SFB 649 Discussion Paper 2017-017

# Generalized Entropy and Model Uncertainty

Alexander Meyer-Gohde *

* Universität Hamburg & Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

# Generalized Entropy and Model Uncertainty*

Alexander Meyer-Gohde†§

This Version: August 4, 2017

**Abstract**

I entertain a generalization of the standard Bolzmann-Gibbs-Shannon measure of entropy in multiplier preferences of model uncertainty. Using this measure, I derive a generalized exponential certainty equivalent, which nests the exponential certainty equivalent of the standard Hansen-Sargent model uncertainty formulation and the power certainty equivalent of the popular Epstein-Zin-Weil recursive preferences as special cases. Besides providing a model uncertainty rationale to these risk-sensitive preferences, the generalized exponential equivalent provides additional flexibility in modeling uncertainty through its introduction of pessimism into agents, causing them to overweight events made more likely in the worst case model when forming expectations. In a standard neoclassical growth model, I close the gap to the Hansen-Jagannathan bounds with plausible detection error probabilities using the generalized exponential equivalent and show that Hansen-Sargent and Epstein-Zin-Weil preferences yield comparable market prices of risk for given detection error probabilities.

*JEL classification*: C61, C63, E17

*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.

†Universität Hamburg, Professur für Volkswirtschaftslehre insb. Wachstum und Konjunktur, Von-Melle-Park 5, 20146 Hamburg, Germany; Tel.: +49-40-42838 3996; E-Mail: alexander.meyer-gohde@wiso.uni-hamburg.de

§Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; E-Mail: alexander.meyer-gohde@wiwi.hu-berlin.de

# 1 Introduction

Model uncertainty in macroeconomic models (see Hansen and Sargent (2001, 2010) and the detailed treatment in the monograph Hansen and Sargent (2007)) places agents in an 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. With agents making intertemporal decisions such as investment in an environment where they distrust the models they use to form expectations about the future, Barillas, Hansen, and Sargent (2009) show that a modest amount of model uncertainty can substitute for a high degree of risk aversion. Tallarini (2000), Barillas, Hansen, and Sargent (2009), and Ju and Miao (2012) among others have emphasized the close relationship between model uncertainty preferences and risk-sensitive preferences such as the popular Epstein and Zin (1989) and Weil (1990) recursive, constant elasticity preferences.[1] Yet an equivalence has only been demonstrated for the specific case of a unit elasticity of intertemporal substitution. This limitation arises due to the differing functional forms of the certainty equivalents in these preferences (exponential for Hansen and Sargent's (2007) model uncertainty and power for Epstein and Zin's (1989) and Weil's (1990) risk-sensitive preferences). Backus, Routledge, and Zin (2005) observe that it is an open question whether the power certainty equivalent underlying Epstein and Zin's (1989) and Weil's (1990) risk-sensitive preferences can be given a model uncertainty foundation that relates the two sets of preferences beyond the known special case.

In this paper, I propose an answer to this open question by generalizing 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. Alongside a generalized exponential certainty equivalent, I derive a power certainty equivalent from model uncertainty preferences and its associated worst-case distribution. With this distribution in hand, I can calibrate risk aversion in Epstein and Zin's (1989) and Weil's (1990) preferences using detection error probabilities as proposed by Anderson, Hansen, and Sargent (2003) and Hansen

[1] Hansen and Marinacci (2016) summarize the connection between Hansen and Sargent's (2007) multiplier preference approach to model uncertainty that I adopt here 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 their multiplier preference and Gilboa and Schmeidler's (1989) multiple priors. Ju and Miao's (2012) generalized smooth ambiguity preferences nests these variational preferences as special cases from a risk sensitive and ambiguity (vis-a-vis unobservable states) perspective.

and Sargent (2007). From the lens of model uncertainty, decreases in risk aversion in Epstein and Zin's (1989) and Weil's (1990) risk-sensitive preferences can be interpreted as a reduction in model uncertainty tempered by an increase in pessimism in the form of an overweighting of the probability of the worst case model. 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). In an application of subadditivity to investment, Dow and Werlang (1992) emphasize that expectations formed under probabilities that do not sum to one reflect both agent's uncertainty and aversion thereto.

Applying the preferences to a standard RBC model[2] under random walk with drift productivity and using the perturbation-based solution and sampling techniques of Bidder and Smith (2012), I find that both Hansen and Sargent's (2007) original formulation and the model uncertainty formulation for Epstein and Zin (1989) and Weil (1990) behave comparably for a given detection error probability with respect to both macroeconomic and asset pricing variables. Examining the worst case density associated with the different specifications, I find that agents with Hansen and Sargent's (2007) formulation fear autocorrelated productivity growth with a lower mean but reduced volatility,[3] those with Epstein and Zin (1989) and Weil (1990) preferences autocorrelated productivity growth with a higher mean but increased volatility, and those with the generalized model uncertainty preferences I introduce here autocorrelated productivity growth with a lower mean and increased volatility.

The remainder of the paper is organized as follows. In section 2, I formulate a general dynamic model and derive the specific conditions under which Epstein and Zin's (1989) and Weil's (1990) risk-sensitive preferences and Hansen and Sargent's (2007) model uncertainty are equivalent. I then turn to the measure of entropy behind model uncertainty and present the generalized measure in section 3. In section 4, I apply this measure to the general dynamic model, derive conditions that

[2] I follow Tallarini's (2000) specification of the RBC model and twist the continuation utility value according to the different certainty equivalents I derive here. 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.

[3] This result is broadly consistent with other studies: Barillas, Hansen, and Sargent (2009), Bidder and Smith (2012), Ellison and Sargent (2015), Bidder and Drew-Becker (2016) all find that the worst case is associated with lower mean growth.

recover both Epstein and Zin's (1989) and Weil's (1990) risk-sensitive preferences as well as Hansen and Sargent's (2007) 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 an otherwise standard RBC model in section 5 and examine the asset pricing and macroeconomic performance of all three frameworks. Section 6 concludes.

## 2 Dynamic Model

In this section, I will lay out a general dynamic model. I review the risk sensitive preferences of Epstein and Zin (1989) and Weil (1990) and the model uncertainty multiplier preferences of Hansen and Sargent (2007), as well as the conditions under which the two coincide.

I will consider a recursive dynamic model where a time-invariant transition density

$$p(x',x,a) \tag{1}$$

gives the joint distribution of the future state, $x' \in X$, 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, is determined by

$$\pi(x',x) \doteq p(x',x,a(x)) \tag{2}$$

the control variable, $a$, is chosen to maximize lifetime utility expressed recursively following Kreps and Porteus (1978) as

$$V(x) = \max_{a\in A} \mathcal{T}(u,\mathcal{R}(V))(x) \tag{3}$$

where $\mathcal{T}$ is a time aggregator and $\mathcal{R}$ a risk aggregator, or certainty equivalent.

The popular risk sensitive preference specification of Epstein and Zin (1989) and Weil (1990) is a constant elasticity time and risk preference formulation, given by

$$V(x) = \max_{a\in A}\left[(1-\beta)\,u(x,a(x))^{1-\rho} + \beta\left(\int V(x')^{1-\gamma}p(x',x,a(x))dx'\right)^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}} \tag{4}$$

where $\beta \in (0,1)$ is the discount factor and, with respect to $u(x,a(x))$, $\rho$ is the inverse of the intertemporal elasticity of substitution and $\gamma$ the coefficient of relative risk aversion.[4] In this case, $\mathcal{R}(V)(x)$

[4]Both of these measures are expressed here with respect to the period utility kernel $u(x,a(x))$ and are misnomers if $u(x,a(x)) \neq C(x)$, where $C(x)$ is the agent's current consumption. See especially, Swanson (2012a) and Swanson (2012b) for measures of relative risk aversion with alternative period utility kernels and under recursive preferences. I

is a power certainty equivalent $E\left[V(x')^{1-\gamma}|x\right]^{\frac{1}{1-\gamma}}$.

Standard expected utility can be recovered using the transformation $\tilde{V}(x) \doteq V(x)^{1-\rho}$) and the limiting case of $\gamma = \rho$

(5) $$\lim_{\gamma \to \rho} \tilde{V}(x) = \max_{a \in A} (1-\beta)\, u(x, a(x))^{1-\rho} + \beta \int \tilde{V}(x') p(x', x, a(x)) dx'$$

In this case, $\mathcal{R}(\tilde{V})(x)$ is the conditional expectations operator $E\left[\tilde{V}(x')|x\right] \doteq \int \tilde{V}(x') p(x', x, a(x)) dx'$.

The risk aggregator, $\mathcal{R}(V)(x)$, can also be given a model uncertainty interpretation using the tools of robust control following Hansen and Sargent (2007). In this approach, 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 utility approach, on value functions over a set of models. This set is constrained by limiting or 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 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 on the worst-case model.

Formally, an agent has preferences in the form of (3) given by

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

where the aggregator $\mathcal{R}(V)(x)$ is derived by considering an agent who entertains a distorted model

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

close to the approximating model, the probability distribution common to other specifications (2). The likelihood ratio between the distorted and approximating models is

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

and the discrepancy between the two models will be calculated as the expected value of this ratio, i.e., their relative entropy or the Kullback-Leibler divergence,

(9) $$\int \ln\left(g(x', x)\right) \tilde{p}(x', x, a(x)) dx'$$

The aggregator $\mathcal{R}$ results from a robustness consideration that selects the density for evaluating the

maintain this misnomer here for expositional expediency.

continuation value as[5]

$$(10)\quad \mathcal{R}(V)(x) \doteq \min_{\substack{\tilde{p}(x',x,a(x))\geq 0 \\ \int \tilde{p}(x',x,a(x))dx'=1}} \int V(x')g(x',x)p(x',x,a(x))dx' + \theta \int \ln\left(g(x',x)\right)\tilde{p}(x',x,a(x))dx'$$

This is Hansen and Sargent's (2007) multiplier preferences approach,[6] which tempers the agent's decisions against models that are pernicious (i.e., reduce her expected continuation value) yet plausible (i.e., are close to the baseline model in the sense of small relative entropy). The worst case model, $\tilde{p}$, that solves the minimization problem balances these two goals, where $\theta$ controls how much weight is assigned to the entropy goal. If this weight is infinite, $\tilde{p}$ is identical to $p$ and $\mathcal{R}$ becomes the conditional expectation operator.

For a finite $\theta$, however, the minimizing model, $\tilde{p}$, will differ from the approximating model, $p$. Rearranging the likelihood ratio, (8), the minimizing model can be expressed as

$$(12)\qquad \tilde{p}(x',x,a(x)) = g(x',x)p(x',x,a(x))$$

where the likelihood ratio, $g$, distorts the approximating model, $p$, to give the minimizing model $\tilde{p}$. Solving the minimization problem, (10), gives

$$(13)\qquad g(x',x) = \frac{\exp\left[-\frac{1}{\theta}V(x')p(x',x,a(x))\right]}{\int \exp\left[-\frac{1}{\theta}V(x')p(x',x,a(x))\right]dx'}$$

as the minimizing distortion. Here, future states $x'$ associated with a lower than average (under the approximating model, $p$) continuation value are assigned a higher probability ($g(x',x) > 1$) than under the approximating model and those $x'$ associated with a higher than average (again, under the approximating model) continuation value a lower probability ($g(x',x) < 1$) than under the approximating model. This distortion of the approximating probability measure is proportional to the expected continuation value, or an agent concerned with the robustness of her decisions operates under the hypothesis that "events occur with probabilities in inverse proportion to their desirability." Hansen and Sargent (2007), following Bucklew (2004), call this a "statistical version of Murphy's

[5]A Bellman-Isaacs condition enables the minimization and maximization operators to be interchanged in formulating the zero-sum game that underlies the selection of the minimizing density, see Hansen and Sargent (2007).

[6]More direct, yet, mathematically less expedient is the constraint preferences approach

$$(11)\qquad \mathcal{R}(V)(x) \doteq \min_{\int \ln(g(x',x))\tilde{p}(x',x,a(x))dx' \leq \eta} \int V(x')g(x',x)p(x',x,a(x))dx'$$

whereby the agent makes her decision rule robust to unstructured uncertainty contained inside the hyperball with a radius $\eta$ centered around her approximating model. $\eta$ thus measures the amount of uncertainty facing an agent. Hansen and Sargent (2001) provide conditions under which this constraint approach is evuivalent to the multiplier approach I use here.

Law." Substituting the minimizing distortion, $g$, back into the minimization problem, (10), gives

$$\mathcal{R}(V)(x) = -\theta \ln \int \exp\left[-\frac{1}{\theta}V(x')p(x',x,a(x))\right]dx' \tag{14}$$

an exponential certainty equivalent. With this certainty equivalent, (6) can be written as

$$V(x) = \max_{a\in A} u(x,a(x)) - \theta\beta \ln \int \exp\left[-\frac{1}{\theta}V(x')p(x',x,a(x))\right]dx' \tag{15}$$

Standard expected utility is recovered in the limiting case of $\theta \to \infty$

$$\lim_{\theta\to\infty} -\theta \ln \int \exp\left[-\frac{1}{\theta}V(x')p(x',x,a(x))\right]dx' = \int V(x')p(x',x,a(x))dx'$$

In this case, $\mathcal{R}(\tilde{V})(x)$ is the conditional expectations operator $E\left[\tilde{V}(x')|x\right] \doteq \int \tilde{V}(x')p(x',x,a(x))dx'$.

The recursive preferences of Epstein and Zin (1989) and Weil (1990) lead to a power certainty equivalent, see (4), whereas those of Hansen and Sargent (2007) lead to an exponential certainty equivalent, see (14). As has been demonstrated by, e.g., Tallarini (2000), 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 2.1.** *Logarithmic Equivalence of Risk Sensitive and Model Uncertainty Preferences*

*If the elasticity of intertemporal substitution in (4) is one, the period utilities are related through a logarithmic transformation*

$$u^{HS}(x,a(x)) = \ln\left(u^{EZ}(x,a(x))\right) \tag{16}$$

*and*

$$-\theta = \frac{1}{(1-\beta)(1-\gamma)} \tag{17}$$

*then*

$$V^{HS}(x) = \frac{1}{1-\beta}\ln\left(V^{EZ}(x)\right) \tag{18}$$

*Proof.* See the Appendix. □

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. Backus, Routledge, and Zin (2005) have pointed out that is an unresolved question how these two preference relate under more general settings. Addressing this question means finding a foundation that recovers both exponential and power certainty equivalents as special cases. I will take the model uncertainty perspective and accomplish exactly this by generalizing the measure of entropy used to compared alternate models.

# 3 Generalized Entropy

To provide a model uncertainty framework that moves beyond the exponential certainty equivalent of Hansen and Sargent (2007) demands that we move past the standard logarithmic relative entropy to measure the distance between two models. I follow the the physics literature on statistical mechanics and replace the standard Boltzmann-Gibbs-Shannon measure of entropy with the generalization introduced by Tsallis (1988). After introducing the basic properties and intuition, I turn 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 entropy

$$S_1(p(x)) \doteq -\int p(x)\ln p(x)dx \tag{19}$$

where the meaning of the subscript in $S_1$ will become apparent shortly, is used in the context of information theory, see, e.g., Cover and Thomas (1991), as a measure of the expected information content[7] of a realization from the distribution $p(x)$—that is, the expected surprisal or unpredictability of a distribution.

The uniqueness theorems of Shannon and Khinchin[8] provide an axiomatic foundation for the function in (19) and prove that its functional form uniquely satisfies their set of axioms. If their axioms are modified to pseudoadditivity[9] 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, introduced by Tsallis (1988), is given by

$$S_q(p(x)) \doteq -\int \left(\frac{1-p(x)^q}{1-q}\right)dx = -\int p(x)^q \ln_q p\left(x\right)dx \tag{20}$$

where the generalized $q$-logarithm, $\ln_q$, is defined as

$$\ln_q\left(x\right) \doteq \frac{x^{1-q}-1}{1-q} \tag{21}$$

It is useful to define the inverse function of $\ln_q$, the generalized $q$-exponential function,

$$\exp_q\left(x\right) \doteq \left[1+\left(1-q\right)x\right]^{\frac{1}{1-q}} \tag{22}$$

Note that both (21) and (22) can be extended over their removable singularities at $q = 1$ to give the standard base $e$ logarithm and exponential function as limiting cases, $\ln_1(x) = \ln(x)$ and $\exp_1(x) =$

[7]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.

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

[9]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 $\lim_{q\to 1} S_q(A+B) = S_1(A)+S_1(B)$.

$\exp(x)$. Thus, Tsallis's (1988) entropy recovers (19) as a limiting case, generalizing Boltzmann-Gibbs-Shannon entropy.

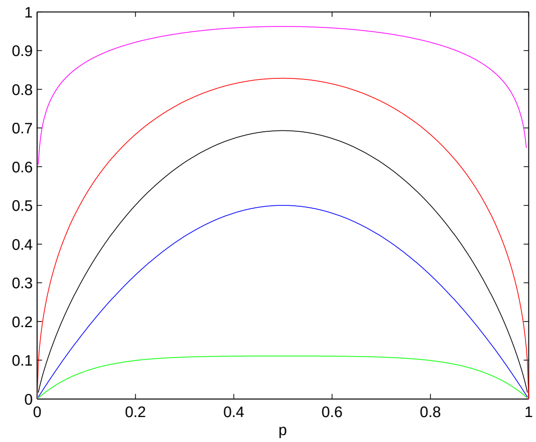
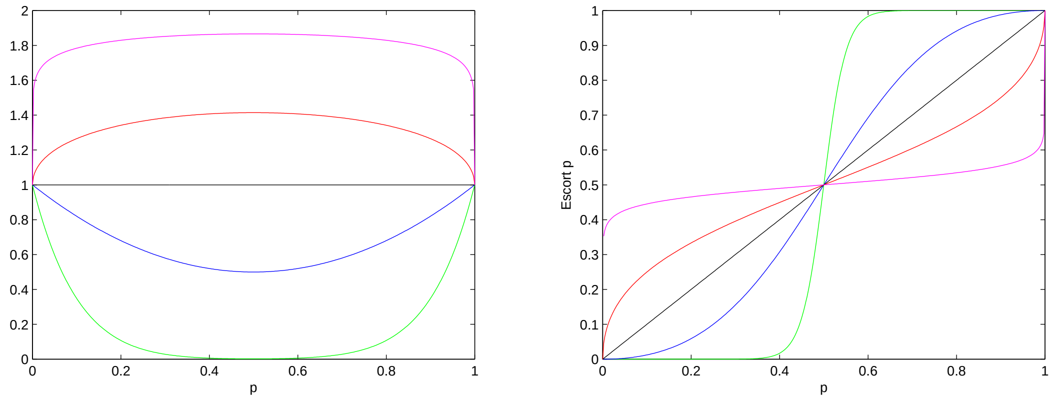



Figure 1: $q$ Entropy or Generalized Expected Surprise
magenta—$q = 0.1$, red—$q = 0.5$, black—$q = 1$, blue—$q = 2$, green—$q = 10$

Figure 1 depicts the generalized entropy (20) for a two state system.[10] The first feature to note is that entropy is concave for all the values of $q$ depicted here; more generally, (20) is concave for $q > 0$ and convex for $q < 0$, see Tsallis (1988) and Tsallis (2009, Ch.3). When the probability of either of the two states is one ($p = 0$ or $p = 1$), entropy is zero as the probability one event will happen with certainty and there is, thus, no expected surprisal. Note that this holds regardless of the value of the entropic index, $q$. As can be seen in figure 1, the expected surprisal is decreasing in $q$; that is, if $q > 1$ then entropy is less than in the standard Boltzmann-Gibbs-Shannon case and if $q < 1$ entropy is greater. The entropic index 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, $0 \leq p_i \leq 1$, $p_i^q \geq p_i$ for $q < 1$ and $p_i^q \leq p_i$ 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 2a 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),

[10] 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).

$q > 1$ can be interpreted as a situation of uncertainty from the perspective of objective probabilities.




(a) Total Probability (b) Probabilities of Escort Distribution

Figure 2: Biased Probabilites
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

$$p_q(x) \doteq \frac{p(x)^q}{\int p(x)^q dx} \tag{23}$$

which normalizes the biased probabilities by the total probability from above. For the two state system, figure 2b 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 $p(x)$

$$E^p[x] \doteq \int x p(x) dx \tag{24}$$

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

$$E^p_q[x] \doteq \int x \frac{p(x)^q}{\int p(x)^q dx} dx \tag{25}$$

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

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

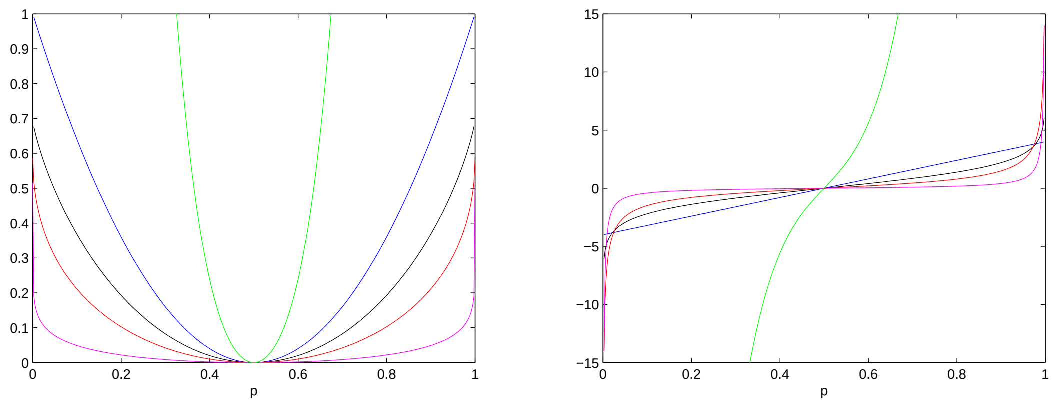
with respect to the reference distribution $p(x)$

(26)
$$I_1\left(\tilde{p}(x), p(x)\right) \doteq \int \tilde{p}(x) \ln \frac{\tilde{p}(x)}{p(x)} dx$$

provides a consistent method of discriminating between two probability distributions by quantifying distance between the two distributions.[11] This can be $q$-generalized following Tsallis (1988), Abe and Bagci (2005), and Tsallis (2009, Ch. 3) as

(27)
$$I_q\left(\tilde{p}(x), p(x)\right) \doteq \int p(x) \left(\frac{\tilde{p}(x)}{p(x)}\right)^q \ln_q \left(\frac{\tilde{p}(x)}{p(x)}\right) dx$$

and is positive and convex (both jointly and individually in $\tilde{p}(x)$ and $p(x)$, see Abe and Bagci (2005), for $q > 0$.




(a) $q$-Relative Entropy

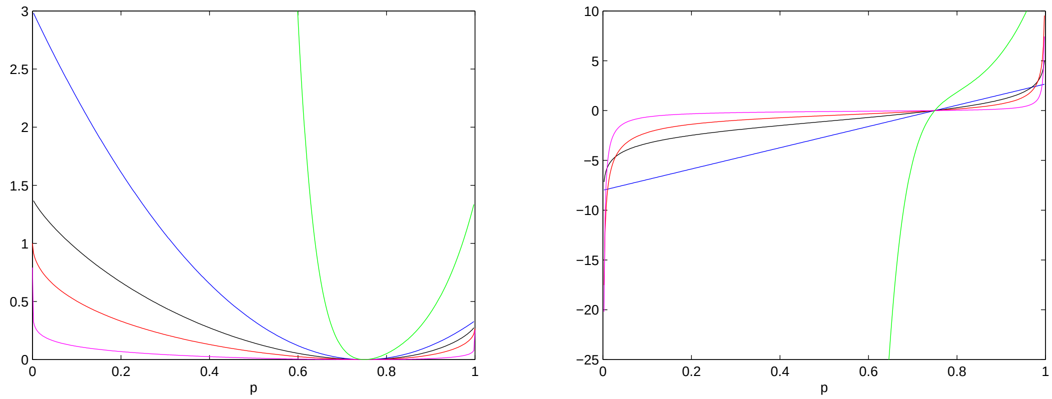
(b) $\frac{\partial}{\partial \tilde{p}(x)}$ of $q$-Relative Entropy

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

Figure 3a plots (27) 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 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, but now $p = 0.75$, as can be deduced by the point of zero relative

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

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.




(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

Again, (27) is a generalization of the standard measure, and the Tsallis (1988) $q$ measure generalizes the standard measure, the relative entropy or the Kullback-Leibler divergence, of discriminating between two distributions.

# 4 Generalized Multiplier Preferences

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, of course, takes this into account when formulating her decision function. My generalization replaces Hansen and Sargent's (2005) and Hansen and Sargent's (2007) Boltzmann-Gibbs-Shannon measure of entropy with the generalized form in (27) from the previous section and allows for a state-dependent weight on the entropy penalty,

$$\mathcal{R}(V)(x) \doteq \min_{\substack{\tilde{p}(x',x,a(x))\geq 0 \\ \int \tilde{p}(x',x,a(x))dx'=1}} \int V(x')\tilde{p}(x',x,a(x))dx'$$

$$+\int \theta(x')\left(\frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}\right)^{q-1}\ln_q\left(\frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}\right)\tilde{p}(x',x,a(x))dx' \tag{28}$$

The first term evaluates continuation utility, conditioning on the current state $x$, under the distorted density. The second term is the generalized relative entropy, conditional on $x$, of the distorted density to the approximating model, reweighted with $\theta(x')$. Indeed, if $\theta(x')$ is independent of $x'$, say $\theta(x') = \overline{\theta}$, this term becomes $\theta I_q\left(\tilde{p}(x',x,a(x)), p(x',x,a(x))|x\right)$.

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

$$\mathcal{R}(V)(x) \doteq \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)p(x',x,a(x))dx'$$

$$+\int \theta(x')g(x',x)^q\ln_q\left(g(x',x)\right)p(x',x,a(x))dx' \tag{29}$$

The likelihood ratio can apparently be interpreted as a distortion to the probability density of the approximating model and distortions are penalized by their entropy weighted by the approximating density. This minimization problem weighs two countervailing forces: the decision maker would like to guard against very painful distortions (those that result in the smallest expected value of her continuation utility, $\int V(x)g(x',x)p(x',x,a(x))dx$); on the other hand, 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 ($\int p(x',x,a(x))g(x',x)\ln g(x',x)dx$), where $\theta(x')$ weights her concern for closeness. Thus, the decision maker is worried that her misspecification is both pernicious and hard to detect.

Specifically, I will set the multiplier, $\theta(x')$, equal to a constant and a term proportional to the continuation utility.

**Assumption 4.1.** *Entropy Multiplier*

*The multiplier* $\theta(x')$ *is given by*

$$\theta(x') \doteq \theta + (q-1)V(x') \tag{30}$$

*where* $\theta$ *and* $q$ *are positive.*

For $q > 1$, this 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 $(q-1)V(x')$ which tilts the minimizing agent's decision further towards pernicious 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.

This assumption on the multiplier allows me to reformulate the zero-sum game expressed in terms of the likelihood ratio, $g(x',x)$, as the sum of an entropy penalty with a constant multiplier and a continuation value evaluated under a weighted worst case density

(31)

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

(32)

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

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 in the second line), but also expresses a form of pessimism. The formulation of Hansen and Sargent (2005) and others with standard Boltzmann-Gibbs-Shannon entropy would set this power to 1, yielding expectations taken with respect to the distorted density $\tilde{p}(x',x,a)$. For $q>1$, events made more likely under the worst case density 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 the distorted model chosen to minimize their continuation utility. In this sense, I interpret $q$ as a measure of agents' pessimism. The resulting minimizing probability distortion is contained in the following

**Proposition 4.2.** *Minimizing Distortion and Risk-Sensitive Operator*

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

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

*and the risk aggregator, or certainty equivalent, by*

$$\mathcal{R}(V)(x) = -\,\theta\ln_q\left[\int \exp_q\left(-\frac{1}{\theta}V(x')\right)p(x',x,a(x))dx'\right] \tag{34}$$

$$= \frac{\theta-\left[\int\left(\theta-(1-q)V(x')\right)^{\frac{1}{1-q}}p(x',x,a(x))dx'\right]^{1-q}}{1-q} \tag{35}$$

*Proof.* See the Appendix. □

Thus, the varying multiplier and generalized entropy lead to a generalized exponential transformation governed jointly by the entropic index $q$ and static multiplier $\theta$ for the risk aggregator. This contrasts with the standard exponential transformation controlled by the static multiplier $\theta$ that results from Hansen and Sargent's (2007) formulation and the power certainty equivalent from Epstein and Zin (1989) and Weil (1990). The interpretation of this generalized form follows more readily from the special cases that capture these two specific preferences.

## 4.1 Equivalence with Hansen-Sargent Multiplier Preferences

In the extensive limit of the multiplier, $\lim_{q\to1}\theta(x')=\theta$, the model uncertainty specification and Hansen and Sargent (2007) is recovered

$$\lim_{q\to1}\mathcal{R}(V)(x) = -\theta\ln\left[\int\exp\left(-\frac{1}{\theta}V(x')\right)p(x',x,a(x))dx'\right] \tag{36}$$

with an exponential certainty equivalent following proposition 4.2 and a minimizing distortion

$$g^{HS}(x',x) = \frac{\exp\left(-\frac{1}{\theta}V(x')\right)}{\exp\left(-\frac{1}{\theta}\mathcal{R}(V)(x)\right)} \tag{37}$$

that tilts the distorted model using the standard exponential function.

This formulation is Hansen and Sargent's (2007) aggregator,

$$\mathcal{R}(V)(x) \doteq \min_{\substack{\tilde{p}(x',x,a(x))\geq0\\ \int\tilde{p}(x',x,a(x))dx'=1}} E^{\tilde{p}}\left[V(x')|x\right]+\theta I_1\left(\tilde{p}(x',x,a(x)),p(x',x,a(x))|x\right) \tag{38}$$

$$= \min_{\substack{\tilde{p}(x',x,a(x))\geq0\\ \int\tilde{p}(x',x,a(x))dx'=1}} \int V(x')\tilde{p}(x',x,a(x))dx'+\theta\int\tilde{p}(x',x,a(x))\ln\frac{\tilde{p}(x',x,a(x))}{p(x',x,a(x))}dx' \tag{39}$$

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

(40)

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

(41)

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

From the perspective of (31), 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.

## 4.2 Equivalence with Epstein-Zin-Weil Risk Sensitive Preferences

In the proportional limit of the multiplier, $\lim_{q\to 1}\theta(x') = \theta$, the risk sensitive specification of Epstein and Zin (1989) and Weil (1990) is recovered

$$\lim_{\theta\to 0} \mathcal{R}(V)(x) = \left[\int V(x')^{\frac{1}{1-q}} p(x',x,a(x))dx'\right]^{1-q} \tag{42}$$

with a power certainty equivalent. 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 (4), $\gamma$, is related to $q$ through $\gamma = -\frac{q}{1-q}$ and 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 4.2 the minimizing distortion associated with Epstein-Zin-Weil preferences is

$$g^{EZW}(x',x) = \left(\frac{V(x')}{\mathcal{R}(V)(x)}\right)^{\frac{1}{1-q}} = \left(\frac{V(x')}{\mathcal{R}(V)(x)}\right)^{1-\gamma} \tag{43}$$

a power tilting instead of the exponential tilting of Hansen-Sargent preferences. Having this minimizing distortion will enable me to parameterize their measure of relative risk aversion, $\gamma$, in Epstein-Zin-Weil preferences from a model uncertainty perspective using detection error probabilities.

From the perspective of (31), note that the $\theta = 0$ specification of Epstein and Zin (1989) and Weil (1990) gives

$$\mathcal{R}(V)(x) = \min_{\substack{g(x',x)>0 \\ \int g(x',x)p(x',x,a(x))dx'=1}} \int V(x')g(x',x)^{q-1}\tilde{p}(x',x,a(x))dx' \tag{44}$$

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.

## 4.3 Atemporal Risk Aversion

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 4.2 is

$$\mathcal{R}(V) = -\theta \ln_q \left( \int \exp_q \left( -\frac{1}{\theta} V(x) \right) p(x) dx \right) \tag{45}$$

and its minimizing density distortion is

$$g(x) = \frac{\exp_q \left( -\frac{1}{\theta} V(x) \right)}{\exp_q \left( -\frac{1}{\theta} \mathcal{R}(V) \right)} \tag{46}$$

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 $\sigma$. The certainty equivalent is

$$\mathcal{R}(V) = -\theta \ln_q \left( 0.5 \exp_q \left( -\frac{1+\sigma}{\theta} \right) + 0.5 \exp_q \left( -\frac{1-\sigma}{\theta} \right) \right) \tag{47}$$

which I will evaluate locally around $\sigma = 0$ out to second order[12]

$$\begin{aligned} \mathcal{R}(V) &\approx \mathcal{R}(V)\Big|_{\sigma=0} + \frac{\partial \mathcal{R}(V)}{\partial \sigma}\Big|_{\sigma=0} + \frac{1}{2} \frac{\partial^2 \mathcal{R}(V)}{\partial \sigma^2}\Big|_{\sigma=0} \\ &= 1 - \frac{q}{\theta + q - 1} \frac{\sigma^2}{2} \end{aligned} \tag{48}$$

As there is no term linear in $\sigma$, 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 term

$$\frac{q}{\theta + q - 1} \tag{49}$$

provides a measure of risk aversion.

In the special case of a power certainty equivalent following Epstein and Zin's (1989) and Weil's

[12] Details of the calculations can be found in the Appendix.

(1990) risk-sensitive preferences, θ is set to zero and the foregoing measure of risk aversion is

(50) $$\left.\frac{q}{\theta+q-1}\right|_{\theta=0}=-\frac{q}{1-q}$$

Which, through comparison with (4) is equal to γ, the coefficient of relative risk aversion.

For the exponential certainty equivalent of Hansen and Sargent's (2007) robust control approach, the entropic index $q$ is set to one, which delivers the following measure of risk aversion

(51) $$\left.\frac{q}{\theta+q-1}\right|_{q=1}=\frac{1}{\theta}$$

See also Hansen and Sargent (2007) and Tallarini (2000).

Returning to the general case in (49), the measure of risk aversion is increasing in θ for $q>0$

(52) $$\frac{\partial\frac{q}{\theta+q-1}}{\partial\theta}=\frac{q}{(\theta+q-1)^2}$$

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

(53) $$\frac{\partial\frac{q}{\theta+q-1}}{\partial q}=-\frac{1-\theta}{(\theta+q-1)^2}$$

## 4.4 Asset Pricing

Consider a household seeking to maximize the following preferences Following

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

where $V_t$ is the households 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

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

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

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

with

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

and

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

combining yields the final form of the pricing kernel

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

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 ($\theta = \infty$), $\Lambda_{t+1}^{U} \doteq g_{t+1}$ is the change of measure under the distorted model, and $\Lambda_{t+1}^{P} \doteq g_{t+1}^{q-1}$ captures the direct effect[13] of the entropic index.

Note that if $q = 1$, $\Lambda_{t+1}^{P}$ is equal to unity and the model uncertainty concerns collapse to Hansen and Sargent's (2007) original formulation (see section 4.1 above). For $q > 1$, agents overweight (underweight) states that have become more (less) likely under the distorted model when pricing assets, embedding a form of pessimism into a non-unity $\Lambda_{t+1}^{P}$. Thus, along with Hansen and Sargent's (2007), Bidder and Smith's (2012), and others' interpretation of $std_t\left(\Lambda_{t+1}^{R}\right)/E_t\left[\Lambda_{t+1}^{R}\right]$ and $std_t\left(\Lambda_{t+1}^{U}\right)$ as the market prices of risk and model uncertainty, respectively, I interpret $std_t\left(\Lambda_{t+1}^{P}\right)/E_t\left[\Lambda_{t+1}^{P}\right]$ as the market price of pessimism.

For Epstein and Zin's (1989) and Weil's (1990) power certainty equivalent, $\theta \to 0$ (see section 4.2 above), and all three components of the stochastic discount factor remain. As the measure of risk aversion is related inversely to $q$ in this case, see section 4.3, an increase in risk aversion is associated with a decrease in pessimism, as $\Lambda_{t+1}^{P}$ approaches unity,

# 5 Business Cycles, Asset Prices, and Model Uncertainty

In this section, I apply the generalized entropy constraint to a stochastic neoclassical growth model with a preference for robustness. I will parameterize the model closely to the production model described in Tallarini (2000). 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

$$U_t = \ln C_t + \psi \ln(1 - N_t) \tag{60}$$

subject to

$$C_t + K_t = W_t N_t + RR_t^K K_{t-1} + (1-\delta)K_{t-1} \tag{61}$$

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

$$I_t = K_t - (1-\delta)K_{t-1} \tag{62}$$

I will assume a perfectly competitive production side of the economy, where output is produced

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

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

$$a_t \equiv Z_t - Z_{t-1} = \overline{a} + \varepsilon_{z,t}, \;\; \varepsilon_{z,t} \sim \mathcal{N}(0,\sigma_z^2) \tag{63}$$

with $\varepsilon_{z,t}$ the innovation to $Z_t$.

The model is detrended with $\begin{bmatrix} y_t & k_t & i_t & c_t & w_t \end{bmatrix} \doteq e^{-Z_t}\begin{bmatrix} Y_t & K_t & I_t & C_t & W_t \end{bmatrix}$, where detrended variables are written in lowercase.

The household's lifetime utility function is expressed recursively using the generalized risk aggregator $\mathcal{R}(V)(x)$ as

$$v_t = \ln c_t + \psi \ln(1-N_t) + \beta \mathcal{R}\left(v_{t+1} + \frac{1}{1-\beta}a_{t+1}\right) \tag{64}$$

$$= \ln c_t + \psi \ln(1-N_t) - \beta\theta \ln_q \left\{ E_t \left[ \exp_q \left\{ -\frac{1}{\theta}\left(v_{t+1} + \frac{1}{1-\beta}a_{t+1}\right)\right\}\right]\right\} \tag{65}$$

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

$$\frac{\psi}{1-N_t} = \frac{1}{c_t}w_t \tag{66}$$

and the second is the intertemporal Euler equation, rearranged as the fundamental asset pricing equation,

$$1 = E_t\left[m_{t+1}R_{t+1}\right] \tag{67}$$

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 pricing kernel (see section 4.4), is given by

$$m_{t+1} \doteq \frac{\partial v_t / \partial C_{t+1}}{\partial v_t / \partial C_t} = \frac{\frac{\partial v_t}{\partial v_{t+1}} \frac{\partial v_{t+1}}{\partial c_{t+1}} e^{Z_{t+1}}}{\frac{\partial v_t}{\partial c_t} e^{Z_t}} \tag{68}$$

with

$$\frac{\partial v_t}{\partial c_t} = \frac{1}{c_t}, \; \frac{\partial v_{t+1}}{\partial c_{t+1}} = \frac{1}{c_{t+1}} \tag{69}$$

and

$$\frac{\partial v_t}{\partial v_{t+1}} = \beta \left( \frac{\exp_q\left\{-\frac{1}{\theta}\left(v_{t+1} + \frac{1}{1-\beta}a_{t+1}\right)\right\}}{E_t\left[\exp_q\left\{-\frac{1}{\theta}\left(v_{t+1}+\frac{1}{1-\beta}a_{t+1}\right)\right\}\right]}\right)^q \tag{70}$$

combining yields the final form of the pricing kernel

$$m_{t+1} = \beta \frac{c_t}{c_{t+1}} e^{a_{t+1}} \left( \frac{\exp_q \left\{ -\frac{1}{\theta} \left( v_{t+1} + \frac{1}{1-\beta} a_{t+1} \right) \right\}}{E_t \left[ \exp_q \left\{ -\frac{1}{\theta} \left( v_{t+1} + \frac{1}{1-\beta} a_{t+1} \right) \right\} \right]} \right)^q \tag{71}$$

The stationarized resource constraint is

$$c_t + k_t = y_t + (1-\delta) \exp(-a_t) k_{t-1} \tag{72}$$

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

$$c_t + k_t = w_t N_t + \left(1 - \delta + RR_t^K\right) \exp(-a_t) k_{t-1} \tag{73}$$

closes the model.

I append the model with the following additional asset pricing variables: the real risk-free rate $R_t^f \equiv E_t(m_{t+1})^{-1}$ and the (ex post) risk premium $rp_t = R_t - R_{t-1}^f$ as the difference between the risky and risk-free rate.

## 5.1 Data and Model Calibration

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

The Sharpe ratio and the market price of risk $\frac{std(m_{t+1})}{E[m_{t+1}]}$ 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 equation (here: $1 = E_t[m_{t+1} R_{t+1}]$ for the risky and $1 = E_t[m_{t+1}] R_t^f$ the risk free return) as

$$\frac{\left| E\left[R_{t+1} - R_t^f\right] \right|}{std\left(R_{t+1} - R_t^f\right)} \leq \frac{std(m_{t+1})}{E[m_{t+1}]} \tag{74}$$

with the Sharpe ratio on the left hand side being empirically observable and given in the lower half of table 1.

Table 2 contains the calibration of the model common to all specifications, where I follow Tallarini (2000) to maintain comparability (see the discussion there). The standard deviation of productivity growth $\sigma_a$ is set to match the post-war U.S. consumption growth volatility in table 1. The remaining parameters, $\theta$ and $q$, will be set using detection error probabilities, following Hansen and

Table 1: Data Moments, 1948:2-2012:4

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 |
| $\Delta \ln I_t$ | 0.004 | 2.536 | 2.558 | 0.336 | 0.248 | 0.043 | 0.662 |
| $\Delta \ln N_t$ | 0.328 | 1.192 | 1.203 | -0.020 | -0.010 | -0.008 | 0.388 |
| $\ln N_t$ | — | 2.778 | 2.802 | 0.999 | 0.998 | 0.997 | -0.139 |
| $\ln C_t - \ln Y_t$ | -0.611 | 5.887 | 5.938 | 0.990 | 0.978 | 0.964 | -0.172 |
| $\ln I_t - \ln Y_t$ | -1.382 | 7.302 | 7.365 | 0.962 | 0.910 | 0.841 | 0.128 |

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 business cycle data was retrieved from the Federal Reserve Economic Data (FRED) database of the Federal Reserve Bank of St. Louis.
All returns are measured as net real quarterly percentage returns.
$R$ 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.

Table 2: Parameter Values

| Parameter | $\beta$ | $\psi$ | $\alpha$ | $\delta$ | $\bar{a}$ | $\sigma_a$ |
|---|---|---|---|---|---|---|
| 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.

Sargent (2007). Specifically, I will use a perturbation solution of the model following Bidder and Smith (2012), but will use the nonlinear moving average policy function of Lan and Meyer-Gohde (2013c) to maintain the stability of the model under nonlinearity.[14] As proposed by Bidder and Smith (2012), I will 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 $\tilde{p}$. Second, I will generate simulations from the distorted model using a sampling importance resampling algorithm and then perform a symmetrical likelihood test.[15]

[14] 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 can be used to initialize the particle filters.

[15] The likelihood calculations are performed by sequential importance sampling-resampling, or particle filtering, with

A value of 0.5 for the detection error probability 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 detection error probability of between 0.15 and 0.2 as lower bound. I will take a conservative perspective and target a detection error probability of 0.25.

## 5.2 Macroeconomic Implications

I begin by comparing the business cycle properties of model uncertainty following Hansen and Sargent (2007) with $q = 1$ and the risk sensitive recursive utility specification of Epstein and Zin's (1989) and Weil's (1990) parameterized via model uncertainty with $\theta = 0$, before turning to the case of the generalized model uncertainty. The calibrations follow the discussion above, where the parameters $q$ and $\theta$ are set according to the specification chosen and to achieve a detection error probability of 0.25 between the approximating and worst case models of each specification. For the generalized model uncertainty case, $q$ is set to 2 (the reason for which will be clear in the next section that addresses asset pricing implications) and $\theta$ is then set to match the detection error probability. 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".

In the upper half of table 3, the business cycle moments for the approximating model are presented for the Hansen and Sargent (2007) specification ($q = 0$) with a detection error probability of 0.25 (which requires $\theta = 15$). The approximating model does a reasonable job in matching the post war U.S. macroeconomic experience, as can be seen by comparing with table 1.

The statistics of the worst case model that agents apparently fear can be found in lower half of table 3. Compared to the approximating model, it can be seen that agents worry about an environment with lower average growth and positive autocorrelation in technology growth. This is a familiar result of the model uncertainty framework, see, e.g., Barillas, Hansen, and Sargent (2009) with the long run risk result echoed by Bidder and Drew-Becker (2016). The detectability of the

a bootstrap proposal except where noted.

Table 3: Business Cycle Moments, Hansen and Sargent (2007) Preferences

Approximating Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta\ln Y_t$ | Cross Corr. w$\Delta\ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta\ln Y_t$ | 0.004 | 1.029 | 1.000 | 0.009 | 0.008 | 0.008 | 1.000 | 1.000 |
| $\Delta\ln C_t$ | 0.004 | 0.566 | 0.550 | 0.085 | 0.080 | 0.076 | 0.988 | 0.984 |
| $\Delta\ln I_t$ | 0.004 | 2.351 | 2.285 | -0.019 | -0.018 | -0.017 | 0.994 | 0.996 |
| $\Delta\ln N_t$ | 0.000 | 0.367 | 0.357 | -0.025 | -0.024 | -0.023 | 0.983 | 0.988 |
| $\ln N_t$ | -1.463 | 1.176 | 1.143 | 0.951 | 0.904 | 0.859 | 0.332 | 0.308 |
| $\ln C_t - \ln Y_t$ | -0.308 | 1.530 | 1.487 | 0.951 | 0.904 | 0.859 | -0.332 | -0.308 |
| $\ln I_t - \ln Y_t$ | -1.330 | 4.271 | 4.152 | 0.951 | 0.904 | 0.859 | 0.331 | 0.307 |
| $\Delta\ln a_t$ | 0.000 | 1.194 | 1.160 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |

Worst-Case Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta\ln Y_t$ | Cross Corr. w$\Delta\ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta\ln Y_t$ | 0.003 | 1.031 | 1.000 | 0.078 | 0.077 | 0.077 | 1.000 | 1.000 |
| $\Delta\ln C_t$ | 0.003 | 0.564 | 0.547 | 0.269 | 0.265 | 0.263 | 0.988 | 0.984 |
| $\Delta\ln I_t$ | 0.003 | 2.413 | 2.341 | -0.004 | -0.004 | -0.003 | 0.994 | 0.996 |
| $\Delta\ln N_t$ | 0.000 | 0.370 | 0.359 | -0.024 | -0.024 | -0.022 | 0.984 | 0.988 |
| $\ln N_t$ | -1.471 | 1.195 | 1.160 | 1.000 | 1.000 | 1.000 | 0.328 | 0.305 |
| $\ln C_t - \ln Y_t$ | -0.298 | 1.552 | 1.506 | 1.000 | 1.000 | 1.000 | -0.329 | -0.305 |
| $\ln I_t - \ln Y_t$ | -1.357 | 4.497 | 4.363 | 1.000 | 1.000 | 1.000 | 0.328 | 0.304 |
| $\Delta\ln a_t$ | -0.001 | 1.193 | 1.157 | 0.009 | 0.008 | 0.009 | 0.934 | 1.000 |

$\theta$ was set to 15 to deliver a detection error probability of 25%

worst case model with negative mean, positively autocorrelated technology growth is balanced with a reduction in the volatility of technology shocks.

In the upper half of table 4, the business cycle moments for the approximating model are presented for the Epstein and Zin (1989) and Weil (1990) specification ($\theta = 0$) with a detection error probability of 0.25 (which requires $q = 1.15$). The results here are essentially identical to those obtained under the approximating model under Hansen and Sargent's (2007) standard model uncertainty framework.

The lower half of table 4 contains the business cycle statistics of the worst case under the model uncertainty foundation for the Epstein and Zin (1989) and Weil (1990) specification. In contrast to the worst case under Hansen and Sargent's (2007) standard model uncertainty framework, agents here fear a technology process with increasing autocorrelations and a more volatile shock. This leads to substantial increases in the autocorrelations of macroeconomic variables and an increase in the volatility of consumption growth. The detectability of the worst case model is now balanced with an increase in the average growth rate of the economy.

Table 4: Business Cycle Moments, Epstein and Zin (1989) Preferences

Approximating Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta \ln Y_t$ | Cross Corr. w$\Delta \ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta \ln Y_t$ | 0.004 | 1.026 | 1.000 | 0.009 | 0.008 | 0.008 | 1.000 | 1.000 |
| $\Delta \ln C_t$ | 0.004 | 0.566 | 0.552 | 0.085 | 0.080 | 0.076 | 0.988 | 0.984 |
| $\Delta \ln I_t$ | 0.004 | 2.360 | 2.301 | -0.019 | -0.018 | -0.017 | 0.994 | 0.996 |
| $\Delta \ln N_t$ | 0.000 | 0.365 | 0.356 | -0.025 | -0.024 | -0.022 | 0.983 | 0.988 |
| $\ln N_t$ | -1.467 | 1.166 | 1.137 | 0.951 | 0.904 | 0.859 | 0.333 | 0.309 |
| $\ln C_t - \ln Y_t$ | -0.304 | 1.516 | 1.478 | 0.951 | 0.904 | 0.859 | -0.333 | -0.309 |
| $\ln I_t - \ln Y_t$ | -1.341 | 4.298 | 4.190 | 0.951 | 0.904 | 0.859 | 0.333 | 0.308 |
| $\Delta \ln a_t$ | 0.000 | 1.191 | 1.161 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |

Worst-Case Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta \ln Y_t$ | Cross Corr. w$\Delta \ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta \ln Y_t$ | 0.005 | 1.025 | 1.000 | 0.207 | 0.206 | 0.208 | 1.000 | 1.000 |
| $\Delta \ln C_t$ | 0.005 | 0.568 | 0.554 | 0.496 | 0.493 | 0.492 | 0.988 | 0.984 |
| $\Delta \ln I_t$ | 0.005 | 2.307 | 2.250 | 0.029 | 0.029 | 0.033 | 0.994 | 0.996 |
| $\Delta \ln N_t$ | 0.000 | 0.362 | 0.353 | -0.026 | -0.025 | -0.021 | 0.983 | 0.987 |
| $\ln N_t$ | -1.460 | 1.145 | 1.117 | 1.000 | 1.000 | 1.000 | 0.337 | 0.312 |
| $\ln C_t - \ln Y_t$ | -0.313 | 1.492 | 1.455 | 1.000 | 1.000 | 1.000 | -0.337 | -0.312 |
| $\ln I_t - \ln Y_t$ | -1.316 | 4.090 | 3.989 | 1.000 | 1.000 | 1.000 | 0.336 | 0.311 |
| $\Delta \ln a_t$ | 0.001 | 1.193 | 1.164 | 0.008 | 0.007 | 0.010 | 0.932 | 1.000 |

$q$ was set to 1.15 to deliver a detection error probability of 25%

The business cycle moments for the approximating model are presented in the upper half of table 5 for the generalized model uncertainty specification with $q = 2$ and with a detection error probability of 0.25 (this requires $\theta = 132.15$). The results here are roughly comparable to those obtained under the approximating model under Hansen and Sargent's (2007) standard model uncertainty framework and the Epstein and Zin (1989) and Weil (1990) specification. With agents pessimistic, $q > 1$, their precautionary behavior is heightened, requiring an increase in the volatility of technology growth (and with it output and the two margins, investment and labor, to smooth the effects of output on consumption) to match the empirical volatility of consumption growth.

The lower half of table 5 contains the business cycle statistics of the worst case under the generalized model uncertainty specification with $q = 2$. Relative to the approximating model, both mechanisms from above are operational, with technology growth having a lowered mean, increased volatility, and heightened autocorrelation compared with the approximating model. The moments of consumption growth, aside from the decrease in the mean here, are nearly identical to those under Hansen and Sargent's (2007) standard model uncertainty framework. Relative to the approximating

Table 5: Business Cycle Moments, Generalized Uncertainty Preferences, $q = 2$

Approximating Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta \ln Y_t$ | Cross Corr. w$\Delta \ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta \ln Y_t$ | 0.004 | 1.233 | 1.000 | 0.008 | 0.008 | 0.007 | 1.000 | 1.000 |
| $\Delta \ln C_t$ | 0.004 | 0.566 | 0.459 | 0.101 | 0.096 | 0.091 | 0.977 | 0.974 |
| $\Delta \ln I_t$ | 0.004 | 2.810 | 2.279 | -0.018 | -0.017 | -0.016 | 0.994 | 0.995 |
| $\Delta \ln N_t$ | 0.000 | 0.526 | 0.426 | -0.024 | -0.023 | -0.022 | 0.985 | 0.987 |
| $\ln N_t$ | -1.418 | 1.644 | 1.333 | 0.950 | 0.903 | 0.858 | 0.330 | 0.315 |
| $\ln C_t - \ln Y_t$ | -0.367 | 2.162 | 1.753 | 0.950 | 0.903 | 0.858 | -0.330 | -0.315 |
| $\ln I_t - \ln Y_t$ | -1.162 | 4.955 | 4.017 | 0.951 | 0.903 | 0.859 | 0.330 | 0.316 |
| $\Delta \ln a_t$ | 0.000 | 1.347 | 1.092 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 |

Worst-Case Model

| Variable | Mean | Std. Dev. % | Relative Std. Dev. | Autocorrelations 1 | 2 | 3 | Cross Corr. w$\Delta \ln Y_t$ | Cross Corr. w$\Delta \ln a_t$ |
|---|---|---|---|---|---|---|---|---|
| $\Delta \ln Y_t$ | 0.003 | 1.240 | 1.000 | 0.047 | 0.045 | 0.046 | 1.000 | 1.000 |
| $\Delta \ln C_t$ | 0.003 | 0.565 | 0.456 | 0.270 | 0.264 | 0.260 | 0.977 | 0.974 |
| $\Delta \ln I_t$ | 0.003 | 2.908 | 2.346 | -0.011 | -0.011 | -0.009 | 0.994 | 0.995 |
| $\Delta \ln N_t$ | 0.000 | 0.532 | 0.429 | -0.024 | -0.024 | -0.022 | 0.985 | 0.987 |
| $\ln N_t$ | -1.429 | 1.673 | 1.350 | 1.000 | 1.000 | 1.000 | 0.326 | 0.312 |
| $\ln C_t - \ln Y_t$ | -0.353 | 2.195 | 1.771 | 1.000 | 1.000 | 0.999 | -0.326 | -0.312 |
| $\ln I_t - \ln Y_t$ | -1.195 | 5.281 | 4.260 | 1.000 | 1.000 | 1.000 | 0.325 | 0.312 |
| $\Delta \ln a_t$ | -0.001 | 1.350 | 1.089 | 0.012 | 0.010 | 0.011 | 0.952 | 1.000 |

$\theta$ was set to 132.15 to deliver a detection error probability of 25%

model, consumption growth volatility goes down in the worst case model despite the increase in the volatility of productivity growth and production, as the pessimistic agents here overweight ($q > 1$) the probability of the worst case and robustify their decision rules more strongly.

Figure 5 plots the joint distributions of the two states, $k_t$ and $\Delta a_t$, for the specifications of Epstein and Zin (1989) and Weil (1990), Hansen and Sargent (2007), and the generalized model uncertainty introduced here. As can be seen in the figure, the mean shift in the distribution of technology growth to the right (indicating higher average growth) is ameliorated by a downward shift in detrended capital for the specification of Epstein and Zin (1989) and Weil (1990) relative to that of Hansen and Sargent (2007). This downward shift along with the increased variability of technology growth highlights that the agents are not necessarily "better off" in the Epstein and Zin (1989) and Weil (1990) specification. The generalized model uncertainty specification with $q = 2$ is associated with a large upward shift in detrended capital. This reflects the overaccumulation of capital (and with it, drop in price through the decreased marginal productivity and increase in return) driven by the agent's overweighting the worst case scenario.

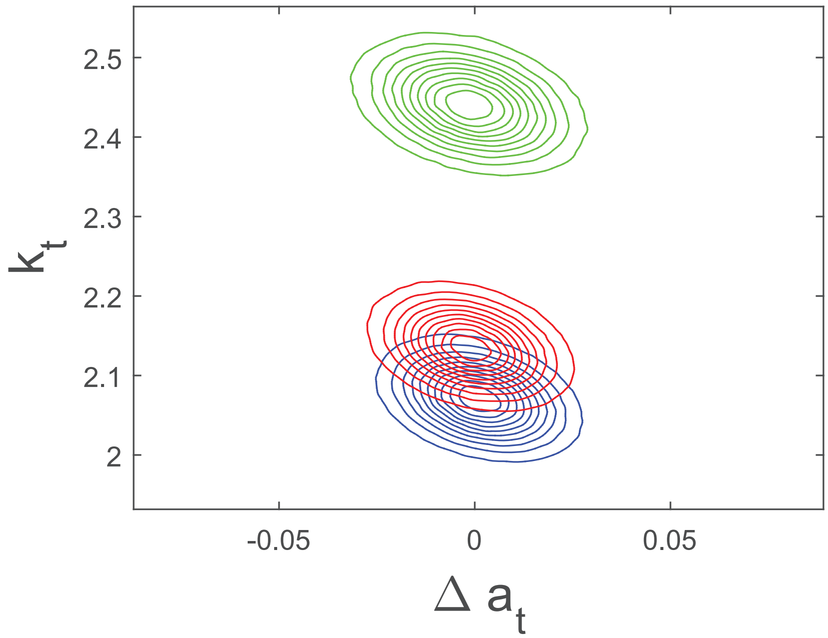



Figure 5:
Red: Hansen and Sargent (2005); Blue: Epstein and Zin (1989); Green: Generalized Uncertainty
Joint unconditional distributions of states, $a$ and $k$.

### 5.3 Asset Pricing Implications

I will first compare the specifications ability to match asset pricing facts, here using the market price of risk, for varying detection error probabilities. This will highlight the close relationship between Epstein and Zin's (1989) and Weil's (1990) risk-sensitive specification and model uncertainty following Hansen and Sargent (2007) when examining empirically plausible market prices of risk for this model. Then I will turn to the generalized model uncertainty introduced here and show that increasing the entropic index $q$ can put the model's asset pricing predictions inside the Hansen and Jagannathan (1997) bounds while maintaining a conservative detection error probability of 0.25.

Under the calibration in the previous section (specifically for detection error probabilities of 25%), both Hansen and Sargent's (2007) and Epstein and Zin's (1989) and Weil's (1990) specifications yield market prices of risk of 0.1. This relation holds more generally, as can be seen in figure 6, which plots the market price of risk of the approximating models against the detection error probabilities[16] for the Hansen and Sargent (2007) and Epstein and Zin (1989) and Weil (1990)

[16]As the particle filter with a reasonable number of particles (1,000,000) still suffers from sampling variation when calculating the likelihood tests for high and low detection error probabilities, I follow Bidder and Drew-Becker (2016)

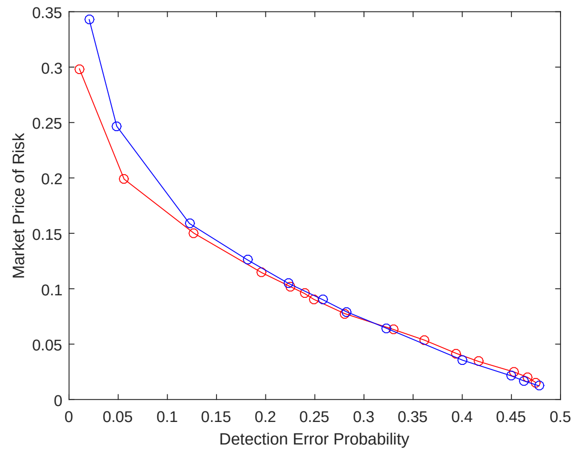



Figure 6: Red: Hansen and Sargent (2005); Blue: Epstein and Zin (1989)
Market Price of Risk and Detection Error Probabilities

specifications. For a detection error probability of 0.25, both specifications yield roughly the same market price of risk of around 0.1. For very low detection error probabilities the specification of Epstein and Zin (1989) and Weil (1990) and for very high detection error probabilities the specification of Hansen and Sargent (2007) produces higher market prices of risk. That these two different specifications yield very similar results when controlling for the detection error probabilities confirms the close relation between these two different preference specifications for the model here.

Table 6: Entropic Index and the Market Price of Risk

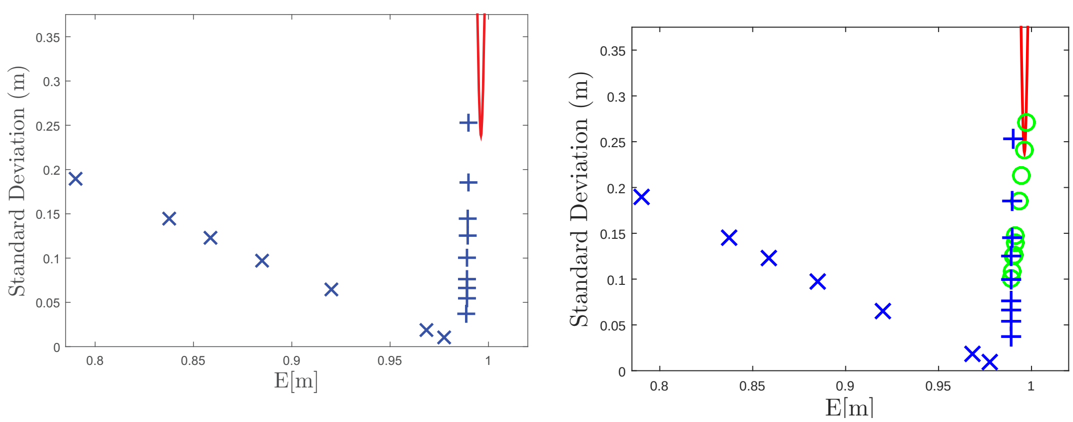
| $q =$ | 1 | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 1.75 | 2 | 2.25 | 2.5 |
|---|---|---|---|---|---|---|---|---|---|---|
| MPR | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 | 0.19 | 0.21 | 0.24 | 0.27 |

$\theta$ is adjusted to keep the detection error probability at 0.25.

Holding the detection error probability 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 6. For the $q = 2$ specification of the previous section, the market price of risk 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 Hansen and Sargent's (2007) and Epstein and Zin's (1989) and

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 detection error probabilities compared with calculations conditional on a subset of the models' variables (i.e., consumption).

Weil's (1990) specifications. That agents overweight the probability of the worst case under the generalized model uncertainty formulation drives up the returns on risky capital relative to the risk free bond. One could object to the fact the econometrician uses the actual likelihood ratio $g$ when calculating the detection error probabilities while the agents in the model overweight $g^q$ the worst case 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 Epstein and Zin (1989) and Weil (1990) specification that operates solely through $q$: the approximate equivalence with Hansen and Sargent's (2007) specification in regards to the market prices of risk and detection error probabilities in figure 6 rests likewise on this discord between the measures of the agents and the econometrician.




(a) $\times$: Expected Utility; $+$: Tallarini (2000) and Hansen and Sargent (2005)

(b) $\bigcirc$: Generalized Entropy, $q = 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.75, 2, 2.25, 2.5$

Figure 7: The Hansen-Jagannathan Bounds

Hansen and Jagannathan (1997) extend the maximal Sharpe ratio point restriction on pricing kernels to a parabola inside which pairs of $std\left(m_{t+1}\right)$ and $E\left[m_{t+1}\right]$ must reside to be consistent with (a vector) of risky assets and the riskless bond. Figure 7a contains this 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 $\theta$). For the expected utility case, the risk-free rate puzzle can be seen through the decrease in $E\left[m_{t+1}\right]$ 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, this degree of risk aversion is associated with a detection error probability of 5%, arguably past the limit of credulity.

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 Epstein and Zin (1989) and Weil (1990) 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 the worse case brings the macroeconomic model's predictions of the market price of risk in line with empirical post war U.S. observations for reasonable detection error probabilities.

# 6 Conclusion

I have derived a generalization of the model uncertainty framework of Hansen and Sargent (2007), using Tsallis's (1988) generalized entropy. The resulting preferences recover Hansen and Sargent's (2007) original formulation with an exponential certainty equivalent as one special case and recover the constant elasticity of substitution risk specification of Epstein and Zin (1989) and Weil (1990) with a power certainty equivalent as another. This latter result is particularly important, as it provides a model uncertainty foundation for Epstein and Zin (1989) and Weil (1990) preferences with arbitrary period utility functions (allowing, e.g., arbitrary intertemporal elasticities of substitution). This is desirable as a small amount of model uncertainty can substitute for a high risk aversion, as demonstrated by Barillas, Hansen, and Sargent (2009).

In an application to a standard RBC model, I find that both Hansen and Sargent's (2007) original formulation and the model uncertainty formulation for Epstein and Zin (1989) and Weil (1990) provide roughly the same predictions for the market price of risk for plausible detection error probabilities. Aside from these limiting cases, the generalization provides a two parameter model approach to model uncertainty, with the new parameter induced by Tsallis's (1988) generalized entropy, the entropic index $q$, determining a form of pessimism that induces agents to overweight the worst case model when forming expectations. As a result, increasing the entropic index (or increasing pessimism) leads to an increase in the market price of risk for a given detection error probability. The empirical value of the market price of risk can be achieved with modest detection error probabilities (25%) and a slightly elevated entropic index ($q = 2$). Future research will seek to discipline this new parameter empirically.

# References


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

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.

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. DREW-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.

BUCKLEW, J. A. (2004): *An Introduction to Rare Event Simulation*. Springer Verlag.

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

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

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., 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.

——— (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.

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,” SFB 649 Discussion Paper 2013-022 April.

——— (2013b): “Pruning in DSGE Perturbation,” SFB 649 Discussion Paper 2013-024 May.

——— (2013c): “Solving DSGE Models with a Nonlinear Moving Average,” *Journal of Economic Dynamics and Control*.

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

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

——— (2012b): "Risk aversion, risk premia, and the labor margin with generalized recursive preferences," Working Paper Series 2012-17, Federal Reserve Bank of San Francisco.

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

# A Appendix

## A.1 Proof of Proposition 2.1

Set the intertemporal elasticity of substitution in (4) 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 (15) completes the proof.

## A.2 Proof of Proposition 4.2

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 in (29) can be written as

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

where $\lambda$ is the multiplier on the constraint that the distorted distribution be a distribution.

The first order condition is

(A-2) $$0 = V(x)p(x) + \theta(x)qg(x)^{q-1}\ln_q(g(x))\,p(x) + \theta(x)g(x)^q g(x)^{-q}p(x) + \lambda p(x)$$

which can be rearranged as

(A-3) $$0 = V(x)p(x) + q\theta(x)g(x)^{q-1}\ln_q(g(x))\,p(x) + \theta(x)p(x) + \lambda p(x)$$

Substituting the form of the entropy multiplier from assumption 4.1

(A-4) $$\theta(x) \doteq \theta + (q-1)V(x)$$

gives

(A-5) $$0 = q\left[V(x)p(x) + \theta(x)g(x)^{q-1}\ln_q(g(x))\,p(x)\right] + \theta p(x) + \lambda p(x)$$

multiplying the foregoing with $g(x)$

(A-6) $$0 = q\left[V(x)p(x)g(x) + \theta(x)g(x)^q\ln_q(g(x))\,p(x)\right] + \theta g(x)p(x) + \lambda g(x)p(x)$$

rearranging

(A-7)
$$0 = q\left[V(x)p(x)g(x) + \theta(x)g(x)^q\ln_q(g(x))\,p(x) + \lambda(g(x)p(x) - 1)\right] + \theta g(x)p(x) + \lambda g(x)p(x)(1-q) + q\lambda$$

and integrating over $x$ yields

(A-8) $$0 = q\tilde{V} + \theta + \lambda$$

Combining the foregoing, (A-8) with the first order condition, (A-5)

$$0 = q\left[V(x) - \tilde{V} + \theta(x)g(x)^{q-1}ln_q\left(g(x)\right)\right]p(x) \tag{A-9}$$

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

$$0 = V(x) - \tilde{V} + \theta(x)\frac{\left(1-g(x)^{q-1}\right)}{1-q} \tag{A-10}$$

which can be rearranged as

$$0 = V(x) - \tilde{V} + \theta\frac{\left(1-g(x)^{q-1}\right)}{(1-q)} - V(x)\left(1-g(x)^{q-1}\right) \tag{A-11}$$

and

$$0 = g(x)^{q-1}V(x) - \tilde{V} + \frac{\theta}{1-q}\left(1-g(x)^{q-1}\right) \tag{A-12}$$

multiplying the foregoing with[17] $\frac{1-q}{\theta}g(x)^{1-q}$ delivers

$$0 = (1-q)\frac{1}{\theta}V(x) - (1-q)\frac{1}{\theta}g(x)^{1-q}\tilde{V} + g(x)^{1-q} - 1 \tag{A-13}$$

or

$$1-(1-q)\frac{1}{\theta}V(x) = g(x)^{1-q}\left(1-(1-q)\frac{1}{\theta}\tilde{V}\right) \tag{A-14}$$

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

$$g(x) = \frac{\left(1-(1-q)\frac{1}{\theta}V(x)\right)^{\frac{1}{1-q}}}{\left(1-(1-q)\frac{1}{\theta}\tilde{V}\right)^{\frac{1}{1-q}}} = \frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q\left(-\frac{1}{\theta}\tilde{V}\right)} \tag{A-15}$$

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

$$\tilde{p}(x) = p(x)\frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q\left(-\frac{1}{\theta}\tilde{V}\right)} \tag{A-16}$$

as was claimed in proposition 4.2.

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

$$1 = \int p(x)\frac{\exp_q\left(-\frac{1}{\theta}V(x)\right)}{\exp_q(-\frac{1}{\theta}\tilde{V})}dx \tag{A-17}$$

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

$$\exp_q(-\frac{1}{\theta}\tilde{V}) = \int p(x)\exp_q\left(-\frac{1}{\theta}V(x)\right)dx \tag{A-18}$$

yielding the risk aggregator or certainty equivalent

$$\tilde{V} = -\theta\ln_q\left[\int \exp_q\left(-\frac{1}{\theta}V(x)\right)p(x)dx\right] \tag{A-19}$$

as was claimed in proposition 4.2.

---

[17] 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-15) with $\theta$ set to zero.

## A.3 Risk Aversion

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

$$= -\theta\frac{\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{1-q} - 1}{1-q} \tag{A-20}$$

$$\frac{\partial\mathcal{R}(V)}{\partial\sigma} = -\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{-q}$$
$$\times\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{q}{1-q}} - 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{q}{1-q}}\right) \tag{A-21}$$

$$\frac{\partial^2\mathcal{R}(V)}{\partial\sigma^2} = q\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{-q-1}$$
$$\times\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{q}{1-q}} - 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{q}{1-q}}\right)^2$$
$$-\frac{q}{\theta}\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{1}{1-q}}\right)^{-q}$$
$$\times\left(0.5\left[1+(1-q)\left(-\frac{1+\sigma}{\theta}\right)\right]^{\frac{2q-1}{1-q}} + 0.5\left[1+(1-q)\left(-\frac{1-\sigma}{\theta}\right)\right]^{\frac{2q-1}{1-q}}\right) \tag{A-22}$$

# SFB 649 Discussion Paper Series 2017

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Fake Alpha" by Marcel Müller, Tobias Rosenberger and Marliese Uhrig-Homburg, January 2017.
002 "Estimating location values of agricultural land" by Georg Helbing, Zhiwei Shen, Martin Odening and Matthias Ritter, January 2017.
003 "FRM: a Financial Risk Meter based on penalizing tail events occurrence" by Lining Yu, Wolfgang Karl Härdle, Lukas Borke and Thijs Benschop, January 2017.
004 "Tail event driven networks of SIFIs" by Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle and Yarema Okhrin, January 2017.
005 "Dynamic Valuation of Weather Derivatives under Default Risk" by Wolfgang Karl Härdle and Maria Osipenko, February 2017.
006 "RiskAnalytics: an R package for real time processing of Nasdaq and Yahoo finance data and parallelized quantile lasso regression methods" by Lukas Borke, February 2017.
007 "Testing Missing at Random using Instrumental Variables" by Christoph Breunig, February 2017.
008 "GitHub API based QuantNet Mining infrastructure in R" by Lukas Borke and Wolfgang K. Härdle, February 2017.
009 "The Economics of German Unification after Twenty-five Years: Lessons for Korea" by Michael C. Burda and Mark Weder, April 2017.
010 "DATA SCIENCE & DIGITAL SOCIETY" by Cathy Yi-Hsuan Chen and Wolfgang Karl Härdle, May 2017.
011 "The impact of news on US household inflation expectations" by Shih-Kang Chao, Wolfgang Karl Härdle, Jeffrey Sheen, Stefan Trück and Ben Zhe Wang, May 2017.
012 "Industry Interdependency Dynamics in a Network Context" by Ya Qian, Wolfgang Karl Härdle and Cathy Yi-Hsuan Chen, May 2017.
013 "Adaptive weights clustering of research papers" by Larisa Adamyan, Kirill Efimov, Cathy Yi-Hsuan Chen, Wolfgang K. Härdle, July 2017.
014 "Investing with cryptocurrencies - A liquidity constrained investment approach" by Simon Trimborn, Mingyang Li and Wolfgang Karl Härdle, July 2017.
015 "(Un)expected Monetary Policy Shocks and Term Premia" by Martin Kliem and Alexander Meyer-Gohde, July 2017.
016 " Conditional moment restrictions and the role of density information in estimated structural models" by Andreas Tryphonides, July 2017.
017 "Generalized Entropy and Model Uncertainty" by Alexander Meyer-Gohde, August 2017.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**
**http://sfb649.wiwi.hu-berlin.de**

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".