooth ambiguity preferences nest these variational preferences as special cases from a risk sensitive and ambiguity (vis-a-vis unobservable states) perspective.

[2] Hansen (2005) confirms that the state dependent multiplier of Maenhout (2004) recovers a model uncertainty foundation for EZ preferences, but notes that this holds only in the continuous time case. For the preferences I derive here, I also find that the multiplier on the entropy constraint can be interpreted as state dependent.

Hansen, and Sargent (2003) and Hansen and Sargent (2007) on their resulting worst case models.

I propose a generalization of the statistics of model uncertainty preferences beyond the logarithmic Bolzmann-Gibbs-Shannon measure of entropy to the measure introduced by Tsallis (1988) for nonextensive statistical mechanics in thermodynamics. This results in a generalized exponential certainty equivalent that encapsulates both the exponential and power certainty equivalents of HS and EZ. From the lens of model uncertainty, decreases in risk aversion in EZ's risk-sensitive preferences can be interpreted as a reduction in model uncertainty tempered by an increase in pessimism in the form of an overweighting the probability of pernicious distortions when formulating robust decision rules. This overweighting of events vis-a-vis objective probabilities relates to the choice-theoretic framework of Quiggin (1982) and results here from the generalized alternative entropy measure and its associated subadditivity of probabilities, the latter found also in Gilboa (1987) and Schmiedler (1989). Dow and Werlang (1992) emphasize that expectations formed under probabilities that do not sum to one (i.e., subadditive) reflect both agent's uncertainty and aversion thereto. My generalization and the introduction of an additional form of uncertainty, pessimism in the sense above, is not costless. While in the EZ case, there is only one free parameter to be calibrated using DEPs, the generalized uncertainty case has two parameters, which does not lead to a unique mapping from DEPs.

Applying these preferences to an endowment economy with long run risk following Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2016) and to an otherwise standard RBC production economy in the vein of Tallarini (2000),[3] I find that both HS and the model uncertainty formulation for EZ behave comparably for a given DEP with respect to their maximum Sharpe ratios. In the endowment economy, I find that EZ preferences attribute *high* amount of model uncertainty to a *low* amount of risk aversion in contrast to the results of Barillas, Hansen, and Sargent (2009) under HS preferences. The EZ specification of model uncertainty entails a degree of probability over/underweighting that involves a violation of the law of total probability – the equivalence with HS preferences necessitates an exclusion of this additional distortion on behalf of the econometrician, implying that agents over/underweight events made more/less likely under this distorted probability distribution.

Examining the worst case density associated with the different specifications, I find that agents with model uncertainty fear lower mean consumption/productivity growth. This is broadly consistent

[3]See Bidder and Smith (2012) for a model uncertainty RBC model with investment adjustment costs, variable capital utilization, stochastic volatility, and labor wealth effect sensitive period utility and Ilut and Schneider (2014) for a model uncertainty New Keynesian model with confidence shocks. Backus, Ferriere, and Zin (2015) provide a thorough analysis of variants of a standard RBC model under risk and ambiguity.

with other studies: Barillas, Hansen, and Sargent (2009), Bidder and Smith (2012), Ellison and Sargent (2015), Bidder and Dew-Becker (2016) find that the worst case is associated with lower mean growth. Under all three specifications, agents are concerned that mean of the volatility process in the endowment economy and that of output growth in the production economy might be higher than specified in the approximating model. In terms of the asset pricing implications, both EZ and HS preferences require a DEP of less than 5% to generate Sharpe ratios comparable to the empirical ratio, whereas the generalized model uncertainty can accomplish this with a conservative detection probability of 25%.

The remainder of the paper is as follows. I begin with the generalized measure of entropy in section 2. In section 3, I apply this measure to a general dynamic model, derive conditions that recover both EZ's as well as HS's original model uncertainty framework, assess atemporal risk aversion in all three frameworks, and examine the asset pricing implications of the generalized model uncertainty specification. I then apply the generalized model uncertainty to both an endowment economy and a production economy in section 4 and examine the asset pricing and macroeconomic performance of all three frameworks. Section 5 concludes.

2 Generalized Entropy

Backus, Routledge, and Zin (2005, p. 361) pose the open question of whether there is a powerlike alternative to the entropy constraint that can relate HS model uncertainty to EZ preferences. I will address the first portion of this question in this section and return to the second portion in the next section. HS model uncertainty produces an exponential certainty via standard (logarithmic) relative entropy. I will appeal to the physics literature on statistical mechanics and the generalization of standard Boltzmann-Gibbs-Shannon (logarithmic) measure of entropy introduced by Tsallis (1988).[4] This measure of entropy is associated with a powerlike relative entropy enduced by subadditivity that under/overweights probabilities resulting in a divergence from the law of total probability that I interpret as pessimism. After introducing the basic properties and intuition (the associated over/underweighting of probabilities will play a crucial role in addressing the second portion of the question above), I turn

[4]While Tsallis (2009) documents a wide, interdisciplinary array of empirical applications consistent with this measure, it is not entirely clear whether this is an appropriate measure even in the physics literature whence it emanated – see the critique of Cho (2002) and the discussion in Abe, Rajagopal, Plastino, Latora, Rapisarda, and Robledo (2003). It will be the single parameter in the model uncertainty formulation of EZ preferences and I will discipline it using DEPs. For the two parameter generalized model uncertainty preferences that I will introduce subsequently, only one parameter can be disciplined using DEPs and I take the entropic index as a free parameter, examine its consequences for model uncertainty, and leave its disciplining for further study.

to the associated measure of relative entropy and compare its properties with those of the standard measure of relative entropy or Kullback-Leibler divergence.

The standard Boltzmann-Gibbs-Shannon measure of (negative) entropy

$$S_1(p(x)) \doteq -\int p(x)\ln p(x)dx \tag{1}$$

is used in the context of information theory, see, e.g., Cover and Thomas (1991), as a measure of the expected information content of a realization from the distribution $p(x)$—that is, the expected surprisal or unpredictability of a distribution.[5]

The uniqueness theorems of Shannon and Khinchin[6] provide an axiomatic foundation for the function in (1) and prove that its functional form uniquely satisfies their set of axioms. If their axioms are modified to pseudoadditivity[7] and biased probabilities $p_{q,i} = p_{1,i}^q$, then there exists an unique measure of entropy for all real values of q—the entropic index.

This measure, q entropy introduced by Tsallis (1988), is given by

$$S_q(p(x)) \doteq -\int \left(\frac{1-p(x)^q}{1-q}\right) dx = -\int p(x)^q \ln_q p(x)\, dx \tag{2}$$

where the generalized q-logarithm and its inverse, the generalized q-exponential function, are

$$\ln_q(x) \doteq \frac{x^{1-q}-1}{1-q}, \quad \exp_q(x) \doteq [1+(1-q)x]^{\frac{1}{1-q}} \tag{3}$$

Note that both the foregoing can be extended over their removable singularities at $q = 1$ to give the standard base e logarithm and exponential function as limiting cases. Thus, Tsallis's (1988) entropy recovers (1) as a limiting case, generalizing Boltzmann-Gibbs-Shannon entropy.

The entropic index q can be interpreted as biasing standard probabilities following Tsallis, Mendes, and Plastino (1998), Tsallis (2003), and Tsallis (2009, Ch. 3) and, as noted above, from the generalization of the Shannon-Khinchin uniqueness theorems. Indeed as a probability is positive and less than one, $0 \le p_i \le 1$, $p_i^q \ge p_i$ for $q < 1$ and $p_i^q \le p_i$ for $q > 1$. Thus, under biased probabilities, one expects more (less) surprisal from a realization of random variable when $q < 1$ ($q > 1$). The total probability under the biased probabilities is depicted in figure 1a for a two state system[8] and clearly shows an increase (decrease) in expected surprisal with $q < 1$ ($q > 1$) stemming from an increase (decrease) in total probability. Following Schmiedler (1989) and Dow and Werlang (1992), $q > 1$ can be interpreted

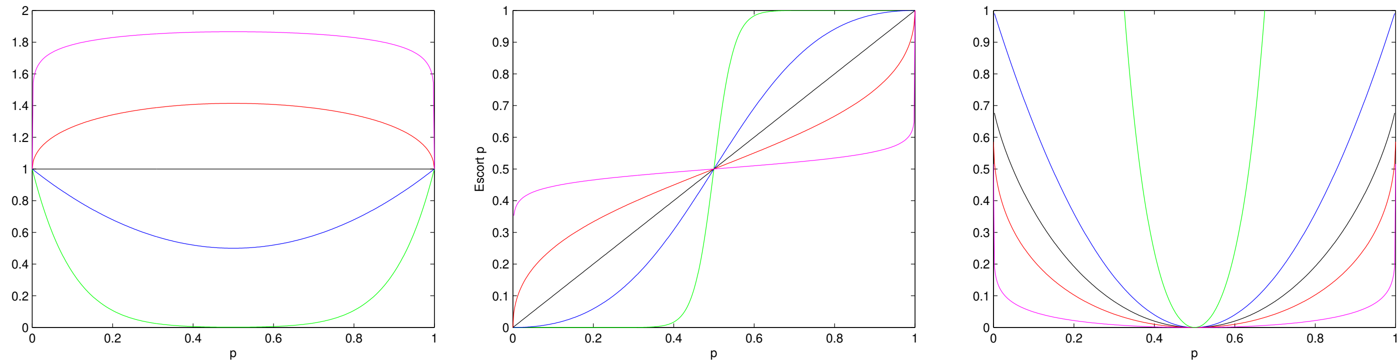
[5]This follows analogously, mathematically and conceptually, with the origin of the term "entropy" as the transformation content in classical thermodynamics and uncertainty or "mixedupness" in statistical mechanics.

[6]See Tsallis (2009, Ch. 2).

[7]For two independent subsystems A and B, pseudoadditivity results in $S_q(A+B) = S_q(A)+S_q(B)+(1-q)S_q(A)S_q(B)$, where standard additivity results in the limiting case $\lim_{q\to 1} S_q(A+B) = S_1(A)+S_1(B)$.

[8]That is, the probability of state one is given by p and that of state two by $1-p$. Of course, the continuous measures above and investigated afterwards are replaced by their discrete counterparts for this example. See Tsallis (2009).

as a situation of uncertainty from the perspective of objective probabilities.



(a) Total Probability (b) Probabilities of Escort Distribution (c) q-Relative Entropy vis-a-vis $p = 0.5$

Figure 1: Biased Probabilites and Relative Entropy
magenta—$q = 0.1$, red—$q = 0.5$, black—$q = 1$, blue—$q = 2$, green—$q = 10$

To preserve the law of total probability, an escort distribution can be defined

$$p_q(x) \doteq \frac{p(x)^q}{\int p(x)^q dx} \tag{4}$$

which normalizes the biased probabilities by the total probability from above. For the two state system, figure 1b plots the probabilities of the escort distribution as a function of the initial probability for different values of the entropic index. As can be seen, the entropic index favors—i.e., increases the probability of—less likely events if $q < 1$ and overweights more likely events if $q > 1$, see also Tsallis, Mendes, and Plastino (1998), Tsallis (2003), and Tsallis (2009, Ch. 3). In contrast to the standard expectations operator with respect to the density $p(x)$

$$E^p[x] \doteq \int x p(x) dx \tag{5}$$

the escort distribution gives a q-generalization of the expectations operator with respect to the density $p(x)$

$$E_q^p[x] \doteq \int x \frac{p(x)^q}{\int p(x)^q dx} dx \tag{6}$$

As shown by Abe and Bagci (2005), this expectation is intricately linked to the functional form of entropy, and this escort expectation leads to a q-generalization of relative entropy that I turn to next.

When comparing two distributions, relative entropy or the Kullback-Leibler divergence of $\tilde{p}(x)$ with respect to the reference distribution $p(x)$

$$I_1(\tilde{p}(x), p(x)) \doteq \int \tilde{p}(x) \ln \frac{\tilde{p}(x)}{p(x)} dx \tag{7}$$

provides a consistent method of discriminating between two probability distributions by quantifying distance between the two distributions.[9] This can be q-generalized following Tsallis (1988), Abe and Bagci (2005), and Tsallis (2009, Ch. 3) as

$$I_q\left(\tilde{p}(x), p(x)\right) \doteq \int p(x) \left(\frac{\tilde{p}(x)}{p(x)}\right)^q \ln_q \left(\frac{\tilde{p}(x)}{p(x)}\right) dx \tag{8}$$

and is positive and convex (both jointly and individually in $\tilde{p}(x)$ and $p(x)$, see Abe and Bagci (2005), for $q > 0$.[10] Figure 3a plots (8) for a two state random variable over possible values of $\tilde{p}$ for differing values of the entropic index with the baseline distribution given by the equiprobable case. Entropy is positive and increasing in the entropic index except when the two distributions match ($\tilde{p} = p = 0.5$). For $q > 1$ ($q < 1$), relative entropy is greater (less) than the Kullback-Leibler divergence.

3 Generalized Multiplier Preferences

In this section I apply the generalized version of entropy and the associated probability measures to the min-max setup of HS. After first presenting and interpreting the general case from two perspectives that share different commonalities with HS's original multiplier preference, I then turn to two special cases. The first recovers HS's original multiplier setup, which serves to link the generalized case to this well studied robustness problem. The second delivers a power certainty equivalent, which presents a model uncertainty foundation for EZ preferences in arbitrary settings (beyond the special unit elasticity of intertemporal substitution and log preference relationship that is already known). Finally, I examine a measure of risk aversion for all three cases (general, HS, and EZ) and explore general asset pricing implications vis-a-vis a decomposition of the pricing kernel.

Here, agents have a preference for robustness; i.e., their decisions are tempered by a fear of model misspecification. This fear is formalized by bounds, derived by a min-max approach, on value functions over a set of models. This set is constrained by penalizing alternative models considered by the agent according to their relative entropy measured vis-a-vis the agent's baseline, or approximating, model. This provides the modeler with a disciplined departure from rational expectations, as agents can have a common approximating model shared with nature, yet demonstrate an ex post divergence by tempering their decisions according to unstructured uncertainty surrounding this model.

Consider a recursive dynamic model where a time-invariant transition density

$$p(x', x, a) \tag{9}$$

[9] Though it is not a metric, as it and the generalization that follows are not symmetric, see Tsallis (1998).

[10] See the online appendix for a discussion of generalized entropy and its derivative in two state settings.

that gives the distribution of the future state, $x' \in X$, conditional on the current state, $x \in X$, and an x measurable control variable, $a \in A$. Thus, the probability distribution over the sequence of states, or model,[11] is determined by

$$\pi(x',x) \doteq p(x',x,a(x)) \tag{10}$$

where the control variable, a, is chosen to maximize lifetime utility expressed recursively by

$$V(x) = \max_{a \in A} u(x,a(x)) + \beta \mathcal{R}(V)(x) \tag{11}$$

where the aggregator $\mathcal{R}(V)(x)$ is derived by considering an agent who entertains a collection of distorted models, each described by a distorted density

$$\tilde{p}(x',x,a(x)) \tag{12}$$

close to the approximating model (10). The likelihood ratio between the models is

$$g(x',x) \doteq \frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))} \tag{13}$$

The decision maker's desire for robustness is formulated as a two player zero sum game, min-max utility, with a minimizing agent, who selects a probability distribution to minimize the decision maker's payoff given her decision or policy function. The decision maker takes this into account when formulating her decision function. I replace HS's Boltzmann-Gibbs-Shannon measure of entropy with the generalized form in (8) from the previous section and calculate distorted expectations using the biased probabilities associated with this generalized measure

$$\begin{aligned} \mathcal{R}(V)(x) &= \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)^q p(x',x,a(x))dx' + \theta I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right) \\ &= \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)^{q-1} \tilde{p}(x',x,a(x))dx' + \theta I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right) \end{aligned} \tag{14}$$

The second term is the generalized relative entropy, conditional on x, of the distorted density to the approximating model. The first term evaluates continuation utility, conditioning on the current state x, under the distorted density weighted by the likelihood ratio between the distorted and approximated densities or Radon-Nikodym derivative $g(x',x)$ to the power of the entropic index q. Thus q is not only the entropic index used in selecting the measure of entropy used to penalize worst case density functions (the second term on the left hand side), but also expresses a form of pessimism. The formulation of Hansen and Sargent (2005) and others with standard Boltzmann-Gibbs-Shannon entropy would set

[11] As I focus on recursive representations – apart from section A.3 of the appendix where I turn to a sequential formulation – I generally refer to a conditional distribution as a model for brevity instead of calling a model a joint distribution over a sequence that such a distribution is an element of.

q to 1, yielding expectations taken with respect to the distorted density $\tilde{p}(x',x,a)$. From the second line when $q>1$, events made more likely under the pernicious distributions they fear are overweighted and those made less likely underweighted when evaluating the expectation of the continuation value under the worst case density (the first term in the second line). Quiggin (1982) deems agents pessimistic if they overweight the probabilities of the worst outcomes on average and if $q>1$ agents will overweight the events in distorted models chosen to minimize their continuation utility. In this sense, I interpret q as a measure of agents' pessimism.

The foregoing zero sum game can be reexpressed in terms of expectations calculated under the escort distribution (see 4) that preserves the law of total probability as[12]

$$\mathcal{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \tilde{E}\left[V(x')\right] + \theta(x) I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right) \tag{15}$$

where

$$\theta(x) = \theta + (q-1)\tilde{E}\left[V(x')\right] \tag{16}$$

and

$$\tilde{E}\left[f(x')\right] \doteq \int f(x') \frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'} dx' \tag{17}$$

The likelihood ratio $g(x',x)$ can apparently be interpreted as a distortion to the probability density of the approximating model where distortions are penalized by their entropy with a nonconstant multiplier.[13] This minimization problem weighs two countervailing forces: the decision maker would like to guard against very painful distortions (those that result in her smallest expected value of her continuation utility, $\tilde{E}[V(x')]$); but a very pernicious distortion that is easy to distinguish, i.e., is far, from her approximating model is considered less likely and adds a large entropy contribution to her objective function ($I_q(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x)$), where $\theta(x')$ weights her concern for closeness. Thus, the decision maker is worried that her misspecification is both pernicious and hard to detect.

For $q>1$, the state dependent multiplier weights future states associated with higher continuation values more strongly; thus, for two competing distorted densities that are equally far from the approximating model, the density associated with a lower continuation value is penalized relatively less. Increasing q increases $(q-1)\tilde{E}[V(x')]$ which tilts the minimizing agent's decision further towards per-

[12] Details can be found in the appendix

[13] See Hansen (2005) and Maenhout (2004) for state dependent multipliers in continuous time formulations. Whereas the multiplier here stems from a rearrangement of (14) that follows from the generalized entropy and its axiomatic foundations, theirs stems from an attempt to restore homotheticity. As I show below, in the homothetic case below, my specification collapses to EZ preferences, just as they recovered the continuous time EZ formulation with their multiplier.

nicious distributions relative to the $q=1$ case. Increasing q, though, also has a countervailing effect: it increases the index in relative entropy, thereby increasing the penalty associated with distorting the probability distribution. Hence changes in q might be interpreted as changes in the shape and not necessarily size of the space of distorted models that agents consider.

Proposition 3.1. *Minimizing Distortion and Risk-Sensitive Operator*

For the generalized entropy measure and multiplier, the minimizing probability distortion is given by

$$g(x',x) = \frac{\exp_q\left(-\frac{1}{\theta}V(x')\right)}{\exp_q\left(-\frac{1}{\theta}\mathcal{R}(V)(x)\right)} = \left(\frac{\theta-(1-q)V(x')}{\theta-(1-q)\mathcal{R}(V)(x)}\right)^{\frac{1}{1-q}} \tag{18}$$

and the risk aggregator, or certainty equivalent, by

$$\mathcal{R}(V)(x) = -\theta \ln_q\left[\int \exp_q\left(-\frac{1}{\theta}V(x')\right)p(x',x,a(x))dx'\right] \tag{19}$$

$$= \frac{\theta - \left[\int(\theta-(1-q)V(x'))^{\frac{1}{1-q}}p(x',x,a(x))dx'\right]^{1-q}}{1-q} \tag{20}$$

Proof. See the Appendix. □

Thus, generalized entropy and its associated distorted probabilities lead to a generalized exponential transformation governed jointly by the entropic index q and static multiplier θ for the risk aggregator. This contrasts with the standard exponential transformation controlled by the static multiplier θ that results from HS's formulation and the power certainty equivalent from EZ.[14] The interpretation of this generalized form follows more readily from the special cases that capture these two specific preferences to which I turn now.

3.1 Equivalence with Hansen-Sargent Multiplier Preferences

In the $q=1$ limit, the standard Kullback-Leibler divergence is used to measure the relative entropy between models, the multiplier in (16) becomes non state dependent ($\lim_{q\to1}\theta(x')=\theta$), and the distorted probabilities of (14) and escort distribution in (15) collapse to the distortion under the induced by the worst-case probability distortion and the model uncertainty specification of HS is recovered

$$\lim_{q\to1}\mathcal{R}(V)(x) = -\theta\ln\left[\int\exp\left(-\frac{1}{\theta}V(x')\right)p(x',x,a(x))dx'\right] \tag{21}$$

with an exponential certainty equivalent following proposition 3.1 and a minimizing distortion

$$g^{HS}(x',x) = \frac{\exp\left(-\frac{1}{\theta}V(x')\right)}{\exp\left(-\frac{1}{\theta}\mathcal{R}(V)(x)\right)} \tag{22}$$

[14]The certainty equivalent in proposition 3.1 is only homogenous (and then linearly) if $\theta=0$. As I show below, $\theta=0$ corresponds to EZ preferences where this homogeneity is well known and $q\to1$ returns the log-exponential twisting of HS that generally does not possess homogeneity.

that tilts the distorted model using the standard exponential function.

This formulation is HS's aggregator,

$$\mathcal{R}(V)(x) \doteq \min_{\substack{\tilde{p}(x',x,a(x))\geq 0 \\ \int \tilde{p}(x',x,a(x))dx'=1}} E^{\tilde{p}}\left[V(x')|x\right] + \theta I_1\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right)$$

$$(23) \qquad = \min_{\substack{\tilde{p}(x',x,a(x))\geq 0 \\ \int \tilde{p}(x',x,a(x))dx'=1}} \int V(x')\tilde{p}(x',x,a(x))dx' + \theta \int \tilde{p}(x',x,a(x)) \ln \frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))} dx'$$

Both the expectation and the relative entropy are with respect to x', conditioning on x. In terms of the likelihood ratio, $g(x',x)$, and the decision maker's approximating model, $p(x',x,a(x))$, the foregoing can be reformulated as

$$\mathcal{R}(V)(x) \doteq \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} E^{g\cdot p}\left[V(x') + \theta \ln\left(g(x',x)\right)\right]$$

$$(24) \qquad = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)p(x',x,a(x))dx' + \theta \int p(x',x,a(x))g(x',x) \ln g(x',x)dx'$$

From the perspective of (14), the formulation here provides decision makers with uncertainty in the modelling sense inasmuch as they entertain deviations from their approximating model. As they use the implied probability distribution of this worst case model, they are not pessimistic in the sense that they do not over- or underweight the ensuing probability distortions.

3.2 Equivalence with EZ Risk Sensitive Preferences

When $\theta = 0$, the multipliers, the Tsallis q relative entropy is used to measure the discrepancy between models, the multiplier in (16) becomes proportional to the expected continuation value, and the distorted probabilities of (14) and escort distribution in (15) maintain a sense of pessimism by overweighting events made more likely under the pernicious distributions they fear and a power certainty equivalent is recovered

$$(25) \qquad \mathcal{R}(V)(x)|_{\theta=0} = \left[\int V(x')^{\frac{1}{1-q}} p(x',x,a(x))dx'\right]^{1-q}$$

To see the equivalence with EZ, recall that their constant elasticity formulation is given by

$$(26) \qquad V(x) = \max_{a\in A}\left[(1-\beta)\,u(x,a(x))^{1-\rho} + \beta\left(\int V(x')^{1-\gamma} p(x',x,a(x))dx'\right)^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}}$$

where $\beta \in (0,1)$ is the discount factor and, with respect to $u(x,a(x))$, ρ is the inverse of the intertemporal elasticity of substitution and γ the coefficient of relative risk aversion.[15] In this case, $\mathcal{R}(V)(x)$ is

[15] Both of these measures are expressed here with respect to the period utility kernel $u(x,a(x))$ and are misnomers if

a power certainty equivalent $E\left[V(x')^{1-\gamma}|x\right]^{\frac{1}{1-\gamma}}$. The preferences in (26) can be rexpressed as[16]

$$\tilde{V}(x) \doteq V(x)^{1-\rho} = \max_{a\in A}\,(1-\beta)\,u(x,a(x))^{1-\rho} + \beta\left(\int \tilde{V}(x')^{\frac{1-\gamma}{1-\rho}} p(x',x,a(x))dx'\right)^{\frac{1-\rho}{1-\gamma}} \tag{27}$$

In this case, $\mathcal{R}(V)(x)$ is a power certainty equivalent $E\left[V(x')^{\frac{1-\gamma}{1-\rho}}|x\right]^{\frac{1-\rho}{1-\gamma}}$.

Backus, Routledge, and Zin (2005, p. 341) restrict $\frac{1}{1-q} < 1$ which translates to $q \in [-\infty, 0] \cup [1, \infty]$. The coefficient of relative risk aversion from (26), γ, is related to q through $\gamma = -\frac{q}{1-q}$ or $\gamma = -\frac{q-\rho}{1-q}$ depending on whether the certainty equivalent is taken to be $\left(\int V(x')^{1-\gamma} p(x',x,a(x))dx'\right)^{\frac{1}{1-\gamma}}$ or $\left(\int \tilde{V}(x')^{\frac{1-\gamma}{1-\rho}} p(x',x,a(x))dx'\right)^{\frac{1-\rho}{1-\gamma}}$. For the former case, values of $q \geq 1$ translate to $\gamma \geq 1$. I will confirm this and provide a measure for risk aversion in the general case in the next section.

Following proposition 3.1, the minimizing distortion associated with EZ preferences is

$$g^{EZW}(x',x) = \left(\frac{V(x')}{\mathcal{R}(V)(x)}\right)^{\frac{1}{1-q}} = \left(\frac{V(x')}{\mathcal{R}(V)(x)}\right)^{1-\gamma} \tag{28}$$

a power tilting instead of the exponential tilting of HS preferences. Having this minimizing distortion will enable me to parameterize their measure of relative risk aversion, γ, in EZ preferences from a model uncertainty perspective using DEPs.

From the perspective of (15), note that the $\theta = 0$ specification of EZ gives

$$\mathcal{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \tilde{E}\left[V(x')\right] + \theta(x) I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right) \tag{29}$$

where $\theta(x) = (q-1)\,\tilde{E}\left[V(x')\right]$ and

$$\tilde{E}\left[f(x')\right] \doteq \int f(x') \frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'} dx' \tag{30}$$

That is the state dependent multiplier becomes proportional to the expected continuation value, analogously to the continuous time result of Hansen (2005) and Maenhout (2004) that recovers a model uncertainty foundation of EZ preferences. To interpret (29), note that the minimizing agent chooses an escort distribution under which the continuation value of the maximizing agent is minimized subject to the entropy constraint. The entropy constraint has two components: the value of the entropy itself and the multiplier. The minimizing agent seeks to find a distribution that is close to the approximating model as measured by (generalized) entropy. Furthermore, the multiplier adds an extra tilt towards pernicious models: for two models equally far from the approximating model (that hence keep the

$u(x,a(x)) \neq C(x)$, where $C(x)$ is the agent's current consumption. See especially, Swanson (2012) and Swanson (2018) for measures of relative risk aversion with alternative period utility kernels and under recursive preferences. I maintain this misnomer here for expositional expediency.

[16] If $\rho > 1$, set $\tilde{V}(x) \doteq -V(x)^{1-\rho}$ and $\tilde{u}(x,a(x))^{1-\rho} \doteq -u(x,a(x))^{1-\rho}$. See Swanson (2018) for details.

value of entropy constant), the minimizing agent would select the more pernicious one (i.e., the one that leads to the smaller expected continuation value), but for ($g > 1$) this more pernicious model is associated with a smaller multiplier ($\theta(x)$), allowing the agent to further twist the worst case away from the approximating model towards even more pernicious distributions.

Yet this tendency towards ever more pernicious distributions is attenuated by the escort distribution itself as is most easily seen from the perspective of (14). Note that here the $\theta = 0$ specification of EZ gives

$$\mathcal{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx' \tag{31}$$

To interpret this, note that if $q = 1$, the minimizing agent would choose an infinitely pernicious distortion $\tilde{p}(x',x,a(x))$ to minimize $\mathcal{R}(V)(x)$. For $q > 1$, this tendency is counterbalanced by the overweighting through q, as making pernicious events more likely increases the value under the integral by increasing $g(x',x) \doteq \frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}$. Recall that q can be interpreted as agents' pessimism: increases in q lead agents to attribute a higher probability to a given pernicious distortion and to more strongly robustify their actions against this distortion, thereby reducing its impact on their continuation value.

While $\theta = 0$ might appear to be a pathological case without an entropy constraint as viewed through (14), the perspective from (15) emphasizes that an entropy constraint remains and that the multiplier is simply linked to the continuation value. Alternatively from the perspective of (14), the preference of EZ can be interpreted as a zero sum game with uncertainty defined via Quiggin's (1982) and Dow and Werlang (1992) subadditive probabilities that lead to overweighting of undesirable outcomes.

In any case, this indicates that the generalized uncertainty preferences conflates two forms of uncertainty, the pessimism associated with overweighting probabilities and the amount of uncertainty through the entropy constraint.

3.3 Risk Aversion

A central goal of this paper is to provide a more general model uncertainty foundation to recursive preferences, especially the preferences of EZ. It is, however, just as instructive to iterate backwards and determine the implied attitudes towards risk provided by model uncertainty. To link the generalized model uncertainty to concepts of risk, I will examine the risk-related properties of the generalized preferences in a static setting.

Abusing notation to minimize clutter by suppressing the dependence on x, the current state, and recycling notation by relabeling the future state, x', with x, the risk aggregator from proposition 3.1 is

$$\mathcal{R}(V) = -\theta \ln_q \left(\int \exp_q \left(-\frac{1}{\theta} V(x) \right) p(x) dx \right) \tag{32}$$

and its minimizing density distortion is

$$g(x) = \frac{\exp_q \left(-\frac{1}{\theta} V(x) \right)}{\exp_q \left(-\frac{1}{\theta} \mathcal{R}(V) \right)} \tag{33}$$

Backus, Routledge, and Zin (2005) calculate the risk aversion with a Taylor expansion of several preferences in a two state equiprobable setup. Accordingly, let there be two states, with outcomes $x_1 = 1+\sigma$ and $x_1 = 1-\sigma$ for positive σ. The certainty equivalent is

$$\mathcal{R}(V) = -\theta \ln_q \left(0.5 \exp_q \left(-\frac{1+\sigma}{\theta} \right) + 0.5 \exp_q \left(-\frac{1-\sigma}{\theta} \right) \right) \tag{34}$$

which I will evaluate locally around $\sigma = 0$ out to second order[17]

$$\mathcal{R}(V) \approx \mathcal{R}(V)\Big|_{\sigma=0} + \frac{\partial \mathcal{R}(V)}{\partial \sigma}\Big|_{\sigma=0} + \frac{1}{2} \frac{\partial^2 \mathcal{R}(V)}{\partial \sigma^2}\Big|_{\sigma=0} = 1 - \frac{q}{\theta+q-1} \frac{\sigma^2}{2} \tag{35}$$

As there is no term linear in σ, risk aversion is second order here. This is not surprising as the generalized exponential risk sensitive preferences are smooth, lacking the kinks responsible for first order risk aversion, see, e.g., Epstein and Zin (1990). The terms

$$\gamma^{\text{Gen. Unc.}} = \frac{q}{\theta+q-1}, \ ,\gamma^{EZ} = \gamma^{\text{Gen. Unc.}}\Big|_{\theta=0} = -\frac{q}{1-q}, \ ,\gamma^{HS} = \gamma^{\text{Gen. Unc.}}\Big|_{\theta=0} = -\frac{q}{1-q} \tag{36}$$

provide measures of risk aversion. In the EZ case, through comparison with (26) this can be seen to correspond to the coefficient of relative risk aversion.[18] In the HS case, this corresponds to Hansen and Sargent (2007) and Tallarini (2000).

Returning to the general case in (36), the measure of risk aversion is decreasing in θ for $q > 0$

$$\frac{\partial \gamma^{\text{Gen. Unc.}}}{\partial \theta} = -\frac{q}{(\theta+q-1)^2}, \ ,\frac{\partial \gamma^{\text{Gen. Unc.}}}{\partial q} = -\frac{1-\theta}{(\theta+q-1)^2} \tag{37}$$

and decreasing in q for θ less than one, but increasing for θ greater than one.

Alternatively for EZ preferences according to (26), the recursive problem can be reexpressed in terms of a linear combination of a period utility function and a power certainty equivalent

$$V(x) = \left[(1-\beta) u(x, a(x))^{1-\rho} + \beta \left(\int V(x')^{1-\gamma} p(x', x, a(x)) dx' \right)^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}} \tag{38}$$

$$\tilde{V}(x) \doteq V(x)^{1-\rho} = (1-\beta) u(x, a(x))^{1-\rho} + \beta \left(\int V(x')^{\frac{1-\gamma}{1-\rho}} p(x', x, a(x)) dx' \right)^{\frac{1-\rho}{1-\gamma}} \tag{39}$$

where ρ is the inverse of the intertemporal elasticity of substitution and γ is the coefficient of relative risk aversion. Comparing with the power certainty equivalent of section 3.2, it follows that $1-q = \frac{1-\rho}{1-\gamma}$

[17] Details of the calculations can be found in the online appendix.
[18] See section 3.2.

or that EZ's coefficient of risk aversion can be mapped through q and ρ as

(40)
$$\gamma^{\text{EZ alternative}} = \frac{q-\rho}{q-1}$$

which is $\frac{-\rho}{q-1}$ less than the measure provided by (36).[19]

3.4 Asset Pricing

A key implication of these preferences (HS, EZ, and generalized uncertainty) is their asset pricing implications as implied by their pricing kernels. I will decompose these kernels (or stochastic discount factors) following the literature into the market prices of risk and uncertainty and a third component – the market price of pessimism – that arises through the entropic index q. To further understand the mechanisms at work, I will explore their general implications for the equity premium puzzle intuitively via log-normal approximations.[20]

Consider a household seeking to maximize the following

(41)
$$V_t = u(C_t, \bullet) - \beta\theta \ln_q\left(E_t\left[\exp_q\left(-\frac{1}{\theta}V_{t+1}\right)\right]\right)$$

where V_t is the household's lifetime discounted utility, $u(C_t, \bullet)$ its period utility function that depends at least on consumption C_t, and $\beta \in (0,1)$ the household's subjective discount factor.

The likelihood ratio between the distorted and approximating models is given by

(42)
$$g_{t+1} = \frac{\exp_q\left(-\frac{1}{\theta}V_{t+1}\right)}{E_t\left[\exp_q\left(-\frac{1}{\theta}V_{t+1}\right)\right]}$$

The household's stochastic discount factor or pricing kernel is given by

(43)
$$M_{t+1} \doteq \frac{\partial V_t/\partial C_{t+1}}{\partial V_t/\partial C_t} = \frac{\frac{\partial V_t}{\partial V_{t+1}}\frac{\partial V_{t+1}}{\partial C_{t+1}}}{\frac{\partial V_t}{\partial C_t}}$$

with

(44)
$$\frac{\partial V_t}{\partial C_t} = u_C(C_t, \bullet),\ \frac{\partial V_{t+1}}{\partial C_{t+1}} = u_C(C_{t+1}, \bullet)$$

and

(45)
$$\frac{\partial V_t}{\partial V_{t+1}} = \beta\left(\frac{\exp_q\left\{-\frac{1}{\theta}V_{t+1}\right\}}{E_t\left[\exp_q\left\{-\frac{1}{\theta}V_{t+1}\right\}\right]}\right)^q = \beta g_{t+1}^q = \beta g_{t+1} g_{t+1}^{q-1}$$

combining yields the final form of the pricing kernel

(46)
$$M_{t+1} = \beta\frac{u_C(C_{t+1}, \bullet)}{u_C(C_t, \bullet)} g_{t+1} g_{t+1}^{q-1} = \Lambda_{t+1}^R \Lambda_{t+1}^U \Lambda_{t+1}^P$$

where $\Lambda_{t+1}^R \doteq \beta\frac{u_C(C_{t+1},\bullet)}{u_C(C_t,\bullet)}$ is the stochastic discount factor under expected utility ($\theta = \infty$), $\Lambda_{t+1}^U \doteq g_{t+1}$

[19] Additional measures of risk aversion based on wealth gambles following Swanson (2018) can be found in the online appendix.

[20] In the section that follows, I will use nonlinear methods and examine their implications for asset prices and the macroeconomy in two specific models.

is the change of measure under the distorted model, and $\Lambda^P_{t+1} \doteq g^{q-1}_{t+1}$ captures the direct effect[21] of the entropic index.

Note that if $q = 1$, Λ^P_{t+1} is equal to unity and the model uncertainty concerns collapse to HS (see section 3.1 above). For $q > 1$, agents overweight (underweight) states that have become more (less) likely under distorted models when pricing assets, embedding a form of pessimism into a non-unity Λ^P_{t+1}. Thus, along with Hansen and Sargent's (2007), Bidder and Smith's (2012), and others' interpretation of $std_t\left(\Lambda^R_{t+1}\right)/E_t\left[\Lambda^R_{t+1}\right]$ and $std_t\left(\Lambda^U_{t+1}\right)$ as the market prices of risk and model uncertainty, respectively, I interpret $std_t\left(\Lambda^P_{t+1}\right)/E_t\left[\Lambda^P_{t+1}\right]$ as the market price of pessimism.

For EZ's power certainty equivalent, $\theta = 0$ (see section 3.2 above) all three components of the stochastic discount factor remain. As risk aversion is related inversely to q in this case, see section 3.3, an increase in risk aversion is associated with a decrease in pessimism, as Λ^P_{t+1} approaches unity.

To understand the consequences of the different preference specifications intuitively, it is instructive to consider the equity premium. The fundamental asset pricing equation gives

$$1 = E_t\left[M_{t+1}R_{t+1}\right], \quad 1 = E_t\left[M_{t+1}\right]R^f_t \tag{47}$$

where R_t is the return on a risky asset and R^f_t is the return on a riskless asset. Log-normal approximations of the foregoing, where lowercases indicate logs and σ's (co)variances, yield

$$E\left[r_{t+1} - r^f_t\right] = -\frac{1}{2}\sigma_r - \sigma_{rm} \tag{48}$$

Applying the decomposition in (46) yields the following expression for the equity premium

$$E\left[r_{t+1} - r^f_t\right] = -\left(\frac{1}{2}\sigma_r + \sigma_{r\lambda^R} + \sigma_{r\lambda^U} + \sigma_{r\lambda^P}\right) = -\left(\frac{1}{2}\sigma_r + \sigma_{r\lambda^R} + q\sigma_{r\lambda^U}\right) \tag{49}$$

The last equality follows as $\sigma_{r\lambda^P} = (q-1)\sigma_{r\lambda^U}$, and this would seem to show that the additional pessimism term is begging the question: The market price of pessimism simply scales up the impact of the market price of uncertainty through q, so clearly we can simply increase q to match the equity premium for a negative covariance $\sigma_{r\lambda^U}$. However, changing q, also changes the market price of uncertainty itself by altering the certainty equivalent in proposition 3.1. Thus if an increase in q is associated with a sufficiently large increase in the covariance $\sigma_{r\lambda^U}$, then an increase in q can lead to an increase in the composite $q\sigma_{r\lambda^U}$ thereby decreasing the equity premium. Indeed a general conclusion from the literature[22] is that increasing risk aversion with EZ will increase the equity premium. From section 3.3, risk aversion is however *decreasing* in q. Thus, increasing q does not simply scale up the effect of the market price of uncertainty and indeed the ensuing change in the market price of

[21] The entropic index, as was shown above, enters into the change of measure g.

[22] See, e.g., Table 1 from Weil (1989) or the review in section 1.3 of Donaldson and Mehra (2008).

uncertainty, at least in the EZ case, moves the equity premium in the *opposite* direction of q.

4 Business Cycles, Asset Prices, and Model Uncertainty

Here I apply the general model uncertainty framework I have derived above to two models of asset pricing. The first is an endowment economy with long-run risk, typically examined in finance with EZ preferences and my generalized model uncertainty allows me to assess the model (and especially its conspicuously low level of risk aversion) in terms of DEPs - a measure of the degree of model uncertainty entertained by agents. I find that the exponential certainty equivalent of HS preferences and the power certainty equivalent of EZ preferences imply similar maximum Sharpe ratios for a given DEP, despite their associated worst case models differing substantially. Surprisingly, the low level of risk aversion under EZ preferences is actually associated with a high degree of model uncertainty. The generalized uncertainty is able to rectify a low degree of model uncertainty with large Sharpe ratios by increasing the pessimism that differentiates EZ from HS preferences.

I then turn to a production economy - that of Tallarini (2000) - and examine the asset pricing and macroeconomic consequences of the different forms of model uncertainty. Unlike the endowment economy, households can endogenously determine their consumption and savings plans. I find that all three preference forms are again able to reconcile asset pricing facts in this environment. EZ and HS preferences again demonstrate a close relationship in terms of maximum Sharpe ratios for given levels of model uncertainty, though they both require substantial amounts of uncertainty to match empirical equity premia. The generalized preference approach with heightened pessimism is, however, able to match the empirical Sharpe ratio with a conservative level of model uncertainty.

4.1 Data and Solution Method

The calibration of the model will focus on matching the first two moments of key macroeconomic indicators and the Sharpe ratio (see table 1) for the U.S. post war period.

The empirical Sharpe ratio and the its theoretical maximum for a given model $\frac{std(m_{t+1})}{E[m_{t+1}]}$ that measures the excess return the household demands for bearing an additional unit of risk can be related through a Cauchy-Schwarz inequality and the fundamental asset pricing equations (47) as

$$\frac{\left|E\left[R_{t+1}-R_t^f\right]\right|}{std\left(R_{t+1}-R_t^f\right)} \leq \frac{std\left(m_{t+1}\right)}{E\left[m_{t+1}\right]} \tag{50}$$

Business Cycle Data

Variable	Mean	Std. Dev. %	Relative Std. Dev.	Autocorrelations 1	2	3	Cross Corr. w$\Delta \ln Y_t$
$\Delta \ln Y_t$	0.004	0.991	1.000	0.380	0.266	0.045	1.000
$\Delta \ln C_t$	0.005	0.566	0.571	0.255	0.201	0.069	0.531

Asset Return Data

Return	Mean	Std. Dev.		
R	2.13	8.26		
R^f	0.26	0.63		
rp	1.87	8.27	Sharpe Ratio	0.2261

All returns are measured as net real quarterly percentage returns. See the online appendix for details on the series used.

Table 1: Empirical Macroeconomic and Asset Pricing Moments, 1948:2-2012:4

with the Sharpe ratio on the left hand side being empirically observable and given in the lower half of table 1. Hansen and Jagannathan (1997) extend the maximum Sharpe ratio point restriction on pricing kernels of the right hand side to a parabola inside which pairs of $std(m_{t+1})$ and $E[m_{t+1}]$ must reside to be consistent with (a vector) of risky assets and the riskless bond.

In calibrating the parameters q and θ, I will proceed in the spirit of Hansen and Sargent (2007) and examine DEPs. A value of 0.5 for the DEP indicates that the two models (approximating and worst-case) are indistinguishable, as the agents have a fifty-fifty chance of correctly identifying the model used to generate the simulations. Barillas, Hansen, and Sargent (2009) argue for a DEP of between 0.15 and 0.2 as lower bound. In general, I will take a conservative perspective and target a DEP of 0.25. For the HS certainty equivalent, this follows Hansen and Sargent (2007) exactly and the DEP pins down the multiplier on the entropy constraint θ. For the EZ specification, $\theta = 0$ and targeting a DEP pins down the only free parameter q, the entropic index. For the generalized specification, I have two free parameters, q and θ, so the DEP will not pin the pair down uniquely. I will focus on the pair that also matches the edge of the Hansen and Jagannathan (1997) bound, which can be interpreted as providing the entropic index q that matches the equity premium for a θ that yields a DEP of 25%.

One can object to the fact the econometrician uses the actual likelihood ratio g when calculating the DEPs while the agents in the model overweight g^q pernicious distributions when forming expectations as perhaps overstating the results for the generalized model uncertainty case. But note that this objection would then also apply to the EZ specification that operates solely through q: the approximate equivalence with HS's specification in regards to the maximum Sharpe ratio and DEPs I will depict in figures 5 and 6 rests likewise on this discord between the measures of the agents and the econometrician.

To solve the models, I use perturbation following Bidder and Smith (2012), but use the nonlinear moving average policy function or pruning of Lan and Meyer-Gohde (2013c) to maintain the stability of the model under nonlinearity.[23] As proposed by Bidder and Smith (2012), I first generate simulations (the length of which will match the length of the post war U.S. data series used) using the perturbation solution of the model and then perform a likelihood ratio test over the agents' approximating model p and the distorted model $\tilde{p}$ using the change of measure directly. Second, I generate simulations from the distorted model using importance sampling and perform a symmetrical likelihood test to calculate the DEPs.[24]

4.2 Endowment Economy with Long Run Risk

In this section, I apply the generalized entropy constraint to an endowment economy with long-run consumption risk. Bansal and Yaron (2004) use EZ preferences and show that long run risk can resolve the equity premium puzzle with reasonable values of risk aversion. Specifically, I follow Bansal, Kiku, and Yaron's (2016) specification that allows for correlation between the consumption and dividend processes and take their quarterly calibration as a starting point for the analysis.

The economy is populated by an infinitely lived household. Under the recursive preferences from EZ as specified in (26)

$$V_t = \left((1-\beta) C_t^{1-\rho} + \beta E_t \left[V_{t+1}^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right)^{\frac{1}{1-\rho}} \to \tilde{V}_t = (1-\beta) C_t^{1-\rho} + \beta E_t \left[\tilde{V}_{t+1}^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} \tag{51}$$

The transformation on the right is discussed surrounding equation (27) and further details can be found in Swanson (2018). Consumption will be nonstationary and exogenously given.[25]

[23] See Lan and Meyer-Gohde (2013b) for a comparison of alternate, so-called pruning, algorithms to deliver this stability. An additional advantage to using a nonlinear moving average or pruning algorithm is that closed-form theoretical moments are available, see Lan and Meyer-Gohde (2013a) and Andreasen, Fernández-Villaverde, and Rubio-Ramírez (2017), which were used to initialize the particle filters.

[24] As the particle filter with 1,000,000 particles still suffers from sampling variation when calculating the likelihood tests for high and low DEPs, I follow Bidder and Dew-Becker (2016) and Bidder and Smith (2018) and calculate the log-likelihood ratios directly from the perturbation approximated changes of measure g. This eliminates the sampling variation and computational burden associated with the particle filter, but assumes that the entire state vector is observable when comparing models. I found that this only slightly reduced the DEPs compared with calculations using the particle filter conditional on a subset of the models' variables (i.e., consumption).

[25] Note from assumption 16, the multiplier on the entropy constraint is a function of continuation utility: expressing the problem with the generalized risk sensitive operator before stationarizing would imply a trend in the multiplier. Essentially, I am stationarizing the multiplier along with the rest of the model by applying the operator afterwards. This follows from Cagetti, Hansen, Sargent, and Williams (2002), who use a stochastically discounted entropy constraint. Details can be

The household's lifetime utility function is expressed recursively as

$$v_t = (1-\beta)\, c_t^{1-\rho} + \beta \mathcal{R}\left(v_{t+1} w_{t+1}^{1-\rho}\right) = (1-\beta)\, c_t^{1-\rho} - \beta\theta \ln_q \left\{ E_t \left[\exp_q \left\{ -\frac{1}{\theta} v_{t+1} w_{t+1}^{1-\rho} \right\} \right] \right\} \tag{52}$$

with $\beta \in (0,1)$ the discount factor, v_t the value function at the optimum, $c_t \doteq \frac{C_t}{W_t}$ detrendend consumption, W_t the detrending factor, and $w_t \doteq \frac{W_t}{W_{t-1}}$ is gross growth rate. The fundamental asset pricing equations in (47) price the return on risky asset $R_t \doteq exp(\Delta d_t) * (1 + PD_t)/PD_{t-1}$ that pays out dividends that grow at the rate Δd_t with PD_t the price dividend ratio and the R_t^f is the risk-free rate. The stochastic discount factor or pricing kernel (see section 3.4), M_{t+1}, is given by[26]

$$M_{t+1} = \beta e^{-\rho \Delta c_{t+1}} \left(\frac{\exp_q \left\{ -\frac{1}{\theta} v_{t+1} e^{(1-\rho)\Delta c_{t+1}} \right\}}{E_t \left[\exp_q \left\{ -\frac{1}{\theta} v_{t+1} e^{(1-\rho)\Delta c_{t+1}} \right\} \right]} \right)^q \tag{53}$$

Consumption and dividend growth are given by the following long-run risk specification

$$\Delta c_t = \mu_c + x_{t-1} + \sigma_{t-1} \varepsilon_t^c \tag{54}$$

$$\Delta d_t = \mu_d + \phi x_{t-1} + \sigma_{t-1} \sigma_d \left[(1-\rho_{dc})^{\frac{1}{2}} \varepsilon_t^d + \rho_{dc} \varepsilon_t^c \right] \tag{55}$$

$$x_t = \rho_x x_{t-1} + \sigma_{t-1} \sigma_x \varepsilon_t^x \tag{56}$$

$$\sigma_t^2 = \bar{\sigma}^2 + \rho_\sigma \left(\sigma_{t-1}^2 - \bar{\sigma}^2 \right) + \sigma_\sigma \varepsilon_t^\sigma \tag{57}$$

where x_t is the long run growth process and ε_t^c, ε_t^d, and ε_t^x are iid standard normals.

The parameterization follows Bansal, Kiku, and Yaron's (2016) quarterly, post-war estimates, but is calibrated to match the standard deviation and first autocorrelation of ΔC and the means and standard deviations of R^f and rp in table 1.[27] Comparing (51) with (25), the values of q and ρ (the entropic index and inverse IES respectively) are related to γ via $\gamma = (q-\rho)/(q-1)$,[28] which gives a value of relative risk aversion equal to 5.72 for the calibrated values of $q = 1.119$ and $\rho = 0.438$. I then calculate the DEP associated with this calibration, which is equal to 4.25%, far below the conservative measure of 25% that I adopted above. For the HS specification, I set $q = 1$ and calculate θ to match this DEP (here $\theta = 5.71$). For the generalized model uncertainty case, I find the q and θ pair (here $q = 2$ and $\theta = -6.29$)[29] that generates a DEP of 25% and a maximum Sharpe ratio of roughly .184 on the edge of the Hansen and Jagannathan (1997) bounds.

found in the online appendix.

[26] See the online appendix for a detailed derivation.

[27] See the online appendix for the complete parameterization.

[28] See the discussion following (25).

[29] Though a negative θ might seem surprising, recall from section 3 that the multiplier on the entropy is a compound that depends on θ as well as the continuation value. In other words, a negative θ need not imply a negative $\theta(x)$.

4.2.1 Long-Run Risk and the Worst Case

	Approximating Model Common to all specifications				Generalized Uncertainty $q = 2$, $\theta = -6.29 \rightarrow$DEP=25%			
Variable	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t
$\Delta \ln c_t$	0.500	0.566	0.255	0.503	0.321	0.556	0.226	0.4711
x_t	0	0.287	0.990	1	-0.177	0.266	0.989	1
$\Delta \ln d_t$	0.200	3	0.114	0.336	-0.364	2.705	0.099	0.311
σ_t	2.370E-03	1.243E-3	0.995	0.000	2.386E-03	1.234E-3	0.994	-0.096

	EZ Preferences $\theta = 0$, $q = 1.119 \rightarrow$DEP=4.25%				HS Preferences $q = 1$, $\theta = 5.71 \rightarrow$DEP=4.25%			
Variable	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t
$\Delta \ln c_t$	0.171	0.552	0.213	0.458	0.301	0.565	0.235	0.482
x_t	-0.324	0.258	0.988	1	-0.194	0.278	0.990	1
$\Delta \ln d_t$	-0.843	2.700	0.093	0.304	-0.428	2.752	0.101	0.317
σ_t	2.388E-03	1.237E-3	0.994	-0.223	2.450E-03	1.189E-3	0.994	-0.117

Table 2: Model Moments for the Approximating and Worst-Case Models of the Endowment Economy

I begin first with the moments of the models and those of their worst cases characterized by g_{t+1}, contained in table 2. [30] Under all three specifications, agents fear a worst case associated with lower mean consumption growth, consistent with the literature on HS preferences (see, e.g., Barillas, Hansen, and Sargent (2009)). Likewise all three models posit a worst case model with a higher mean volatility of exogenous innovations. This does not translate to higher volatility in worst case consumption growth due to its reduced correlation with the long run risk process. Comparing HS and EZ worst cases, both with the same DEP of 4.25%, the reduction in mean consumption growth is larger in the latter. This is compensated for under HS preferences with higher relative volatility and autocorrelation in consumption growth, which stems from the larger role they assign to the long run risk process in the worst case relative to EZ. The generalized uncertainty preferences introduced above with $q = 2$ (the reason for this value will be apparent in when I turn to asset pricing in the next section), likewise induces fears of heightened volatility and reduced means. In contrast to the HS and EZ preferences that were calibrated to match the 4.25% DEP commensurate with Bansal, Kiku, and Yaron's (2016) calibration of the EZ preferences, the generalized uncertainty preferences were set to induce a DEP of

[30] Moments for the worst cases of EZ and my generalized preferences under the g^q_{t+1} measure can be found in the online appendix. The differences are quantitative and not qualitative to those presented here.

25%. Hence, it is not surprising that the deviations from the approximating model are not as large in this case.

4.2.2 Asset Pricing Implications

Figure 2 traces out the maximum Sharpe ratios and the associated DEPs for HS (red) and EZ (blue) preferences. As discussed in section 3.4 the pricing kernel for $q \neq 1$ multiplies the market price of risk with $g_{t+1}g_{t+1}^{q-1}$, where g_{t+1} is the change of measure induced by the worst case (in the context of asset pricing, the market price of model uncertainty) and the term g_{t+1}^{q-1} that captures the over-/underweighting of outcomes made more/less likely under distorted densities (which captures pessimism in the sense of Quiggin (1982)). This is a discord here in the sense that an econometrician might be calculating the DEPs implied by g_{t+1}, whereas the agent could also be thought of as possessing concerns about the potentially larger (for $q > 1$) distortion implied by $g_{t+1}g_{t+1}^{q-1} = g_{t+1}^{q}$. To investigate this, figure 2 also contains the maximum Sharpe ratios and the associated DEPs associated with the distortion g_{t+1}^{q} for the EZ preferences (green).

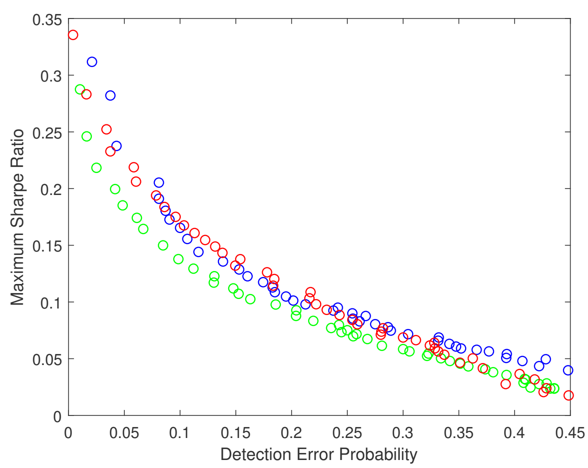


Figure 2: Maximum Sharpe Ratio and DEPs for the Endowment Economy
Red: HS (g); Blue: EZ (g); Green: EZ (g^q)

As can be seen in the figure, HS (red) and EZ (blue) preferences (both under their respective g_{t+1}'s) trace out a very similar relationship between the maximum Sharpe ratios and DEPs, implying that the difference between the exponential and power certainty equivalents derived from these two preferences are minimal in terms of their asset pricing implications, so long as they are parameterized by DEPs. The picture is changed, however, if one focuses on the distortion implied by combining the agent's model uncertainty and pessimism contained in the distortion g_{t+1}^{q} for EZ preferences. Here the green curve traces out a lower maximum Sharpe ratio for a given DEP.

The decomposition of the pricing kernel from section 3.4 can shed some light on the sources of the different preferences' abilities to match the maximum Sharpe ratio. Despite the general similarity in the mappings under both EZ and HS between the maximum Sharpe ratios and DEPs, at such a low DEP (4.25%) as implied by the calibrated EZ specification, HS preferences display a somewhat lower maximum Sharpe ratio. The negative correlation between the market price of risk and the market price of model uncertainty account for why the two sum to a larger variance than that of the kernel under HS preferences and stems from the different correlations between the market price of uncertainty and long run risk. With the value of q at 1.119, the role of the market price of pessimism (defined as the market price of model uncertainty to the $q-1$'th power) is necessarily small, contributing only a little more than one-tenth of the variance contributed by the market price of model uncertainty under EZ preferences. Under generalized model uncertainty, the market prices of risk and uncertainty contribute equally to the variance of the kernel (this follows by definition for the case of $q=2$ considered here: $\Lambda_{t+1}^{U}=g_{t+1}$ and $\Lambda_{t+1}^{P}=g_{t+1}^{q-1}$).

Model	σ_M	σ_Λ^R	σ_Λ^U	σ_Λ^P	$\rho\left(M,\Lambda^R\right)$	$\rho\left(\Lambda^U,X\right)$
HS ($q=1$)	0.236	0.002	0	0.237	-0.080	0.127
EZ ($\theta=0$)	0.267	0.002	0.028	0.239	0.092	-0.086
Gen. Unc.	0.184	0.002	0.092	0.092	0.115	-0.134

See the main text.

Table 3: Pricing Kernel Decomposition for the Endowment Economy

As the role of pessimism is tied positively to the entropic index q under EZ preferences and the maximum Sharpe ratio for a given level of model uncertainty is roughly equivalent to that implied by HS preferences, one might suspect that decreasing q (i.e., increasing risk aversion) serves to increase the amount of model uncertainty under EZ preferences while holding the role of pessimism constant (recall the discussion in section 3.4, a decrease in q reduces the exponent in $\Lambda_{t+1}^{P}\doteq g_{t+1}^{q-1}$, but also changes g_{t+1} itself, which one would expect would increase based on the intuition of a decrease in q going hand in hand with an increase in risk aversion). In contrast, table 4 shows that the proportion of the variance of the pricing kernel under EZ preferences from the market price of pessimism is decreasing in risk aversion (increasing in q). Thus EZ preferences and low levels of risk aversion attribute a substantial role to pessimism from the model uncertainty I presented above.

q	$\gamma^{\text{EZ}} = q/(q-1)$	$\gamma^{\text{EZ alternative}} = (q-\rho)/(q-1)$	σ_M	$100\frac{\sigma_\Delta^R}{\sigma_M}$	$100\frac{\sigma_\Delta^U}{\sigma_M}$	$100\frac{\sigma_\Delta^P}{\sigma_M}$
1.1	11	6.615	0.343	0.719	9.018	91.038
1.25	5	3.246	0.116	2.132	19.940	80.003
1.5	3	2.123	0.068	3.655	33.273	66.596
2	2	1.561	0.045	5.436	49.846	49.846

Table 4: Risk and Pessimism under EZ in the Endowment Economy

Table 5 displays the role of the entropic index q under generalized model uncertainty while holding the DEP constant at 25%. Increasing q, while holding the DEP constant, increases the maximum Sharpe ratio and moving inside the Hansen and Jagannathan (1997) bounds just past $q =$.[31] The $q = 2$ specification that gives equal weight to the market price of model uncertainty and pessimism is associated with a maximum Sharpe ratio of 0.184, not far from the empirical Sharpe ratio of 0.2261 in table 1. One the one hand, this specification attributes a conservative value to model uncertainty – recall that under HS and EZ preferences a DEP of less than 5% was necessary to approach the empirical Sharpe ratio. Yet on the other, this implies that agents' pessimism lead them to subjectively double the probability of events made more likely under the distorted distribution.

$q=$	1	1.1	1.2	1.3	1.4	1.5	1.75	2	2.25	2.5
MSR	0.095	0.096	0.099	0.103	0.112	0.121	0.1643	0.184	0.219	0.235

θ is adjusted to keep the DEP at 0.25.

Table 5: The Entropic Index and the Maximum Sharpe Ratio for Generalized Model Uncertainty in the Endowment Economy

The endowment economy, while informative, denies the maximizing agent a decision margin, with which she can react to the different environments of uncertainty, leaving her only to price the risks, uncertainty, and pessimism that confront her. Following Cochrane (2008, p. 295), one would expect agents to try and smooth out the distortions the fear when confronted with these pernicious distortion and, placing herself in a potentially different situation with a different valuation of the risks, uncertainties, and pessimism with which she is confronted. To address this, I will examine the robustness of the results to a production setting, where consumers try to do just that.

[31] See the online appendix.

4.3 Production Economy with Stochastic Productivity

In the endowment economy, agents take their consumptions streams as given and the different preference specifications lead to different valuations of theses streams. To analyze the question as to whether and how agents might choose different consumption streams to optimally robustify their decisions to the different forms of uncertainty they face requires that these consumption streams be chosen endogenously. To do so, I apply the generalized entropy constraint to a stochastic neoclassical growth model with a preference for robustness. The goal here is to see whether the asset pricing results from the endowment economy carry over when agents can react to their environment without sacrificing the fit to the macroeconomy.

I follow Tallarini (2000) and examine a neoclassical production model with consumption and labor margins and follow his parameterization closely (see the online appendix for details). The economy is populated by an infinitely lived household that optimizes over consumption C_t and labor supply N_t with the period utility function

$$U_t = \ln C_t + \psi \ln(1 - N_t) \tag{58}$$

subject to the standard budget constraint (see the online appendix).

Output Y_t is produced under perfect competition using the labor augmented Cobb-Douglas technology $Y_t = K_{t-1}^{\alpha}\left(e^{Z_t}N_t\right)^{1-\alpha}$ and K_{t-1} is the capital stock. Z_t is a stochastic productivity process and $\alpha \in [0,1]$ the capital share. Productivity is assumed to be a random walk with drift

$$a_t \equiv Z_t - Z_{t-1} = \overline{a} + \varepsilon_{z,t}, \;\; \varepsilon_{z,t} \sim \mathcal{N}(0, \sigma_z^2) \tag{59}$$

with $\varepsilon_{z,t}$ the innovation to Z_t. The endogenous variables are detrended with e^{-Z_t}.

The household's lifetime utility function is expressed recursively as[32]

$$v_t = \ln c_t + \psi \ln(1 - N_t) - \beta\theta \ln_q \left\{ E_t \left[\exp_q \left\{ -\frac{1}{\theta}\left(v_{t+1} + \frac{1}{1-\beta} a_{t+1} \right) \right\} \right] \right\} \tag{60}$$

with $\beta \in (0,1)$ the discount factor and v_t the value function at the optimum. The first of household's two optimality conditions is the intratemporal labor supply/productivity condition equalizing the utility cost of marginally increasing labor supply to the utility value of the additional consumption

$$\frac{\psi}{1 - N_t} = \frac{1}{c_t} w_t \tag{61}$$

and the second is the intertemporal Euler equation, rearranged as the first equation in (47) where $R_t \doteq RR_t^K + 1 - \delta$ is the return on capital and m_{t+1}, the stochastic discount factor of the household or

[32] See footnote 25 for a discussion of detrending.

pricing kernel (see section 3.4), is given by[33]

$$m_{t+1} = \beta \frac{c_t}{c_{t+1}} e^{-a_{t+1}} \left(\frac{\exp_q \left\{ -\frac{1}{\theta} \left(v_{t+1} + \frac{1}{1-\beta} a_{t+1} \right) \right\}}{E_t \left[\exp_q \left\{ -\frac{1}{\theta} \left(v_{t+1} + \frac{1}{1-\beta} a_{t+1} \right) \right\} \right]} \right)^q \tag{62}$$

4.3.1 Macroeconomic and Asset Pricing Implications

I begin by comparing the business cycle properties of model uncertainty following HS with $q = 1$ and the risk sensitive recursive utility specification of EZ parameterized via model uncertainty with $\theta = 0$, before turning to the case of the generalized model uncertainty. I then compare the specifications' ability to match asset pricing facts, here using the maximum Sharpe ratio, for varying DEPs. This reiterates the close relationship between EZ and HS preferences from the endowment economy, noting again the potential discord between the agent's and the econometrician's change of measure. Under generalized model uncertainty, increasing the entropic index q can put the model's asset pricing predictions inside the Hansen and Jagannathan (1997) bounds while maintaining a conservative DEP of 0.25.

The parameters q and θ are set to achieve a DEP of 0.25 between the approximating and worst case models of each specification. For the generalized model uncertainty case, q is set to 2 (set to bring the maximum Sharpe ratio to the bounds of the Hansen and Jagannathan (1997) bounds for a DEP of 25%) and θ is then set to match the DEP. The volatility of productivity growth is adjusted under each preference specification such that the volatility of consumption growth matches its empirical target in table 1. The approximating models for all three specifications do a comparably good job in matching the data, despite their different uncertainty specifications, consistent with what Backus, Ferriere, and Zin (2015) deem the "Tallarini property".

The business cycle moments for all three specifications with DEPs equal to 25% are contained in table 11, with the associated approximating models in the left and the worst cases in the right panels. In contrast to the endowment economy, agents choose their consumption streams endogenously. To facilitate comparability, the approximating models are calibrated such that agents choose consumption with identical means and standard deviations. As can be seen in the table, this requires different assumptions regarding the mean and volatility of productivity, especially for generalized model uncertainty preferences, where agents robustify their consumption decisions more strongly due to the overweighting of pernicious probability distributions under $q = 2$. That is, their precautionary motive is stronger,

[33] See the online appendix for a detailed derivation.

HS Preferences

$q = 1$, DEP=25%$\rightarrow \theta = 15$

	Approximating Model			Worst-Case Model		
Variable	Mean %	Std. Dev. %	Autocorr.	Mean %	Std. Dev. %	Autocorr.
$\Delta \ln Y_t$	0.400	1.029	0.009	0.286	1.030	0.008
$\Delta \ln C_t$	0.400	0.566	0.085	0.286	0.564	0.084
a_t	0.000	1.194	0.000	-0.114	1.192	0.000

EZ Preferences

$\theta = 0$, DEP=25%$\rightarrow q = 1.082$

	Approximating Model			Worst-Case Model		
Variable	Mean %	Std. Dev. %	Autocorr.	Mean %	Std. Dev. %	Autocorr.
$\Delta \ln Y_t$	0.400	1.023	0.009	0.298	1.026	0.007
$\Delta \ln C_t$	0.400	0.566	0.084	0.298	0.565	0.080
a_t	0.000	1.189	0.000	-0.102	1.190	-0.002

Generalized Uncertainty

$q = 2$, DEP=25%$\rightarrow \theta = 132.15$

	Approximating Model			Worst-Case Model		
Variable	Mean %	Std. Dev. %	Autocorr.	Mean %	Std. Dev. %	Autocorr.
$\Delta \ln Y_t$	0.400	1.233	0.008	0.272	1.235	0.006
$\Delta \ln C_t$	0.400	0.566	0.101	0.272	0.564	0.120
a_t	0.000	1.347	0.000	-0.128	1.345	0.001

For all three specifications, σ_z is adjusted to match the empirical volatility of $\Delta \ln C_t$ in the approximating model and the free parameter in the preference specification to yield a DEP of 25%.

Table 6: Business Cycle Moments for the Approximating and Worst Case Models in the Production Economy

leading them to more strongly smooth consumption, which in turn necessitates a lower mean and more volatile productivity process to result in the volatility of consumption growth as measured in the data in table 1.

The moments of the worst cases under the g_{t+1} measures can be found in the right panels of table 11. All three specifications agree on a reduction in mean consumption growth, achieved through a reduction in the mean productivity growth process of the worst cases. While the mean consumption growth rate under EZ is less than under HS, the increase in the volatility of production is larger under EZ relative to its worst case. This follows from the increase in production growth volatility whereas HS finds a decrease in the worst case. and an increase its volatility as properties of the worst case. Under generalized model uncertainty with $q = 2$, I find the most substantial reduction in mean consumption growth reiterating the stronger precautionary motive noted above in the worst case with an increase in production growth volatility resulting despite the decrease in productivity growth volatility. For the g^q_{t+1} measures for EZ and generalized uncertainty (see the online appendix for the tables), the worst case for EZ preferences further reduces the feared mean and increases the autocorrelation of consumption growth, bringing them nearly in line with the HS version in table 11. The decrease in mean consumption growth under this g^q_{t+1} measure is substantial for the $q = 2$ generalized uncertainty preferences, consistent with the high amount of pessimistic overweighted implied by the value of q.

Using the maximum Sharpe ratio to assess the asset pricing implications for varying DEPs essentially mirrors the results from the endowment economy.[34] There is a close relationship between EZ and HS preferences if the DEPs of the latter are calculated using only the margin of model uncertainty. In this case, both specifications close about half of the gap to the Hansen and Jagannathan (1997) bounds while maintaining a conservative DEP of 0.25. Likewise confirmed from the endowment economy is that using the combined distortion g^q_{t+1} for EZ preferences generates a much lower maximum Sharpe ratio for a given DEP than under the HS case, underlining the caveat that this close relation hinges on using only the change of measure g_{t+1} associated with EZ preferences. The two parameter generalized model uncertainty is able to move inside the bounds with an entropic index q of 2, the maximum Sharpe ratio is 0.21, just shy of the empirical Sharpe ratio of 0.2261, see the lower half of table 1, and more than twice the value obtained under both HS's and EZ's specifications. That agents overweight the probability of pernicious distributions including the worst case under the generalized model uncertainty formulation drives up the returns on risky capital relative to the risk free bond.

Model	σ_M	σ^R_Λ	σ^U_Λ	σ^P_Λ	$\rho\left(M, \Lambda^R\right)$
HS ($q = 1$)	0.101	0.006	0.096	0	0.983
EZ ($\theta = 0$)	0.101	0.006	0.089	0.007	0.984
Gen. Unc.	0.213	0.006	0.104	0.104	0.969

See the main text.

Table 7: Pricing Kernel Decomposition for the Production Economy

The variances of the decomposition of the pricing kernel from section 3.4 for the three preferences can be found in table 7. All three attribute a minimal amount of the kernels variation to the market price of risk and instead find model uncertainty and pessimism to be the primary contributors. EZ preferences do attribute the majority of the variance as stemming from uncertainty, but the role of pessimism (its variance is roughly 1/4 of that of uncertainty) is far from negligible – that the sum of the variances of the three components exceed the variance of the kernel itself is a result of the

[34] See the online appendix.

negative correlation of the market price of risk with the remaining components. Under generalized model uncertainty as in the endowment economy, the market prices of risk and uncertainty contribute equally to the variance of the kernel and here even individually demonstrate a higher volatility than the market price of model uncertainty in either the HS or EZ case.

As was the case in the endowment model, the proportion of the variance of the pricing kernel under EZ preferences that is attributable to pessimism depends positively on q (and hence negatively on risk aversion).[35] Indeed for relative risk aversions less than two, the variance of the market price of pessimism is the predominant contributor to the variance of the kernel and for plausible levels of risk aversion in the range of, say, 2 to 5, the market price of pessimism remains a significant contributor to the volatility of the kernel.

From an asset pricing perspective, the approach of generalized model uncertainty is of interest beyond its ability to provide a model uncertainty foundation for the EZ specification with arbitrary felicity functions. The combination of model uncertainty and pessimism in the formulation of expectations by overweighting the probability of events made more likely under pernicious distributions brings the macroeconomic model's predictions of the maximum Sharpe ratio in line with empirical post war U.S. observations for reasonable DEPs.

5 Conclusion

I have derived a generalization of the model uncertainty framework of HS, using Tsallis's (1988) generalized entropy. The resulting preferences recover HS's original formulation with an exponential certainty equivalent as one special case and recover the constant elasticity of substitution risk specification of EZ with a power certainty equivalent as another. This latter result is particularly important, as it provides a model uncertainty foundation for EZ preferences with arbitrary period utility functions (allowing, e.g., arbitrary intertemporal elasticities of substitution). While my generalized model uncertainty is able relax the special set of assumptions needed to link EZ and HS preferences, it comes with its own set of special assumptions. Namely, the inclusion of an additional margin of model uncertainty via the new parameter induced by Tsallis's (1988) generalized entropy, the entropic index q, that I deem a form of pessimism that induces agents to overweight events made more likely under the pernicious distributions they fear when forming expectations. This might in and of itself be viewed

[35] See table 13 in the online appendix.

as a negative result that demonstrates the difficulty in providing a clear link between EZ and HS in general settings.

In applications to both a standard RBC production model and a standard long-run risk endowment model, I find that both HS's original formulation and the model uncertainty formulation for EZ provide roughly the same predictions for the maximum Sharpe ratios for a given DEP. This supports the notion that EZ and HS preferences are closely related, but this comes with a caveat, as the Sharpe ratios are only equivalent when these probabilities are calculated for EZ using the approximating model without acknowledging the additional margin of pessimism. Aside from these limiting cases, increasing the entropic index (or increasing pessimism) leads to an increase in the maximum Sharpe ratio for a given DEP under the two parameter generalized model uncertainty I introduce here. The bounds of Hansen and Jagannathan (1997) can be entered with modest DEPs (25%) and an elevated entropic index ($q = 2$). This opens a new question as to how this entropic index is to be disciplined and whether uncertainty and pessimism can be separated empirically at all.

References

ABE, S., AND G. BAGCI (2005): "Necessity of q-Expectation Value in Nonextensive Statistical Mechanics," *Physical Review E*, 71(1), 016139.

ABE, S., A. K. RAJAGOPAL, A. PLASTINO, V. LATORA, A. RAPISARDA, AND A. ROBLEDO (2003): "Revisiting Disorder and Tsallis Statistics," *Science*, 300(5617), 249–251.

ANDERSON, E. W., L. P. HANSEN, AND T. 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.

ANDREASEN, M. M., J. FERNÁNDEZ-VILLAVERDE, AND J. RUBIO-RAMÍREZ (2017): "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications," *Review of Economic Studies*, Forthcoming.

BACKUS, D., A. FERRIERE, AND S. ZIN (2015): "Risk and ambiguity in models of business cycles," *Journal of Monetary Economics*, 69(C), 42–63.

BACKUS, D. K., B. R. ROUTLEDGE, AND S. E. ZIN (2005): "Exotic Preferences for Macroeconomists," in *NBER Macroeconomics Annual 2004, Volume 19*, NBER Chapters, pp. 319–414. National Bureau of Economic Research, Inc.

BANSAL, R., D. KIKU, AND A. YARON (2016): "Risks for the long run: Estimation with time aggregation," *Journal of Monetary Economics*, 82(7), 507–532.

BANSAL, R., AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59(4), 1481–1509.

BARILLAS, F., L. P. HANSEN, AND T. J. SARGENT (2009): "Doubts or Variability?," *Journal of Economic Theory*, 144(6), 2388–2418.

BIDDER, R., AND I. DEW-BECKER (2016): "Long-Run Risk Is the Worst-Case Scenario," *American Economic Review*, 106(9), 2494–2527.

BIDDER, R. M., AND M. E. SMITH (2012): "Robust Animal Spirits," *Journal of Monetary Economics*, 59(8), 738–750.

——— (2018): "Doubts and variability: A robust perspective on exotic consumption series," *Journal of Economic Theory*, 175, 689 – 712.

CAGETTI, M., L. P. HANSEN, T. J. SARGENT, AND N. WILLIAMS (2002): "Robustness and Pricing with Uncertain Growth," *The Review of Financial Studies*, 15(2), 363–404.

CHO, A. (2002): "A Fresh Take on Disorder, Or Disorderly Science?," *Science*, 297(5585), 1268–1269.

COCHRANE, J. H. (2008): "Financial Markets and the Real Economy," in *Handbook of the Equity Risk Premium*, ed. by R. Mehra, Handbooks in Finance, pp. 237 – 325. Elsevier.

COVER, T. M., AND J. A. THOMAS (1991): *Elements of Information Theory*. John Wiley and Sons, Inc.

DONALDSON, J., AND R. MEHRA (2008): "Risk-Based Explanations of the Equity Premium," in *Handbook of the Equity Risk Premium*, ed. by R. Mehra, Handbooks in Finance, chap. 2, pp. 37 – 99. Elsevier, San Diego.

DOW, J., AND S. WERLANG (1992): "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, 60(1), 197–204.

DUFFIE, D., AND L. G. EPSTEIN (1992): "Stochastic Differential Utility," *Econometrica*, 60(2), 353–394.

ELLISON, M., AND T. J. SARGENT (2015): "Welfare Cost of Business Cycles with Idiosyncratic Consumption Risk and a Preference for Robustness," *American Economic Journal: Macroeconomics*, 7(2), 40–57.

EPSTEIN, L., AND S. ZIN (1990): "'First-order' risk aversion and the equity premium puzzle," *Journal of Monetary Economics*, 26(3), 387–407.

EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–69.

GILBOA, I. (1987): "Expected utility with purely subjective non-additive probabilities," *Journal of Mathematical Economics*, 16(1), 65–88.

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

HANSEN, L. P. (2005): "Comment," in *NBER Macroeconomics Annual 2004, Volume 19*, NBER Chapters, pp. 391–404. National Bureau of Economic Research, Inc.

HANSEN, L. P., J. HEATON, J. LEE, AND N. ROUSSANOV (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6 of *Handbook of Econometrics*, chap. 61, pp. 3967–4056. Elsevier.

HANSEN, L. P., AND R. JAGANNATHAN (1997): "Assessing specification errors in stochastic discount factor models," *The Journal of Finance*, 52(2), 557–590.

HANSEN, L. P., AND M. MARINACCI (2016): "Ambiguity Aversion and Model Misspecification: An Economic Perspective," *Statistical Science*, 31(4), 511–515.

HANSEN, L. P., AND T. J. SARGENT (2001): "Robust Control and Model Uncertainty," *The American Economic Review*, 91(2), 60–66.

——— (2005): "Robust Estimation and Control under Commitment," *Journal of Economic Theory*, 124(2), 258–301.

——— (2007): *Robustness*. Princeton University Press.

——— (2009): "Wanting Robustness to Misspecification," mimeo.

——— (2010): "Wanting Robustness in Macroeconomics," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3 of *Handbook of Monetary Economics*, chap. 20, pp. 1097–1157. Elsevier.

ILUT, C. L., AND M. SCHNEIDER (2014): "Ambiguous Business Cycles," *American Economic Review*, 104(8), 2368–99.

JU, N., AND J. MIAO (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80(2), 559–591.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

KOCHERLAKOTA, N. R. (1990): "On the 'discount' factor in growth economies," *Journal of Monetary Economics*, 25(1), 43–47.

KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

LAN, H., AND A. MEYER-GOHDE (2013a): "Decomposing Risk in Dynamic Stochastic General Equilibrium," Discussion paper, SFB 649 2013-022.

——— (2013b): "Pruning in DSGE Perturbation," Discussion paper, SFB 649 2013-024.

——— (2013c): "Solving DSGE Models with a Nonlinear Moving Average," *Journal of Economic Dynamics and Control*, 37(12), 2643–2667.

MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 1447–1498.

MAENHOUT, P. J. (2004): "Robust Portfolio Rules and Asset Pricing," *Review of Financial Studies*, 17(4), 951–983.

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3(4), 323–43.

SCHMIEDLER, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57(3), 571–87.

SWANSON, E. T. (2012): "Risk Aversion and the Labor Margin in Dynamic Equilibrium Models," *American Economic Review*, 102(4), 1663–1691.

——— (2016): "A Macroeconomic Model of Equities and Real, Nominal, and Defaultable Debt," Discussion paper, mimeo.

——— (2018): "Risk Aversion, Risk Premia, and the Labor Margin with Generalized Recursive Preferences," *Review of Economic Dynamics*, 28, 290–321.

TALLARINI, JR., T. D. (2000): "Risk-sensitive real business cycles," *Journal of Monetary Economics*, 45(3), 507–532.

TSALLIS, C. (1988): "Possible Generalization of Bolzmann-Gibbs Statistics," *Journal of Statistical Physics*, 52(1/2), 479–487.

——— (1998): "Generalized Entropy Based Criterion for Consistent Testing," *Physical Review E*, 58(2), 479–487.

——— (2003): "Nonextensive Statistical Mechanics and Economics," *Physica A*, 324, 89–100.

——— (2009): *Introduction to Nonextensive Statistical Mechanics*. Springer, New York, NY.

TSALLIS, C., R. S. MENDES, AND A. PLASTINO (1998): "The Role of Constraints within Generalized Nonextensive Statistics," *Physica A*, 261, 534–554.

WEIL, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401–421.

——— (1990): "Nonexpected Utility in Macroeconomics," *The Quarterly Journal of Economics*, 105(1), 29–42.

A Appendix

A.1 Details of Equation (15)

The generalized risk-aggregator (14) can be rewritten as follows

$$\mathscr{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx' + \theta I_q\left(\tilde{p}(x',x,a(x)),p(x',x,a(x))|x\right)$$

(A-1)

$$\begin{aligned} = & \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')\frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'}dx' \\ & + \left(1-\frac{1}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'}\right)\int V(x')g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx' \\ & + \theta I_q\left(\tilde{p}(x',x,a(x)),p(x',x,a(x))|x\right) \end{aligned}$$

(A-2)

$$\begin{aligned} = & \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')\frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'}dx' \\ & + \left(\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'-1\right)\int V(x')\frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'}dx' \\ & + \theta I_q\left(\tilde{p}(x',x,a(x)),p(x',x,a(x))|x\right) \end{aligned}$$

(A-3)

As the distorted density $\tilde{p}(x',x,a(x)) = g(x',x)p(x',x,a(x))$ is restricted to indeed be a probability density function $\int \tilde{p}(x',x,a(x))dx' = \int g(x',x)p(x',x,a(x))dx' = 1$, the term

$$\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'-1 = \int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx' - \int \tilde{p}(x',x,a(x))dx' \tag{A-4}$$

$$= \int \left(g(x',x)^{q-1}-1\right)\tilde{p}(x',x,a(x))dx' \tag{A-5}$$

is proportional to q-relative entropy

$$I_q\left(\tilde{p}(x',x,a(x)),p(x',x,a(x))\right) \doteq \int p(x)\left(\frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}\right)^q \ln_q\left(\frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}\right)dx' \tag{A-6}$$

$$= \int p(x',x,a(x))g(x',x)^q \ln_q\left(g(x',x)\right)dx' \tag{A-7}$$

$$= \int p(x',x,a(x))g(x',x)^q\frac{g(x',x)^{1-q}-1}{1-q}dx' \tag{A-8}$$

$$= \int p(x',x,a(x))\frac{g(x',x)-g(x',x)^q}{1-q}dx' \tag{A-9}$$

$$= \int \frac{1-g(x',x)^{q-1}}{1-q}\tilde{p}(x',x,a(x))dx' \tag{A-10}$$

Substituting into the risk aggregator above gives

$$\mathcal{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x') \frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'} dx'$$
$$+ (q-1) I_q \left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right) \int V(x') \frac{g(x',x)^{q-1}\tilde{p}(x',x,a(x))}{\int g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx'} dx'$$
$$+ \theta I_q \left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right)$$

(A-11)

Collecting terms in $I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right)$ delivers (15) in the main text.

A.2 Details of Proposition 3.1

Abusing notation to minimize clutter by suppressing the dependence on x, the current state, and recycling notation by relabeling the future state, x', with x, the aggregator of (14) can be rewritten as

(A-12) $$\tilde{V} \doteq \min_{g(x)>0} \int V(x) g(x)^q p(x) dx + \theta \int g(x)^q \ln_q\left(g(x)\right) p(x) dx + \lambda \left(\int g(x) p(x) dx - 1 \right)$$

The first order condition is

(A-13) $$0 = qV(x)g(x)^{q-1}p(x) + \theta \frac{1 - qg(x)^{q-1}}{1-q} + \lambda p(x)$$

multiplying the foregoing with $g(x)$

(A-14) $$0 = qV(x)g(x)^q p(x) + \theta \frac{g(x) - qg(x)^q}{1-q} + \lambda p(x) g(x)$$

and rearranging yields

(A-15)
$$0 = q\left[V(x)g(x)^q p(x)dx + \theta(x)^q \ln_q\left(g(x)\right)p(x) + \lambda\left(g(x)p(x) - 1\right)\right] + \theta g(x)p(x) + (1-q)\lambda p(x)g(x) + \lambda q$$

Integrating over x yields

(A-16) $$0 = q\tilde{V} + \theta + \lambda$$

Combining the foregoing, (A-16) with the first order condition, (A-13)

(A-17) $$0 = q\left[V(x) - \tilde{V} + \theta(x)g(x)^{q-1} ln_q\left(g(x)\right)\right] p(x)$$

noting that $p(x)$ and q are assumed nonzero gives

(A-18) $$0 = V(x) - \tilde{V} + \theta(x) \frac{\left(1 - g(x)^{q-1}\right)}{1-q}$$

which can be rearranged as

(A-19) $$0 = V(x) - \tilde{V} + \theta \frac{\left(1 - g(x)^{q-1}\right)}{(1-q)} - V(x)\left(1 - g(x)^{q-1}\right)$$

and

(A-20) $$0 = g(x)^{q-1}V(x) - \tilde{V} + \frac{\theta}{1-q}\left(1 - g(x)^{q-1}\right)$$

multiplying the foregoing with[36] $\frac{1-q}{\theta}g(x)^{1-q}$ delivers

$$0 = (1-q)\frac{1}{\theta}V(x) - (1-q)\frac{1}{\theta}g(x)^{1-q}\tilde{V} + g(x)^{1-q} - 1 \tag{A-21}$$

or

$$1-(1-q)\frac{1}{\theta}V(x) = g(x)^{1-q}\left(1-(1-q)\frac{1}{\theta}\tilde{V}\right) \tag{A-22}$$

from which the minimizing likelihood ratio, $g(x)$, follows as

$$g(x) = \frac{\left(1-(1-q)\frac{1}{\theta}V(x)\right)^{\frac{1}{1-q}}}{\left(1-(1-q)\frac{1}{\theta}\tilde{V}\right)^{\frac{1}{1-q}}} = \frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q\left(-\frac{1}{\theta}\tilde{V}\right)} \tag{A-23}$$

and the minimizing, or worst-case, probability distribution is then

$$\tilde{p}(x) = p(x)\frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q\left(-\frac{1}{\theta}\tilde{V}\right)} \tag{A-24}$$

as was claimed in proposition 3.1.

Integrating both sides of the previous equation with respect to x gives

$$1 = \int p(x)\frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q(-\frac{1}{\theta}\tilde{V})}dx \tag{A-25}$$

which, as $\tilde{V}$ is independent of x, can be written as

$$\exp_q(-\frac{1}{\theta}\tilde{V}) = \int p(x)\exp_q\left(-\frac{1}{\theta}V(x)\right)dx \tag{A-26}$$

yielding the risk aggregator or certainty equivalent

$$\tilde{V} = -\theta\ln_q\left[\int \exp_q\left(-\frac{1}{\theta}V(x)\right)p(x)dx\right] \tag{A-27}$$

as was claimed in proposition 3.1.

A.3 Sequential Formulation

To also examine sequential decision problems, it will be useful to consider sequential formulations of g generalized entropy. I follow Hansen and Sargent (2005) closely and rederive their sequential entropy formulations under q generalizations. Relabel the likelihood ratio or Radon-Nikodym derivative $g_{t+1} \doteq g(x',x) \doteq \frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}$ and distributions $\tilde{p}_{t+1} \doteq \tilde{p}(x',x,a(x))$ and $p_{t+1} \doteq p(x',x,a(x))$ from (13) for notational ease and define the martingale

$$G_t = g_t G_{t-1} = G_0\prod_{s=1}^{t} g_s \tag{A-28}$$

[36] Note that if $\theta = 0$, the foregoing reduces to $0 = g(x)^{q-1}V(x) - \tilde{V}$, which can be solved for the minimizing likelihood ratio $g(x)$ as $g(x) = \left(\frac{V(x)}{\tilde{V}}\right)^{\frac{1}{1-q}}$, which is the same as (A-23) with θ set to zero.

Following Hansen and Sargent (2005), I will decompose $-E_0\left[S_q(G_t)\right] = E_t\left[G_t^q \ln_q G_t\right]$, the q generalized (negative) entropy of a time t distortion conditional on time 0 information as[37]

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[g_{j+1}^q \ln_q g_{j+1}\right]\right] \tag{A-29}$$

$$= E_0\left[\sum_{j=0}^{t-1} G_j^q I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right] \tag{A-30}$$

or the relative entropies of the distortions to the approximating model. Likewise following Hansen and Sargent (2005), it will be useful to have a q generalized discounted relative entropy, as this will be used to construct an entropy penalty for sequential decision problems.

$$(1-\rho)\sum_{t=1}^{\infty} \rho^t E_0\left[G_t^q \ln_q G_t\right] = \rho \sum_{t=1}^{\infty} \rho^t E_0\left[G_t^q E_t\left[g_{t+1}^q \ln_q g_{t+1}\right]\right] \tag{A-31}$$

$$= \rho \sum_{t=1}^{\infty} \rho^t E_0\left[G_t^q I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right] \tag{A-32}$$

which follows from (A-29) and summation by parts.

Here I state and interpret a sequential decision problem that leads to the recursive formulations above.[38] Consider the following zero-sum two player game with a minimizer choosing a martingale to distort the maximizer's model

$$\max_{\{a_t\}} \min_{\{g_{t+1}\}} E_0\left[\sum_{t=0}^{\infty} G_t^q \beta^t\left(u(x_t, a_t) + \theta\beta\left[g_{t+1}^q \ln_q g_{t+1}\right]\right)\right] \tag{A-33}$$

where G_t is a martingale defined in (A-28) and x_t is distributed as in (10). Under a Bellman-Isaacs condition, I am free to exchange the minimization and maximizations. Doing so and considering a given sequence of controls, $\{a_t \in A_t\}$, defining $W_0 \doteq E_0\left[\sum_{t=0}^{\infty} G_t^q \beta^t u(x_t, a_t)\right]$, the minimizer solves the following value equation

$$V_0 = \min_{\{g_{t+1}\}} W_0 + \theta\beta E_0\left[\sum_{t=0}^{\infty} G_t^q \beta^t g_{t+1}^q \ln_q g_{t+1}\right] \tag{A-34}$$

using the law of iterated expectations, the second term in the foregoing is (A-31), the q generalized discounted relative entropy with β as the discount factor. The value function solves the following Bellman equation

$$G_t^q V_t = \min_{g_{t+1}} G_t^q u(x_t, a_t) + \beta\theta G_t^q E_t\left[g_{t+1}^q \ln_q g_{t+1}\right] + \beta E_t\left[G_{t+1}^q \tilde{V}_{t+1}\right] \tag{A-35}$$

where $\tilde{V}_{t+1}$ is next period's value function at the optimum. Dividing through by G_t^q gives

$$V_t = \min_{g_{t+1}}\left(u(x_t, a_t) + \beta\theta E_t\left[g_{t+1}^q \ln_q g_{t+1}\right] + \beta E_t\left[\frac{G_{t+1}^q}{G_t^q}\tilde{V}_{t+1}\right]\right) \tag{A-36}$$

[37]See the online appendix.

[38]See especially Hansen and Sargent (2007) and Hansen and Sargent (2005), whom I follow closely here.

$$= u(x_t, a_t) + \beta \min_{g_{t+1}} \left(E_t \left[g_{t+1}^q \tilde{V}_{t+1} \right] + \theta I_q \left(\tilde{p}_{j+1}, p_{j+1} \right) \right) \tag{A-37}$$

Inspection shows that the minimization problem is identical to (14). Using the solution to the minimizer's problem from proposition (3.1) delivers

$$\tilde{V}_t = u(x_t, a_t) - \beta\theta \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\theta} \tilde{V}_{t+1} \right) \right] \right) \tag{A-38}$$

a recursive representation of the objective function of the maximizer, who then chooses the sequence $\{a_t\}$ accordingly.

B Online Appendix [Not for Publication]

B.1 Assumptions for Equivalence between HS and EZ

The recursive preferences of EZ lead to a power certainty equivalent, see section 3.2, whereas those of HS lead to an exponential certainty equivalent, see section 3.1. As has been demonstrated by, e.g., Tallarini (2000), Hansen, Heaton, Lee, and Roussanov (2007), Barillas, Hansen, and Sargent (2009), and Ju and Miao (2012), the two are closely related under special restrictions on the parameters and the period utility function. I review this in the following proposition.

Proposition B.1. *Logarithmic Equivalence of Risk Sensitive and Model Uncertainty Preferences If the elasticity of intertemporal substitution in (26) is one, the period utilities are related through a logarithmic transformation*

$$-\theta = \frac{1}{(1-\beta)(1-\gamma)} \quad \text{and} \quad u^{HS}(x,a(x)) = \ln\left(u^{EZ}(x,a(x))\right) \tag{B-39}$$

then

$$V^{HS}(x) = \frac{1}{1-\beta}\ln\left(V^{EZ}(x)\right) \tag{B-40}$$

Proof. Setting the intertemporal elasticity of substitution in (26) to one ($\rho = 1$) and taking logs yields

$$\ln((V^{EZ}(x)) = (1-\beta)\ln(u^{EZ}(x,a(x))) + \frac{\beta}{1-\gamma}\ln\left(\int V^{EZ}(x')^{1-\gamma} p(x',x,a(x))dx'\right)$$

Defining $\tilde{V}^{EZ}(x) = \frac{\ln((V^{EZ}(x))}{1-\beta}$ and dividing the foregoing by $(1-\beta)$ gives

$$\tilde{V}^{EZ}(x) = \ln(u^{EZ}(x,a(x))) + \frac{\beta}{(1-\beta)(1-\gamma)}\ln\left(\int \exp \tilde{V}^{EZ}(x)(1-\beta)(1-\gamma)p(x',x,a(x))dx'\right)$$

comparison with the recursive formulation under HS

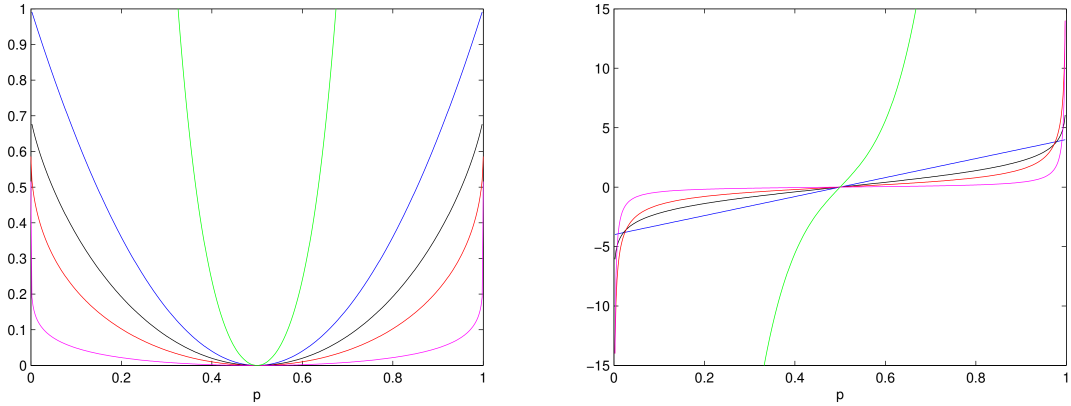
$$V^{HS}(x) = u^{HS}(x,a(x)) - \theta\beta\ln\int\exp\left[-\frac{1}{\theta}V^{HS}(x')p(x',x,a(x))\right]dx' \tag{B-41}$$

completes the proof. □

Risk sensitive and uncertainty averse preferences coincide but only in the special case of an intertemporal elasticity of substitution of one and a logarithmic relationship between the period utility functions – Hansen, Heaton, Lee, and Roussanov's (2007, p. 3975) "special set of assumptions."

B.2 Generalized Relative Entropy: Two State Examples

Figure 3a plots (8) for a two state random variable over possible values of $\tilde{p}$ for differing values of the entropic index with the baseline distribution given by the equiprobable case. When the two distributions match ($\tilde{p} = p = 0.5$), relative entropy is zero. Elsewhere, entropy is positive and increasing in the entropic index. For $q > 1$ ($q < 1$), relative entropy is greater (less) than the Kullback-Leibler



(a) q-Relative Entropy

(b) $\frac{\partial}{\partial \tilde{p}(x)}$ of q-Relative Entropy

Figure 3: q-Relative Entropy or Generalized Kullback-Leibler Divergence
magenta—$q = 0.1$, red—$q = 0.5$, black—$q = 1$, blue—$q = 2$, green—$q = 10$
$p(x) = 0.5$—Two State Equiprobable

divergence. Figure 3b plots the derivative with respect to $\tilde{p}$, which also varies with q. Note that for the case $q = 2$, the derivative is linear in $\tilde{p}$ given by $-\frac{2}{1-p} + \frac{2}{p(1-p)}\tilde{p}$. Thus, the entropic index does more than just scale standard relative entropy, but also changes the margin.

Figure 4 provides the same picture as figure 3, but now $p = 0.75$, as can be deduced by the point of zero relative entropy. This change not only shifts the picture from before to the right, but also tilts the measures to the right, as can be confirmed using the linear relationship for the $q = 2$ case above.

B.3 Conditional Entropy

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] \tag{B-42}$$

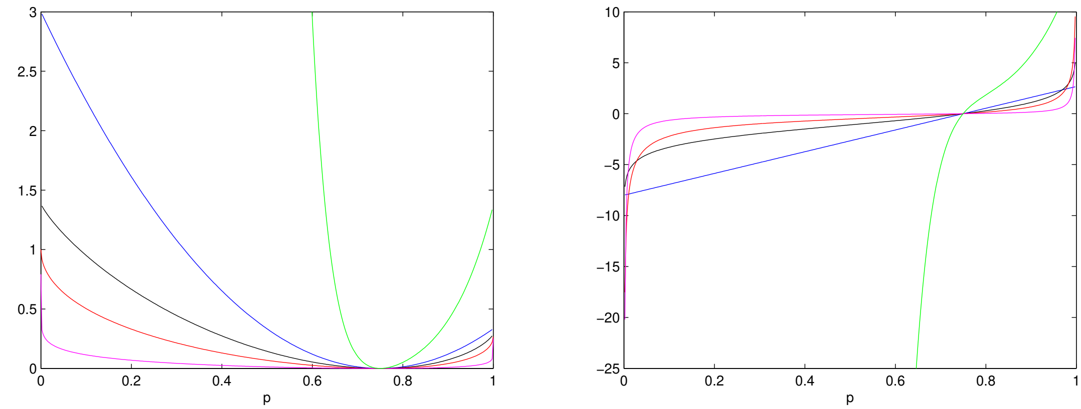
can be written as

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] - \frac{1}{1-q}E_0\left[G_0^q\left(G_0^{1-q}-1\right)\right] \tag{B-43}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] + \frac{1}{1-q}E_0\left[G_0^q\left(1-G_tG_0^{-q}\right)\right] \tag{B-44}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] + \frac{1}{1-q}E_0\left[\left(G_0^q - G_t\right)\right] \tag{B-45}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_0^q G_t^{-q}-1\right)\right] \tag{B-46}$$



(a) q-Relative Entropy (b) $\frac{\partial}{\partial \tilde{p}(x)}$ of q-Relative Entropy

Figure 4: q-Relative Entropy or Generalized Kullback-Leibler Divergence
magenta—$q = 0.1$, red—$q = 0.5$, black—$q = 1$, blue—$q = 2$, green—$q = 10$
$p(x) = 0.75$—Two State Nonequiprobable

where the second equality follows as $E_0[G_t] = G_0$. Define the following q martingale

(B-47) $$\tilde{G}_t \doteq G_t^q = g_t^q \tilde{G}_{t-1} = G_0 \prod_{s=1}^{t} g_s^q$$

which is a sub-/super-martingale as q is greater/less than zero. Equation (B-60) can be written as

(B-48) $$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = \frac{1}{1-q} E_0\left[\tilde{G}_t\left(\tilde{G}_0 \tilde{G}_t^{-1} - 1\right)\right]$$

(B-49) $$= \frac{1}{1-q} E_0\left[\tilde{G}_0 - \tilde{G}_t\right]$$

(B-50) $$= \frac{1}{1-q} E_0\left[\tilde{G}_0 + \sum_{j=0}^{t-1}\left(\tilde{G}_{j+1} - E_j\left[\tilde{G}_{j+1}\right]\right) - E_{t-1}\left[\tilde{G}_t\right]\right]$$

(B-51) $$= \frac{1}{1-q} E_0\left[\sum_{j=0}^{t-1} E_j\left[\tilde{G}_j g_{j+1} - \tilde{G}_{j+1}\right]\right]$$

(B-52) $$= E_0\left[\sum_{j=0}^{t-1} \tilde{G}_j E_j\left[\frac{g_{j+1} - \frac{\tilde{G}_{j+1}}{\tilde{G}_j}}{1-q}\right]\right]$$

(B-53) $$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[\frac{g_{j+1} - g_{j+1}^q}{1-q}\right]\right]$$

(B-54) $$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[g_{j+1}^q \frac{g_{j+1}^{1-q} - 1}{1-q}\right]\right]$$

(B-55) $$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[g_{j+1}^q \ln_q g_{j+1}\right]\right]$$

or, using the definition of q relative entropy in (8),

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = E_0\left[\sum_{j=0}^{t-1} G_j^q I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right] \tag{B-56}$$

B.4 Discounted Conditional Entropy

$$(1-\rho)\sum_{t=1}^{\infty}\rho^t E_0\left[G_t^q \ln_q G_t\right] = \rho\sum_{t=1}^{\infty}\rho^t E_0\left[G_t^q E_t\left[g_{t+1}^q \ln_q g_{t+1}\right]\right] \tag{B-57}$$

$$= \rho\sum_{t=1}^{\infty}\rho^t E_0\left[G_t^q I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right] \tag{B-58}$$

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] \tag{B-59}$$

can be written as

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] - \frac{1}{1-q}E_0\left[G_0^q\left(G_0^{1-q}-1\right)\right] \tag{B-60}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] + \frac{1}{1-q}E_0\left[G_0^q\left(1-G_t G_0^{-q}\right)\right] \tag{B-61}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_t^{1-q}-1\right)\right] + \frac{1}{1-q}E_0\left[\left(G_0^q - G_t\right)\right] \tag{B-62}$$

$$= \frac{1}{1-q}E_0\left[G_t^q\left(G_0^q G_t^{-q}-1\right)\right] \tag{B-63}$$

where the second equality follows as $E_0\left[G_t\right] = G_0$. Define the following q martingale

$$\tilde{G}_t \doteq G_t^q = g_t^q \tilde{G}_{t-1} = G_0 \prod_{s=1}^{t} g_s^q \tag{B-64}$$

which is a sub-/super-martingale as q is greater/less than zero. Equation (B-60) can be written as

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = \frac{1}{1-q}E_0\left[\tilde{G}_t\left(\tilde{G}_0 \tilde{G}_t^{-1}-1\right)\right] \tag{B-65}$$

$$= \frac{1}{1-q}E_0\left[\tilde{G}_0 - \tilde{G}_t\right] \tag{B-66}$$

$$= \frac{1}{1-q}E_0\left[\tilde{G}_0 + \sum_{j=0}^{t-1}\left(\tilde{G}_{j+1} - E_j\left[\tilde{G}_{j+1}\right]\right) - E_{t-1}\left[\tilde{G}_t\right]\right] \tag{B-67}$$

$$= \frac{1}{1-q}E_0\left[\sum_{j=0}^{t-1} E_j\left[\tilde{G}_j g_{j+1} - \tilde{G}_{j+1}\right]\right] \tag{B-68}$$

$$= E_0\left[\sum_{j=0}^{t-1}\tilde{G}_j E_j\left[\frac{g_{j+1} - \frac{\tilde{G}_{j+1}}{\tilde{G}_j}}{1-q}\right]\right] \tag{B-69}$$

$$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[\frac{g_{j+1} - g_{j+1}^q}{1-q}\right]\right] \tag{B-70}$$

$$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[g_{j+1}^q \frac{g_{j+1}^{1-q}-1}{1-q}\right]\right] \tag{B-71}$$

$$= E_0\left[\sum_{j=0}^{t-1} G_j^q E_j\left[g_{j+1}^q \ln_q g_{j+1}\right]\right] \tag{B-72}$$

or, using the definition of q relative entropy in (8),

$$E_0\left[G_t^q \ln_q G_t\right] - E_0\left[G_0^q \ln_q G_0\right] = E_0\left[\sum_{j=0}^{t-1} G_j^q I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right] \tag{B-73}$$

B.5 Discounted Entropy and Detrended Certainty Equivalents

Consider the sequential formulation of section A.3 where the period utility grows with the factor W_t and relative entropy is discounted with the same factor following Cagetti, Hansen, Sargent, and Williams (2002).

The necessity of discounting the entropy penalty with the growth factor of period utility follows the same logic that discounts entropy in the absence of growth with the idiosyncratic discount factor. As Hansen and Sargent (2005) point out, a failure to discount the entropy penalty in the presence of utility discounting will cause concerns for robustness to wear off over time, causing agents to front load their the probability distortions associated with these concerns.[39] Consider now a situation with constant growth of period utility. From Kocherlakota (1990), we know that the effective discount factor is now the product of the idiosyncratic discount factor the growth factor of the period utility function. Hence, not discounting with the growth factor causes agents to grow out of their concerns for robustness, analogously to Hansen and Sargent (2005). In the presence of stochastic growth, if the entropy penalty is not discounted with the product of the now stochastic growth and idiosyncratic discount factor, a positive shock to growth will lead to the same, albeit now conditionally, wearing off of robustness concerns and front loading of probability distortions.

The zero-sum two player game with a minimizer choosing a martingale to distort the maximizer's model is now

$$V_t = \min_{g_{t+1}} \left(u(x_t, a_t) + W_t \theta E_t\left[g_{t+1}^q \ln_q g_{t+1}\right] + \beta E_t\left[g_{t+1}^q \tilde{V}_{t+1}\right]\right) \tag{B-74}$$

$$= u(x_t, a_t) + \beta \min_{g_{t+1}} \left(E_t\left[g_{t+1}^q \tilde{V}_{t+1}\right] + \theta \frac{W_t}{\beta} I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right) \tag{B-75}$$

$$= u(x_t, a_t) + \beta \min_{g_{t+1}} \left(E_t\left[g_{t+1}^q \tilde{V}_{t+1}\right] + \tilde{\theta}_t I_q\left(\tilde{p}_{j+1}, p_{j+1}\right)\right) \tag{B-76}$$

[39] Hansen and Sargent (2009) show that failing to discount entropy leads to an indifference towards the timing of the resolution of uncertainty in the two-state dynamic example of Kreps and Porteus (1978) and towards persistence in the example of Duffie and Epstein (1992).

Inspection shows that the minimization problem is identical to (14) but with a time varying multiplier on the entropy constraint $\tilde{\theta}_t \doteq \theta \frac{W_t}{\beta}$. Using the solution to the minimizer's problem from proposition (3.1) delivers

$$V_t = u(x_t, a_t) - \beta \tilde{\theta}_t \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\tilde{\theta}_t} V_{t+1} \right) \right] \right) \tag{B-77}$$

Let $\tilde{u}(x_t, a_t) \doteq u(x_t, a_t)/F_t$ be stationarized or detrended period utility. Dividing the foregoing through by F_t yields

$$\tilde{V}_t = \tilde{u}(x_t, a_t) - \beta \frac{\tilde{\theta}_t}{F_t} \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\tilde{\theta}_t} V_{t+1} \right) \right] \right) \tag{B-78}$$

where $\tilde{V}_t \doteq u(x_t, a_t)/F_t$ is the stationarized or detrended value function. Using the definition of $\tilde{\theta}_t$ above and requiring W_t be equal to βF_t following Cagetti, Hansen, Sargent, and Williams (2002) gives

$$\tilde{V}_t = \tilde{u}(x_t, a_t) - \beta \frac{\theta \frac{W_t}{\beta}}{F_t} \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\theta \frac{W_t}{\beta}} \tilde{V}_{t+1} F_{t+1} \right) \right] \right) \tag{B-79}$$

$$= \tilde{u}(x_t, a_t) - \beta \frac{\theta \frac{W_t}{\beta}}{F_t} \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\theta \frac{W_t}{\beta}} \tilde{V}_{t+1} F_{t+1} \right) \right] \right) \tag{B-80}$$

$$= \tilde{u}(x_t, a_t) - \beta \frac{\theta F_t}{F_t} \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\theta F_t} \tilde{V}_{t+1} F_{t+1} \right) \right] \right) \tag{B-81}$$

$$= \tilde{u}(x_t, a_t) - \beta \theta \ln_q \left(E_t \left[\exp_q \left(-\frac{1}{\theta} \tilde{V}_{t+1} \frac{F_{t+1}}{F_t} \right) \right] \right) \tag{B-82}$$

That is, discounting entropy following Cagetti, Hansen, Sargent, and Williams (2002) with the stochastic factor compatible with time preference and stochastic growth is equivalent to first stationarizing the approximating model and then applying the concern for robustness.

B.6 Risk Aversion

$$\begin{aligned} \mathcal{R}(V) &= -\theta \ln_q \left(0.5 \exp_q \left(-\frac{1+\sigma}{\theta} \right) + 0.5 \exp_q \left(-\frac{1-\sigma}{\theta} \right) \right) \\ &= -\theta \frac{\left(0.5 \left[1 + (1-q) \left(-\frac{1+\sigma}{\theta} \right) \right]^{\frac{1}{1-q}} + 0.5 \left[1 + (1-q) \left(-\frac{1-\sigma}{\theta} \right) \right]^{\frac{1}{1-q}} \right)^{1-q} - 1}{1-q} \end{aligned} \tag{B-83}$$

$$\begin{aligned} \frac{\partial \mathcal{R}(V)}{\partial \sigma} &= -\left(0.5 \left[1 + (1-q) \left(-\frac{1+\sigma}{\theta} \right) \right]^{\frac{1}{1-q}} + 0.5 \left[1 + (1-q) \left(-\frac{1-\sigma}{\theta} \right) \right]^{\frac{1}{1-q}} \right)^{-q} \\ &\times \left(0.5 \left[1 + (1-q) \left(-\frac{1+\sigma}{\theta} \right) \right]^{\frac{q}{1-q}} - 0.5 \left[1 + (1-q) \left(-\frac{1-\sigma}{\theta} \right) \right]^{\frac{q}{1-q}} \right) \end{aligned} \tag{B-84}$$

$$\frac{\partial^2 \mathcal{R}(V)}{\partial \sigma^2} = q\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{-q-1}$$
$$\times\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{q}{1-q}} - 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{q}{1-q}}\right)^{2}$$
$$-\frac{q}{\theta}\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{-q}$$
$$\times\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{2q-1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{2q-1}{1-q}}\right) \tag{B-85}$$

B.7 Risk Aversion following Swanson (2018)

In many macroeconomic models, agents possess a flexible labor margin that can be adjusted to attenuate the consequences of risk on their utility. Swanson (2018) incorporates these flexible labor margins into measures of risk aversion defined over wealth gambles, which lead to different attitudes towards risk by agents. I extend his measure to the generalized uncertainty preferences presented above.

Swanson (2018) calculates the risk aversion with a Taylor expansion of an agent's value function with respect to a risky wealth gamble with period utility potentially a function of both consumption and labor. Accordingly, let the agents value function be given by

$$V(a,x) = u(c,l) - \beta\theta \ln_q\left(\int \exp_q\left(-\frac{1}{\theta}V(a',x')\right)p(x)dx\right) \tag{B-86}$$

where the agents assets next period, a', are given by

$$a' = (1+r)a + wl + d - c + \sigma\varepsilon' \tag{B-87}$$

with r the return, w the wage, l the agent's labor supply, d net transfers, c her consumption, and $\sigma\varepsilon'$ a one-shot risky gamble. The state vector x is exogenous to the household and governs the wage, returns, and transfers. Swanson (2018) determines the one-time fee μ the agent would be willing to pay to avoid the risky gamble

$$a' = (1+r)a + wl + d - c - \mu \tag{B-88}$$

The agent's coefficient of absolute wealth-gamble risk aversion $R^a(a,x) = \lim_{\sigma\to 0}\mu(a,x,\sigma)/(\sigma^2/2)$, where $\mu(a,x,\sigma)$ is the fee the agent would be willing to pay to avoid the gamble σ in state x with wealth a. Given the assumptions 1 and 3-8 of Swanson (2018), this measure of risk aversion is given in the nonstochastic steady state by

$$R^a(a,x) = -\frac{V'_{11}}{V'_1} + \frac{q}{\theta}\left(\exp_q\left(-\frac{1}{\theta}V'\right)\right)^{q-1} V'_1 \tag{B-89}$$

$$= -\frac{V'_{11}}{V'_1} + q\frac{V'_1}{\theta-(1-q)V'} \tag{B-90}$$

$$= \begin{cases} -\frac{u_{cc}+\lambda u_{cl}}{u_c}\frac{r}{1+w\lambda} + q\frac{ru_c}{\theta\frac{r}{1+r}-(1-q)u} & \text{for } A=1 \\ -\frac{u_{cc}+\lambda u_{cl}}{u_c}\frac{r}{1+w\lambda} + q\frac{ru_c}{\theta-(1-q)u} & \text{for } A=(1-\beta) \end{cases} \tag{B-91}$$

where $\lambda = (u_c u_{cl} - u_l u_{cc} \big/ (u_c u_{ll} - u_l u_{cl})$.

The agent's consumption-wealth and consumption-and-leisure-wealth coefficients of relative wealth-gamble risk aversion are given through the following transformations that incorporate the agent's human wealth into the total wealth consideration

$$R^c(a,x) = \frac{c}{r}R^a(a,x) \tag{B-92}$$

$$R^{cl}(a,x) = \frac{c+w\left(\bar{l}-l\right)}{r}R^a(a,x) \tag{B-93}$$

where $\bar{l}$ is the agent's time endowment.

Simple inspection of (B-89) for the cases $\theta = 0$ and $q = 1$ confirm that these are identical to the measures provided in Swanson (2018).[40]

B.8 Data

All business cycle data was retrieved from the Federal Reserve Economic Data (FRED) database of the Federal Reserve Bank of St. Louis.

R is the return on the NYSE value weighted portfolio from the CRSP dataset and R^f is the secondary market rate on the three month Treasury bill. Both returns have been deflated by the implicit deflator of the PCE Nondurables and Services series.

B.9 Calibration of the Endowment Economy

The endowment economy is paramterized for the EZ case as in table 8, where θ necessarily equals zero.[41] The values of θ and q for the HS and generalized uncertainty preferences are discussed in the main text.

[40] The one caveat being $\frac{r}{1+r}$ in the last term on the last line of (B-89). As discussed by Swanson (2018), however, this term, which is equal to $1-\beta$, ought to be present when the agent's period utility function is not scaled by $1-\beta$.

[41] See section 3.2.

Parameter	β	q	ρ	ϕ	ρ_x	ρ_σ	ρ_{dc}	σ_x	σ_d	σ_c	σ_σ	μ_c	μ_d
Value	0.997	1.119	0.438	3.209	0.991	0.995	0.057	0.081	5.39	0.005	1.3e-6	0.005	0.002

See the main text.

Table 8: Parameter Values

B.10 Detailed Derivation of the Pricing Kernel for the Endowment Economy

The stochastic discount factor of the household or pricing kernel (see section 3.4), M_{t+1}, is given by

$$M_{t+1} \doteq \frac{\partial v_t/\partial C_{t+1}}{\partial v_t/\partial C_t} = \frac{\frac{\partial v_t}{\partial v_{t+1}}\frac{\partial v_{t+1}}{\partial C_{t+1}}}{\frac{\partial v_t}{\partial C_t}} \tag{B-94}$$

with

$$\frac{\partial v_t}{\partial C_t} = (1-\beta)(1-\rho)\left(\frac{C_t}{W_t}\right)^{-\rho}\frac{1}{W_t},\ \frac{\partial v_{t+1}}{\partial C_{t+1}} = (1-\beta)(1-\rho)\left(\frac{C_{t+1}}{W_{t+1}}\right)^{-\rho}\frac{1}{W_{t+1}} \tag{B-95}$$

and

$$\frac{\partial v_t}{\partial v_{t+1}} = \beta\left(\frac{\exp_q\left\{-\frac{1}{\theta}v_{t+1}\left(\frac{W_{t+1}}{W_t}\right)^{1-\rho}\right\}}{E_t\left[\exp_q\left\{-\frac{1}{\theta}v_{t+1}\left(\frac{W_{t+1}}{W_t}\right)^{1-\rho}\right\}\right]}\right)^q\left(\frac{W_{t+1}}{W_t}\right)^{1-\rho} \tag{B-96}$$

combining and deflating with consumption ($W_t = C_t$) yields the final form of the pricing kernel in the main text.

B.11 Worst Case Moments under q Measure for the Endowment Economy

	EZ Preferences $\theta = 0$, $q = 1.119 \to$DEP=4.25%				Generalized Uncertainty $q = 2$, $\theta = -6.29 \to$DEP=25%			
Variable	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t	Mean %	Std. Dev. %	Autocorr.	Corr. w/ x_t
$\Delta \ln c_t$	0.127	0.541	0.182	0.425	0.206	0.563	0.231	0.479
x_t	-0.367	0.234	0.986	1	-0.290	0.274	0.989	1
$\Delta \ln d_t$	-0.982	2.673	0.078	0.279	-0.731	2.732	0.103	0.320
σ_t	2.382E-03	1.233E-03	0.994	-0.270	2.427E-03	1.235E-03	0.994	-0.203

Worst case moments calculated under the g^q measure

Table 9: Worst Case Moments

B.12 The Hansen and Jagannathan (1997) Bounds for the Endowment Economy

Figure 5 depict the Hansen and Jagannathan (1997) bounds from the data and the maximum Sharpe ratios from the different specifications here. The case of expected utility clearly replicates Weil's (1989) risk-free rate puzzle, with the increase in the volatility of the pricing kernel associated with a decrease in its mean. Separating risk aversion and intertemporal substitution with the preferences of EZ or HS breaks this puzzle and moves vertically towards the bounds, here depicted with DEPs of 5 to 40%. The generalized model uncertainty is depicted in green circles with the DEPs fixed at 25% and approaches the bounds with q's above 2.

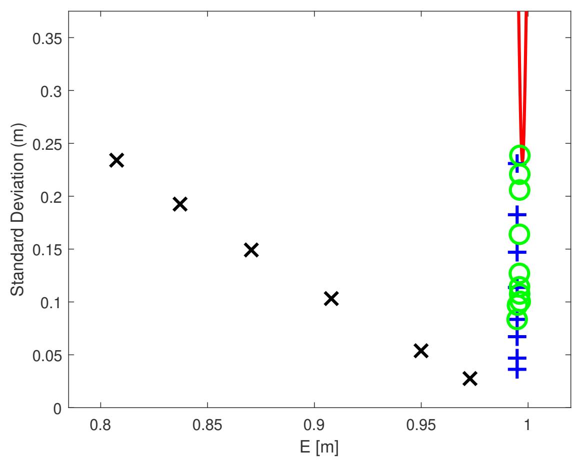


Figure 5: The Hansen-Jagannathan Bounds for the Endowment Economy
$\times$: Expected Utility; $+$: HS and EZ;
$\bigcirc$: Generalized Entropy, DEP$=25\%$ and $q = 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.75, 2, 2.25, 2.5$

B.13 Details of the Production Economy

The household's maximization is subject to

(B-97) $$C_t + K_t = W_t N_t + RR_t^K K_{t-1} + (1-\delta)K_{t-1}$$

where K_t is capital stock accumulated today for productive purpose tomorrow, W_t real wage, RR_t^K the capital rental rate and $\delta \in [0,1]$ the depreciation rate. Investment is the difference between the current capital stock and the capital stock in the previous period after depreciation

(B-98) $$I_t = K_t - (1-\delta)K_{t-1}$$

The stationarized resource constraint is

(B-99) $$c_t + k_t = y_t + (1-\delta)\exp(-a_t)\, k_{t-1}$$

where $y_t = e^{-\alpha a_t} k_{t-1}^{\alpha} N_t^{1-\alpha}$ follows from profit maximization, with the stationarized wage $w_t = (1-\alpha)\, e^{-\alpha a_t}$ $k_{t-1}^{\alpha} N_t^{-\alpha}$ and rental rate $RR_t = \alpha e^{-(1-\alpha)a_t} k_{t-1}^{\alpha-1} N_t^{1-\alpha}$ and the household's budget constraint

(B-100) $$c_t + k_t = w_t N_t + \left(1 - \delta + RR_t^K\right)\exp(-a_t)\, k_{t-1}$$

closing the model.

The calibration of the model follows Tallarini (2000) to maintain comparability (see the discussion there and the online appendix for the calibration common to all models here). The standard deviation of productivity growth σ_a is set to match the post-war U.S. consumption growth volatility in table 1 and the preference parameters, θ and q, are set using DEPs; see the discussion in section 4.1.

Parameter	β	ψ	α	δ	$\overline{a}$	σ_a
Value	0.9926	$\bar{N} = 0.2305$	0.339	0.021	0.004	Std. Dev. $\Delta \ln c_t = 0.566\%$

See Tallarini (2000) and the main text.

Table 10: Parameter Values

Table 10 contains the calibration of the model common to all specifications that follows Tallarini (2000).

B.14 Detailed Derivation of the Pricing Kernel for the Production Economy

(B-101) $$m_{t+1} \doteq \frac{\partial v_t / \partial C_{t+1}}{\partial v_t / \partial C_t} = \frac{\frac{\partial v_t}{\partial v_{t+1}} \frac{\partial v_{t+1}}{\partial c_{t+1}} e^{-Z_{t+1}}}{\frac{\partial v_t}{\partial c_t} e^{-Z_t}}$$

with

(B-102) $$\frac{\partial v_t}{\partial c_t} = \frac{1}{c_t},\ \frac{\partial v_{t+1}}{\partial c_{t+1}} = \frac{1}{c_{t+1}}$$

and

(B-103) $$\frac{\partial v_t}{\partial v_{t+1}} = \beta \left(\frac{\exp_q\left\{-\frac{1}{\theta}\left(v_{t+1} + \frac{1}{1-\beta} a_{t+1}\right)\right\}}{E_t\left[\exp_q\left\{-\frac{1}{\theta}\left(v_{t+1} + \frac{1}{1-\beta} a_{t+1}\right)\right\}\right]} \right)^q$$

combining yields the final form of the pricing kernel in the main text.

EZ Preferences

$\theta = 0$, DEP=25%$\rightarrow q = 1.082$

	Approximating Model			Worst-Case Model q		
Variable	Mean %	Std. Dev. %	Autocorr.	Mean %	Std. Dev. %	Autocorr.
$\Delta \ln Y_t$	0.400	1.023	0.009	0.291	1.023	0.012
$\Delta \ln C_t$	0.400	0.566	0.084	0.291	0.562	0.082
a_t	0.000	1.189	0.000	-0.109	1.187	0.004

Generalized Uncertainty

$q = 2$, DEP=25%$\rightarrow \theta = 132.15$

	Approximating Model			Worst-Case Model		
Variable	Mean %	Std. Dev. %	Autocorr.	Mean %	Std. Dev. %	Autocorr.
$\Delta \ln Y_t$	0.400	1.233	0.008	0.146	1.237	0.005
$\Delta \ln C_t$	0.400	0.566	0.101	0.146	0.562	0.119
a_t	0.000	1.347	0.000	-0.254	1.344	0.001

For all three specifications, σ_z is adjusted to match the empirical volatility of $\Delta \ln C_t$ in the approximating model and the free parameter in the preference specification to yield a detection error probability of 25%.

Table 11: Business Cycle Moments using the q measure

B.15 Worst Case Moments under q Measure for the Production Economy

B.16 Details of the Asset Pricing Implications of the Production Economy

Under the calibration in the previous section (specifically for DEPs of 25%), both HS's and EZ's specifications yield maximum Sharpe ratios of 0.1. This relation holds more generally, as can be seen in figure 6, which plots the maximum Sharpe ratios of the approximating models against the DEPs for the HS and EZ specifications. For a DEP of 0.25, both specifications yield roughly the same maximum Sharpe ratio of around 0.1. For very low DEPs the specification of EZ and for very high DEPs the specification of HS produces higher maximum Sharpe ratios. That these two different specifications yield very similar results when controlling for the DEPs confirms the close relation between these two different preference specifications carries over from the endowment economy studied above. Likewise confirmed from the endowment economy is that the combined distortion g^q_{t+1} for EZ preferences generates a much lower maximum Sharpe ratio for a given DEP than under the HS case, underlining the caveat that this close relation hinges on using only the change of measure g_{t+1} associated with EZ preferences.

Holding the DEP constant at 25%, the generalized model uncertainty present in this paper moves directly towards the bounds and enters them with a $q = 2.25$, as can be seen in table 12. For the $q = 2$ specification, the maximum Sharpe ratio is 0.21, just shy of the empirical Sharpe ratio of 0.2261, see

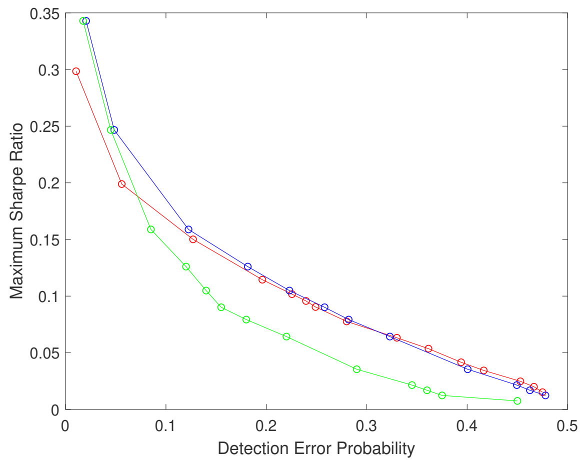


Figure 6: Maximum Sharpe Ratio and DEPs
Red: HS (g); Blue: EZ (g); Green: EZ (g^q)

$q =$	1	1.1	1.2	1.3	1.4	1.5	1.75	2	2.25	2.5
MSR	0.10	0.11	0.12	0.13	0.14	0.15	0.19	0.21	0.24	0.27

θ is adjusted to keep the DEP at 0.25.

Table 12: Entropic Index and the Maximum Sharpe Ratio

the lower half of table 1, and more than twice the value obtained under both HS's and EZ's specifications. That agents overweight the probability of pernicious distributions including the worst case under the generalized model uncertainty formulation drives up the returns on risky capital relative to the risk free bond.

Figure 7 contains the Hansen and Jagannathan (1997) bound for the assets in table 1 and both expected utility ($\theta = \infty$ and $q = 1$) and for recursive utility using the exponential certainty equivalent ($q = 1$ and varying θ). For the expected utility case, the risk-free rate puzzle can be seen through the decrease in $E[m_{t+1}]$ with risk aversion is increased from 5, 10, 20, 30, 40, 50, and finally to 100. By holding the elasticity of intertemporal substitution constant at one, Tallarini (2000) is able to march up to the bounds, but only for a degree of risk aversion equal to 100. Under the Hansen and Sargent (2005) interpretation of calibrating using model uncertainty, this degree of risk aversion is associated with a DEP of less than 5% for both HS and EZ, arguably past the limit of credulity. My generalized model uncertainty is again able to approach and then enter the bounds with q's above 2 for the DEPs fixed at 25%.

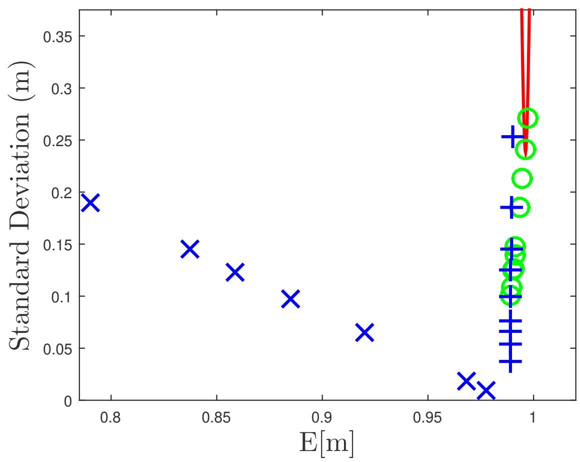


Figure 7: The Hansen-Jagannathan Bounds
$\times$: Expected Utility; $+$: HS and EZ;
$\bigcirc$: Generalized Entropy, DEP$=25\%$ and $q=1,1.1,1.2,1.3,1.4,1.5,1.75,2,2.25,2.5$

q	$q/(q-1)$	σ_M	$100\frac{\sigma_\Lambda^R}{\sigma_M}$	$100\frac{\sigma_\Lambda^U}{\sigma_M}$	$100\frac{\sigma_\Lambda^P}{\sigma_M}$
1.1	11	0.085	6.612	85.842	8.567
1.25	5	0.041	13.752	69.890	17.464
1.5	3	0.026	21.139	53.329	26.661
2	2	0.019	28.778	36.171	36.171

Table 13: Production: Risk and Pessimism under EZ