

FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

**Doubts and Variability: A Robust Perspective on  
Exotic Consumption Series**

Rhys Bidder  
Federal Reserve Bank of San Francisco

Matthew E. Smith  
Hutchin Hill Capital

January 2018

Working Paper 2013-28

<http://www.frbsf.org/economic-research/publications/working-papers/2013/28/>

**Suggested citation:**

Bidder, Rhys, Matthew E. Smith. 2018. “Doubts and Variability: A Robust Perspective on Exotic Consumption Series.” Federal Reserve Bank of San Francisco Working Paper 2013-28. <https://doi.org/10.24148/wp2013-28>

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

# Doubts and Variability: A Robust Perspective on Exotic Consumption Series

Rhys Bidder\*

Federal Reserve Bank of San Francisco

Matthew E. Smith<sup>†</sup>

Hutchin Hill Capital

This draft: January 25, 2018

## Abstract

Consumption-based asset-pricing models have experienced success in recent years by augmenting the consumption process in ‘exotic’ ways. Two notable examples are the Long-Run Risk and rare disaster frameworks. Such models are difficult to characterize from consumption data alone. Accordingly, concerns have been raised regarding their specification. Acknowledging that both phenomena are naturally subject to ambiguity, we show that an ambiguity-averse agent may behave as if Long-Run Risk and disasters exist even if they do not or exaggerate them if they do. Consequently, prices may be misleading in characterizing these phenomena since they encode a pessimistic perspective of the data-generating process.

---

\*Economic Research Department, Federal Reserve Bank of San Francisco, 101 Market Street, San Francisco, CA 94115; Email: rhys.bidder@sf.frb.org

<sup>†</sup>Disclaimer: The views expressed in this paper, and all errors and omissions, should be regarded as those solely of the authors and not necessarily those of the Federal Reserve Bank of San Francisco, the Federal Reserve Board of Governors or the Federal Reserve System. Disclaimer: This paper represents the personal work product and views of Matthew Smith and have neither been reviewed nor endorsed by Hutchin Hill Capital, LP, its affiliates, any funds advised by them, or any of their officers, directors, employees, agents, partners, principals, owners, or shareholders (collectively, “Hutchin Hill Parties”). No representation or warranty is made by Matthew Smith or any Hutchin Hill Party about the completeness or accuracy of any information provided herein, and such parties expressly disclaim all liability for any of the information contained in these materials. These materials have been prepared for educational use only and may not be relied upon for any other purpose (including without limitation any commercial purpose). These materials do not constitute an offer, or the solicitation of an offer, to buy or sell any security or financial instrument or investment advisory services. No Hutchin Hill Party shall be liable or have any obligation with respect to this paper or any reliance thereon.

## 1 Introduction

A principal challenge for the early generations of consumption-based asset pricing models was to generate sufficiently high prices and quantities of risk while respecting a plausible degree of risk aversion and the apparent smoothness in consumption growth from period to period (Mehra and Prescott (1985) and Hansen and Jagannathan (1991)). One popular response to this challenge has been to introduce ‘exotic’ elements in consumption dynamics. Notably, the Long-Run Risk (LRR) model of Bansal and Yaron (2004) asserts a small but persistent component in consumption growth that allows the process to exhibit considerable risk over longer horizons without introducing counterfactual volatility at high frequencies. An alternative strategy is to appeal to rare disasters in consumption growth (Rietz (1988), Barro (2006), Gourio (2012), Gabaix (2012), Nakamura, Steinsson, Barro, and Ursa (2013) and Wachter (2013)). However, despite the success of these approaches, concerns have been raised as to their specification, or even existence. By definition, direct evidence of the LRR component is hard to detect in post-war consumption data, leading to the question of whether the component actually exists (Hansen, Heaton, and Li (2008), Marakani (2009), Beeler and Campbell (2012) and Croce, Lettau, and Ludvigson (2015)). With regard to disasters, as noted in Dolmas (2013), the rarity of the phenomenon in question undermines empirical analysis.

We take a different approach. Rather than positing the existence of Long-run Risk or taking a firm stance on the calibration of rare disasters we show that an ambiguity averse agent’s fear of model misspecification can generate or exaggerate these phenomena endogenously in the mind of the agent. Consequently, one explanation for why these exotic properties appear to be encoded in prices, but are simultaneously difficult to identify in consumption data directly, is that prices reflect not only the true model of consumption but also the agent’s fear of misspecification.

Our agent does not fully trust her ‘benchmark’ model of consumption growth and the probability distributions it implies. She acknowledges that the benchmark is an approximation to the true data generating process but fears it is misspecified in some unknown way. She expresses these fears by envisaging alternative probability distributions (implicitly capturing misspecifications in her benchmark), which she thinks may plausibly describe consumption. Formally, we endow the agent with a desire for robustness to model misspecification, as captured by the multiplier preferences of Hansen and Sargent (2008). To construct a robust evaluation of random payoffs the agent envisages adverse misspecifications, balancing the damage they could cause against their plausibility. A particular ‘worst case’ distribution emerges from the agent’s optimization problem, allowing insight into the sort of misspeci-

fications against which she desires robustness. The agent then evaluates risky payoffs *as if* this worst case is generating the data. We show that this worst case will naturally encode phenomena akin to LRR and disasters.

We begin with a specification of the benchmark that features white noise consumption growth with persistent variation in conditional variance. The worst case reflects a fear of misspecifications that would imply lower growth and higher volatility, as represented by negative and positive mean shifts in the marginal distributions for endowment and volatility innovations, respectively. Most importantly, the agent’s pessimism becomes more extreme when volatility is high, as captured by a greater negative distortion to the mean of the endowment innovation. Since the volatility process is persistent, the consumption growth process under the worst case inherits this persistence. Consequently, the worst case exhibits the hallmark of LRR models - a small but persistent component in consumption growth. The association of high volatility with low growth under the worst case also induces negative skewness in the consumption growth process so that ‘disasters’ are more common, in terms of a longer left tail of the unconditional distribution of consumption growth.

An extreme draw from the *unconditional* distribution is perhaps not the most natural concept of a disaster as it does not align with our intuition of an abrupt decline. In order to allow for more a more standard, ‘conditional’ concept of disasters we go on to consider a benchmark model featuring a non-normal ‘jump’ component in consumption growth to allow for occasional dramatic declines in consumption. In this context we show that disasters are a more powerful phenomenon to interact with robustness, in the sense that a more reasonable calibration of ambiguity aversion can attain stylized asset pricing facts, such as elevated premia and prices of risk than in the heteroscedastic Gaussian case. Under our calibrations it is entirely reasonable for an agent with a plausible degree of ambiguity aversion to behave as if disasters arrive with significantly greater frequency than implied by the model and, thus, not allowing for this fear of misspecification could undermine inference. In addition, we show that allowing for time variation in disaster risk leads to predictability in returns that reflects in part a tendency for the agent to over-extrapolate shocks to the probability of jumps.

To obtain our characterization of the worst case model we show how one can use importance sampling techniques to draw from the worst case distribution. We also make use of cumulant generating functions to obtain clear expressions for worst case moments. Both approaches are founded on an exponential affine approximation to the worst case change in measure relative to the benchmark. These methods allow Robust Control to be taken to classes of models beyond simple linear-quadratic or log-linear Gaussian setups that account for most discrete time models of robustness in the literature.

## 2 Literature

Our paper relates to several important bodies of work. Foremost, we build on the Robust Control (RC) literature and particularly Barillas, Hansen, and Sargent (2009) (BHS) who also address how allowing for a reasonable degree of model uncertainty can help with asset-pricing. However, they work with a very particular homoscedastic and log-linear endowment economy that leads to a highly restricted worst case model, characterized by a pessimistically shifted mean in endowment growth that is constant over time and analytically expressed. By introducing heteroskedasticity, we lose the simplicity of the worst case and analytic tractability.

Regarding our more elaborate worst case (featuring LRR) it is perhaps undesirable that the omission by BHS of a fairly uncontroversial phenomenon like stochastic volatility has such an important impact on the qualitative properties of the worst case, given that robustness is largely about doubting the specification of the benchmark model. Our heteroscedastic setup shows how a robust agent will distort dynamics if given enough flexibility and emphasizes that the benchmark specification can importantly restrict the usefulness of the RC approach. In response to the loss of tractability, we exploit cumulant generating functions to obtain characterizations of the worst case in discrete time and develop methods to draw from it using importance sampling. Setting aside the log-linear setup of BHS, much of the work in economics on robustness has been LQ-Gaussian, again for tractability, thereby imposing unnecessary modeling restrictions that can be relaxed using the methods in this paper.<sup>1</sup>

Alternative formulations of ambiguity have had asset-pricing success. Ju and Miao (2012) illustrate how smooth ambiguity and learning about latent states can jointly explain an array of moments. Learning also features in Epstein and Schneider (2008). We choose not to introduce learning - though surely it is an important phenomenon - as we attempt to remain within an (extremely) parsimonious RC framework in which the agent's doubts are expressed solely in terms of the model's dynamics. As an example of models using both robustness *and* learning, Hansen and Sargent (2010) explore uncertainty over the filtering distribution of a latent state in addition to uncertainty over dynamics. Allowing a second explicit type of uncertainty over the filtering distribution provides an extra degree of freedom that we do

---

<sup>1</sup>To some degree these restrictions can be overcome if one works in continuous time. Kleshchelski and Vincent (2008) focus primarily on the term structure with stochastic volatility while Xu, Wu, and Li (2010) and Anderson, Hansen, and Sargent (2003) are additional treatments of RC in continuous time. However, since discrete time models are often a workhorse in financial and particularly macro-financial applications it is useful to develop tools for these contexts. Also applicable to discrete time, Bidder and Smith (2012) use perturbation methods to approximate worst case moments (which can then be used to enhance proposal densities for Monte Carlo routines) and one may also solve models beginning with a general discretized shock distribution, as in Boyarchenko (2012).

not exploit here.

Ju and Miao (2012) also have greater freedom *via* their separation of risk aversion, ambiguity aversion and elasticity of intertemporal substitution (EIS), while we operate within a log-utility case. Their extra degrees of freedom allow empirical success but it is nevertheless important to show how the most basic RC formulation provides striking insights despite its parsimony. Furthermore, Ju and Miao (2012)’s framework does not generically yield a ‘worst case’ model. Their agent retains a pessimistic outlook but, unlike in our case, cannot be mimicked by a pessimistic agent without ambiguity aversion. It is often useful for insight into the agent’s behavior to have such an object and in some cases, such as stress testing applications (see Breuer and Csiszr (2013), Glasserman and Xu (2014) and Bidder, Giacomini, and McKenna (2016)) the worst case is practically useful. It also helps calibrate ambiguity aversion by appealing to the difficulty of distinguishing the benchmark and worst case models through statistical tests using simulated data from each model.

Long-run risk features in several papers on ambiguity aversion. The agent in Hansen and Sargent (2010) focuses on two explicit models (the latent ‘states’) as possibilities under the benchmark: an iid consumption growth model and LRR. The agent’s fears of misspecification are partly represented with distorted filtered probabilities of which model is true. It is important that in our model LRR emerges endogenously *only* in the worst case - we do not need to ‘prime the pump’ by asserting it directly as a primitive, which seems appropriate if ambiguity is to be about ‘unknown unknowns’. Bidder and Dew-Becker (2016) also generate long-run risk in the mind of an ambiguity averse agent without including it as a primitive. However, they formulate an alternative and more flexible approach to ambiguity aversion while we remain within the popular and axiomatized RC paradigm. Like Bidder and Dew-Becker (2016), we can also speak to the literature on extrapolative expectations (see Cecchetti, Lam, and Mark (1997), Fuster, Hebert, and Laibson (2011) and Greenwood and Shleifer (2014)) as the agent’s worst case features exaggerated persistence.

In analyzing the interaction of ‘jumps’ or disasters with ambiguity aversion our work relates to Drechsler (2013) and Liu, Pan, and Wang (2004). While there are several differences between their formulations and ours, most notable is the fact that neither are Hansen-Sargent-type robustness and both add substantial structure to the class of models considered plausible by the agent. Both works assert that certain aspects of the model are fully trusted - thereby focusing ambiguity in a way that the fully unstructured approach of RC does not. Since the unstructured nature of uncertainty is intuitively an important aspect of a truly ambiguous situation, it is again worthwhile to consider the more restricted RC framework despite the cost of reduced flexibility to match observed behavior and data.

### 3 Robustness

There is much evidence suggesting that, when faced with situations that are ‘ambiguous’ or uncertain in the Knightian sense, agents do not behave in accordance with standard axioms of choice (Knight (1921)). An important literature has emerged suggesting possible formalizations for how agents make decisions in such contexts. One of the more prominent and intuitively appealing formalizations is the Robust Control framework, which has been adapted for economic applications by Hansen and Sargent (2008).

#### 3.1 Environment

A robust agent entertains a benchmark model in which the state, control and innovation sequences are related according to the (possibly nonlinear) vector valued equation

$$s_{t+1} = g(s_t, u_t, w_{t+1}) \tag{1}$$

where  $s_t$  is the state vector,  $u_t$  is a vector of controls and  $w_t \equiv (w_{1,t}, w_{2,t}, \dots, w_{N,t})'$  is a vector-valued random innovation.

Given a control law,  $u_t = u(s_t)$ , and a density,  $p_w(w_{t+1}|s_t)$ , for  $w_{t+1}$ , equation (1) implies a benchmark transition density  $p(s_{t+1}|s_t)$ . It is convenient to partition the state,  $s_t$  into elements unknown on entering the period, which we identify with  $w_t$ , and those elements that are predetermined, denoted  $\hat{s}_t$ . We capture the dependence of  $\hat{s}_t$  on the state prevailing in the previous period by the function  $f$ , such that  $\hat{s}_t = f(s_{t-1})$ . With this decomposition we have  $p(s_{t+1}|s_t) = p_w(w_{t+1}|s_t) \mathbb{1}_{f(s_t)}(\hat{s}_{t+1})$ .

#### 3.2 Preferences

We endow the agent with multiplier preferences, which are discussed extensively in Hansen and Sargent (2008) and axiomatized in Strzalecki (2011). Our agent is not a Bayesian - her problem progresses from a situation of multiple models to making a decision not by integrating over the models with respect to a unique prior (essentially resolving the multiple models to a single hyper-model) but by a penalized max-min approach. Formally, the decision problem of the agent takes the following form:

$$\max_{\{u_t\}} \min_{\{m_{t+1}\}} \sum_{t=0}^{\infty} E [\beta^t M_t \{ \Upsilon(s_t, u_t) + \beta \theta E(m_{t+1} \log m_{t+1} | \mathfrak{S}_t) \} | \mathfrak{S}_0] \tag{2}$$

where  $\Upsilon(\cdot, \cdot)$  is the period payoff function,  $\beta$  is the rate of pure time preference and the problem is subject to equation (1),  $M_{t+1} = m_{t+1} M_t$ ,  $E[m_{t+1} | \mathfrak{S}_t] = 1$ ,  $m_{t+1} \geq 0$  and  $M_0 = 1$ .

We assume that the robust agent's information set,  $\mathfrak{S}_t$ , contains the entire history of states. Thus,  $\{m_{t+1}, t \geq 0\}$  is a sequence of Martingale increments that recursively define a non-negative Martingale,  $M_t = M_0 \prod_{j=1}^t m_j$ .

$M_t$  defines Radon-Nikodym derivatives that twist the measures implicit in the benchmark model to yield absolutely continuous measures representing alternative distributions considered by the agent. The agent's desire for robustness is reflected in the minimization over the sequence of martingale increments,  $m_t$ . The minimization implies that the agent considers twists in the distributions towards realizations of the state sequence that are painful. Intuitively, the agent assesses her vulnerability to model misspecification by envisaging adverse alternative models or distributions.

Clearly, without some discipline on this search for an adverse distribution, the minimization would be ill-posed or lead to pathological results. This discipline is captured by the second term in the sum, which represents the relative entropy associated with any alternative distribution considered (identified with a particular Martingale), premultiplied by a parameter  $\theta > 0$ . The degree of robustness is controlled by this term as it penalizes consideration of models that diverge from the benchmark - the greater is  $\theta$ , the greater the penalty.

The penalty captures the intuition that the agent considers models that, although different, are somehow 'near' the benchmark, which is implicitly considered reasonable. A particular alternative distribution, associated with a particular Martingale, may be especially painful in the sense of implying a very low expected payoff, but may not solve the minimization problem due to the offsetting effect of the entropy penalty. Thus, the two components in equation (2) capture the way in which the robust agent balances pain and plausibility.

We seek a recursive expression of the problem and, invoking results in Hansen and Sargent (2008), obtain a value function of the following form

$$V(w_t, \hat{s}_t) = \max_{u_t} \min_{m(w_{t+1}, \hat{s}_{t+1})} \Upsilon(s_t, u_t) \quad (3)$$

$$+ \beta \int m(w_{t+1}, \hat{s}_{t+1}) V(w_{t+1}, \hat{s}_{t+1}) p_w(w_{t+1} | s_t) \quad (4)$$

$$+ \theta m(w_{t+1}, \hat{s}_{t+1}) \log m(w_{t+1}, \hat{s}_{t+1}) p_w(w_{t+1} | s_t) dw_{t+1} \quad (5)$$

subject to  $\int m(w_{t+1}, \hat{s}_{t+1}) p_w(w_{t+1} | s_t) dw_{t+1} = 1$  for all values of  $\hat{s}_{t+1}$ .

Solving the inner minimization problem we obtain the minimizing, or 'worst case', Martingale increment, which has the form

$$m(w_{t+1}, \hat{s}_{t+1}) = \frac{e^{-\frac{V(w_{t+1}, \hat{s}_{t+1})}{\theta}}}{E_t \left[ e^{-\frac{V(w_{t+1}, \hat{s}_{t+1})}{\theta}} \right]} \quad (6)$$

Substituting this solution into the original problem we obtain the following expression

$$V(s_t) = \max_{C_t} \log C_t - \beta \theta \log E_t \left[ \exp \left( -\frac{V(s_{t+1})}{\theta} \right) \right] \quad (7)$$

where we have also asserted log period utility over consumption,  $C_t$ , that we will use in our applications below.<sup>2</sup> Note that this is algebraically equivalent to the utility recursion of an Epstein-Zin agent with log utility and risk aversion of  $1 - \alpha$  where  $\alpha = -\frac{1}{\theta}$ .

We can construct worst case analogues of the various conditional distributions defined above for the benchmark. Specifically,  $\tilde{p}(s_{t+1}|s_t) \equiv m(w_{t+1}, \hat{s}_{t+1})p(s_{t+1}|s_t)$  is the worst case conditional distribution of the state and implicit in this is the conditional worst case distribution over innovations,  $\tilde{p}_w(w_{t+1}|s_t) \equiv m(w_{t+1}, \hat{s}_{t+1})p_w(w_{t+1}|s_t)$ .

As shown in Barillas, Hansen, and Sargent (2009), our agent's stochastic discount factor is of the following form

$$\Lambda_{t,t+1} = \Lambda_{t,t+1}^r \Lambda_{t,t+1}^u \quad (8)$$

$$\Lambda_{t,t+1}^r \equiv \beta \left( \frac{C_{t+1}}{C_t} \right)^{-1} \quad (9)$$

$$\Lambda_{t,t+1}^u \equiv \frac{\exp \left( \frac{-V_{t+1}}{\theta} \right)}{E_t \left[ \exp \left( \frac{-V_{t+1}}{\theta} \right) \right]} \quad (10)$$

Thus, our agent prices assets as if she has logarithmic period utility, but under a distorted conditional expectations operator informed by the worst case, since the second component of the discount factor is the minimizing Martingale increment.

### 3.3 Characterizing the worst case

The worst case distribution is an important object beyond simply being an artifact of a decision problem since it is suggestive of the nature of other plausible and damaging models that might concern the agent. Much of the robustness literature in economics, following Hansen and Sargent (2008), adopts the Linear-Quadratic Gaussian (LQG) approach, in which the laws of motion characterizing the evolution and interaction of state and control are linear, while the period payoff to the agent is quadratic and shocks are Gaussian. This brings substantial tractability and, in particular, results in a worst case distribution that can be directly

---

<sup>2</sup>We will be considering an endowment economy such that in equilibrium prices are such that desired consumption and the realized endowment align. Although log period preferences are not strictly necessary it preserves the homogeneity properties of the value function necessary for the scaling of the value function (by consumption) that we will implement below to obtain a stationary problem. Relatedly, it also allows aggregation of such preferences to that of a representative agent with preferences of the same form.

calculated (rather than approximated). The worst case retains Gaussianity and can be represented then by a distorted mean vector and covariance matrix of the Gaussian innovations. While useful for expositional purposes - and based on an enormous control literature - many problems do not have reasonable LQG representations, rendering this approach inadequate.

In fact, robustness can be applied to much broader class of models provided one can solve for an approximation to the robust agent's value function.<sup>3</sup> Based on this, one can construct an approximation to the minimizing Martingale increment and, thus, the worst case distribution over innovations. Since the worst case distribution over sequences can be constructed from the worst case conditional distributions of innovations, combined with the law of motion, we will spend time deriving the properties of these innovation distributions. To do this we will make use of cumulant generating functions (CGF) and Monte Carlo techniques, showing how one can go well beyond the limitations of the LQG framework.

### 3.3.1 Cumulant generating functions

Innovation  $w_{i,t+1}$  has conditional CGF

$$k_{i,t}(z_i) = \log E_t [\exp \{z_i w_{i,t+1}\}] \quad (11)$$

The  $j^{\text{th}}$  cumulant is obtained by evaluating the  $j^{\text{th}}$  derivative of the CGF at 0 with the first and second cumulants being the mean and variance respectively. Moments can be derived given knowledge of the cumulants. It is convenient to set up a benchmark model such that the elements of  $w_t$  are conditionally independent. In that case, the conditional CGF for  $w_{t+1}$  is given by the sum of the CGFs of its elements.

To obtain the CGF for  $w_t$  under the worst case we apply the worst case change in measure captured in  $\Lambda_{t+1}^u$

$$\tilde{k}_t(z) = \log E_t [\exp \{z \cdot w_{t+1}\} \Lambda_{t,t+1}^u] \quad (12)$$

where  $z$  is vector valued. We cannot guarantee *a priori* that the elements of  $w_{t+1}$  are conditionally independent under the worst case. Indeed, typically they will not be as this depends upon the minimizing increment's dependence on the innovations. However, if one has an expression for the (log of)  $\Lambda_{t,t+1}^u$  that is conditionally linear in  $w_{t+1}$ , then we retain conditional

---

<sup>3</sup>Indeed, if one can solve a model with an Epstein-Zin agent with log utility, one need only reinterpret certain objects from the robustness perspective and use them to characterize the worst case.

independence, even under the worst case. Specifically, we obtain

$$\tilde{k}_t(z) = \log E_t[\exp\{(z + \lambda_t) \cdot w_{t+1}\}] + t.i.z \quad (13)$$

$$= k_t(z + \lambda_t) + t.i.z \quad (14)$$

$$\Lambda_{t,t+1}^u = \exp\{\lambda_t \cdot w_{t+1} + \xi_t\} \quad (15)$$

where  $\xi_t$  is known in  $t$  and  $\lambda_t$  is a vector, also known in  $t$ . We thus see that the worst case CGF is equal to the benchmark CGF evaluated at a shifted origin, plus terms independent of the point of evaluation such that the independence of innovations under the benchmark is inherited by the worst case. Thus, one can consider each innovation's ‘marginal’ worst case CGF in isolation when deriving its moments:

$$\tilde{k}_t(z) = \sum_{i=1}^N \tilde{k}_{i,t}(z_i) = \sum_{i=1}^N k_{i,t}(z_i + \lambda_{i,t}) t.i.z_i \quad (16)$$

As will be discussed below, under exponential affine approximations to the stochastic discount factor, we will indeed be working with linear expressions for  $\log \Lambda_{t,t+1}^u$ .

### 3.3.2 Importance sampling

Depending on the structure of the value function approximation and on the underlying benchmark innovation distributions one may or may not be able to fully characterize the worst case innovation distributions. If one can show that the worst case distribution belongs to a recognized parametric family, then one may be able to calculate explicitly enough moments to restrict all the parameters describing the distribution. For example, if one can show the worst case distributions are Gaussian then one need only calculate the mean and variance for full characterization. Then, if the distribution is of a known class, then one can draw from it directly using well known random number generators. However, even if one can calculate a large set of moments, one may not be able to fully characterize the distribution, in which case one must find a way to draw from it using alternative methods.

Provided that one can evaluate an approximation to the worst case density  $\tilde{p}_w(w_{t+1}|s_t)$ , we can draw from it using a variety of Monte Carlo techniques. This requires simply that we can obtain an approximation to the value function since we can then construct the minimizing Martingale increment and use it to pre-multiply the benchmark density. For example, we can employ Sampling Importance Resampling (SIR) algorithm of Rubin (1987) and Smith and Gelfand (1992) for drawing from the worst case distribution. This entails obtaining draws from  $p_w(w_{t+1}|s_t)$ , computing associated importance weights (given by the minimizing

Martingale increment) and then resampling with replacement according to those weights. This yields approximate draws from  $\tilde{p}_w(w_{t+1}|s_t)$ . By iteratively drawing from  $\tilde{p}_w(w_{t+1}|s_t)$  and evolving the state according to the law of motion (1) we obtain draws from the worst case distribution of sequences of the state.<sup>4</sup>

Thus, by using these methods one can apply robust control analysis to a broad class of discrete time non-linear models that admit non-standard worst case distributions, assuming one is able to obtain a reasonable approximation to the value function.

## 4 Results

We begin with a simple example, augmenting a basic endowment economy model with the presence of stochastic volatility, and show explicitly how fears of Long-Run Risk may emerge endogenously in the mind of the agent even if there is none in the benchmark. We then address models featuring disasters and show that the interaction of disaster risk with robustness is particularly powerful.

### 4.1 An illustrative example: Endowment economy with stochastic volatility

Consider a simple endowment economy in which consumption growth features stochastic volatility

$$\log g_{t+1} = G_0 + v_t^{0.5} w_{g,t+1} \quad (17)$$

$$v_{t+1} = (1 - \varphi_v) \bar{v} + \varphi_v v_t + \sigma_v w_{v,t+1} \quad (18)$$

$$w_{i,t+1} \sim N(0, 1) \text{ for } i \in \{g, v\} \quad (19)$$

where  $g_t \equiv \frac{C_t}{C_{t-1}}$  is consumption growth,  $G_0$  controls the trend,  $v_t$  is the conditional variance of shocks to consumption growth in  $t + 1$  and  $w_{g,t+1}$  and  $w_{v,t+1}$  are independent innovations.

We guess and verify an exponential affine approximation to the agent's utility function,  $u_t \equiv \frac{U_t}{C_t}$ , where we scale by current consumption to obtain a stationary problem.

$$u_t = \exp \{F_0 + F_v v_t\} \quad (20)$$

$$F_0 \equiv \frac{\beta}{1 - \beta} \left( G_0 + F_v (1 - \varphi_v) - \frac{1}{2} \frac{1}{\theta} F_v^2 \sigma_v^2 \right) \quad (21)$$

$$F_v \equiv -\frac{1}{2} \frac{1}{\theta} \frac{\beta}{1 - \beta \varphi_v} \quad (22)$$

---

<sup>4</sup>In Bidder, Giacomini, and McKenna (2016) it is emphasized that one need not even know the proposal distribution to apply similar methods, one need simply resample from draws from the proposal using a constructed change in measure.

The implied solutions for the components of the stochastic discount factor clearly make the distinction between the influence of the standard, expected utility elements of the agent's preferences, and those reflecting her preference for robustness

$$\log \Lambda_{t+1}^r = k_t + \lambda_{g,t}^r w_{g,t+1} \quad (23)$$

$$\log \Lambda_{t+1}^u = \xi_t + \lambda_{g,t}^u w_{g,t+1} + \lambda_{v,t}^u w_{v,t+1} \quad (24)$$

$$\lambda_{g,t}^r \equiv -v_t^{0.5} \quad (25)$$

$$\lambda_{g,t}^u \equiv -\frac{1}{\theta} v_t^{0.5} \quad (26)$$

$$\lambda_v^u \equiv -\frac{1}{\theta} F_v \sigma_v \quad (27)$$

We will refer to  $\lambda_{g,t}^u$  and  $\lambda_{v,t}^u$  as ‘uncertainty’ prices and  $\lambda_{g,t}^r$  as a ‘risk’ price. As the concern for robustness vanishes ( $\theta \rightarrow \infty$ ) the uncertainty prices tend to zero, while  $\lambda_{g,t}^r$  is independent of the robustness parameter and represents the standard risk price of an expected utility agent with log utility.<sup>5</sup>

#### 4.1.1 Worst case innovation distribution: Fluctuating pessimism

Under our benchmark the growth and volatility innovations are both standard Normal, such that  $k_t^g = k_t^v = k$  where  $k(z) = \frac{z^2}{2}$ . The derivatives of the worst case CGF are, for  $i \in \{g, v\}$ ,

$$\tilde{k}_{i,t}^1(z) = k^1(z + \lambda_{i,t}^u) = u + \lambda_{i,t}^u \quad (28)$$

$$\tilde{k}_{i,t}^2(z) = k^2(z + \lambda_{i,t}^u) = 1 \quad (29)$$

$$\tilde{k}_{i,t}^j(z) = k^j(z + \lambda_{i,t}^u) = 0 \text{ for } j > 2 \quad (30)$$

The first two cumulants are the mean and variance, respectively. Thus, evaluating the derivatives at the origin, we see that the only effects on the innovation distributions under the worst case are mean shifts. The properties of unit standard deviation and Normality (as indicated by higher order cumulants being zero) are preserved under the worst case. The means are pessimistically shifted by an amount equal to the uncertainty prices and, in the

---

<sup>5</sup> $k_t$  and  $\xi_t$  are functions of deep parameters and states, as shown in the appendix. Remaining endogenous variables can typically be calculated in a similar way, utilizing exponential affine approximations to the true equilibrium objects. In the case of equity returns we utilize the approximation discussed in Campbell and Shiller (1989) and, where necessary, an approximation based on the zero-coupon term structure of equity, advocated by Lettau and Wachter (2011).

case of  $w_{g,t+1}$ , the shift is state dependent, through the dependence of  $\lambda_{g,t}^u$  on  $v_t$ :

$$\tilde{E}_t[w_{g,t+1}] = \lambda_{g,t}^u = -\frac{1}{\theta}v_t^{0.5} \quad (31)$$

$$\tilde{E}[w_{v,t+1}] = \lambda_v^u = -\frac{1}{\theta}F_v\sigma_v \quad (32)$$

The mean of the endowment innovation is distorted downwards, while the mean of the volatility innovation is distorted upwards. These patterns are to be expected given the undesirability of low consumption growth and high volatility. The pessimistic shift to the mean of  $w_{g,t+1}$  is more intense, the higher is volatility. Intuitively, as the agent is more exposed to misspecifications in her model for growth (because  $w_{g,t+1}$  is pre-multiplied by a larger volatility term) she envisages greater distortions to the benchmark model as a by-product of her robust decision problem.

Although the agent doubts her model at all times, the *manifestation* of these doubts varies. Equation (31) shows that when volatility is high the pessimistic (negative) mean distortion to  $w_{g,t+1}$  is more intense. In some sense, as volatility fluctuates, the (conditional) worst case in her mind does also. This is perhaps a partial response to concerns that models of ambiguity are counterfactual in their prediction of pessimistic behavior. The argument often is that in the real world people do not typically appear pessimistic. However, we see here that robustness is quite capable of generating variations in apparent pessimism. As in the real world, pessimism can be more extreme in occasional periods of unusually high volatility.

#### 4.1.2 Worst case sequence distribution: Fears of long-run risk and ‘disasters’

The properties of the worst case innovation distributions discussed above are interesting but also, at least qualitatively, to be expected. Less obvious, however, are the implications of our model for the worst case distribution over sequences, induced by iterating on the distorted conditional distributions and the law of motion for the economy.

A consequence of the retention of Gaussianity under the worst case is that the implied distribution over sequences can be represented by a system of the same form as the benchmark model (17) – (19) but with different innovation distributions. One can re-express  $w_{i,t+1} = \varepsilon_{i,t+1} + \lambda_{i,t}^u$  with  $\varepsilon_{i,t+1} \sim N(0, 1)$  and then incorporate the  $\lambda_{i,t}^u$  term into the conditional mean dynamics of the system. Thus we can obtain the following convenient representation of the

worst case<sup>6</sup>

$$\log g_{t+1} = G_0 + \zeta_{g,v} v_t + v_t^{0.5} \varepsilon_{g,t+1} \quad (33)$$

$$v_{t+1} = (1 - \varphi_v) \tilde{v} + \varphi_v v_t + \sigma_v \varepsilon_{v,t+1} \quad (34)$$

$$\varepsilon_{i,t+1} \sim N(0, 1) \text{ for } i \in \{g, v\} \quad (35)$$

Since  $\tilde{v} > \bar{v}$ , the average level of volatility will be higher than under the benchmark. However, most notable is the additional term, linear in  $v_t$ , in equation (33). Consumption growth acquires a contribution to its conditional mean that depends on the level of volatility.

The volatility component  $v_t$  is persistent. Thus, the conditional mean of consumption growth inherits this persistence under the worst case. This is transparent in the worst case system, which makes clear that a scaled version of  $v_t$  is acting as a Long-Run Risk component.<sup>7</sup> In this important sense the agent prices assets as if there is Long-Run Risk, even though there is none in the true model. The manner in which the persistent component enters the model (*via* the concerns of the agent) is highly restricted however, since the persistence of the component is pinned down by the volatility process and its size by the degree of ambiguity aversion.

This provides a cautionary econometric tale in that it emphasizes it is important to allow for the presence of ambiguity if estimating such models using financial data. If one matches a limited set of moments and/or imposes restrictions on a subset of parameters (*via* priors and/or direct imposition) then the encoding of the worst case beliefs in asset prices could lead estimates to exaggerate the true degree of persistence under the benchmark.<sup>8</sup>

The worst case system also makes clear that consumption growth will feature unconditional skewness.  $v_{t+1}$  drives the conditional mean *and* conditional volatility so that, on average, periods of low growth will be associated with higher volatility. This raises the probability of consumption growth being thrown even further into the left tail, even while *conditional* skewness in  $t + 1$ , given information in  $t$ , remains zero. We recognize, however, that it is arguably in a conditional sense that people typically think of disasters, and we will address this in section 4.2 below, by enriching the endowment process.

---

<sup>6</sup>Similar derivations are used in Drechsler and Yaron (2011) where they interpret dynamics under a risk neutral (not worst case) measure and Backus, Chernov, and Martin (2011) also derive risk neutral object using the formalism of CGFs. Also, in continuous time, Drechsler (2013) (online appendix) distinguishes physical, risk neutral and worst case processes, expressed explicitly.

<sup>7</sup>For an approach to generating subjective long-run consumption risks outside the ambiguity framework, see Collin-Dufresne, Johannes, and Lochstoer (2013).

<sup>8</sup>See Schorfheide, Song, and Yaron (2014) for a full information approach to estimating models allowing for long run risk and a flexible specification of Epstein-Zin preferences.

### 4.1.3 Discussion: Pricing of risk and uncertainty

We note from expressions (20) – (22) that the agent exhibits no concern for time variation in  $v_t$  *per se* in the absence of robustness ( $\theta \rightarrow \infty$ ). With robust preferences, the influence of the stochastic nature of  $v_t$  is manifested in  $F_0$  and  $F_v$ . Even if  $v_t$  featured no persistence ( $\varphi_v = 0$ ) the value function would still load on  $v_t$ . This is because the recursive representation of the worst case entails a conditional mean shift in  $w_{g,t+1}$ . Therefore,  $v_t$  is a relevant state because it pre-multiplies  $w_{g,t+1}$  and thus transmits the pessimistic mean distortion to (worst case) expected consumption growth. The relevance of  $v_t$  is clearly greater if it features persistence ( $\varphi_v > 0$ ).<sup>9</sup>

Intuitively, the presence of  $v_t$  as a welfare-relevant state reflects how, from the perspective of the time-0 sequence problem, certain path realizations expose the agent to more painful misspecifications and, therefore, feature more twisting by the minimizing Martingale. Thus,  $v_t$  enables a recursive representation: once the agent has reached a ‘node’ (or history) featuring high volatility it signifies that the worst case joint distribution over sequences requires substantial distortions to the conditional distributions stemming from that node.

Now, considering the stochastic discount factor, we see that, even within a fully trusted model,  $v_t$  controls compensation for exposure to  $w_{g,t+1}$  as reflected in  $\lambda_{g,t}^r$ . However, there are additional effects in our case because  $v_t$  enhances exposure to misspecifications represented by mean shifts to  $w_{g,t+1}$  ( $\lambda_{g,t}^u \equiv -\frac{1}{\theta}v_t^{0.5}$ ). Furthermore, recalling the presence of  $v_t$  in the value function discussed above,  $w_{v,t+1}$  has the appearance of a ‘risk’ being priced ( $\lambda_v^u \equiv -\frac{1}{\theta}F_v\sigma_v$ ) because the innovation to  $v_{t+1}$  will determine how pessimistic a change in measure the agent will be operating under from the next period forwards.

Thus, the algebraic equivalence between an Epstein-Zin agent and the robust agent is natural. The Epstein-Zin agent wants to hedge against shocks to future lifetime utility. The robust agent is concerned with exposure to realizations that make her particularly vulnerable to misspecifications that are damaging for future lifetime utility. Assuming that the data generating process is the benchmark model, a ‘wedge’ is introduced that lowers the prices of assets whose payoffs depend on  $w_{g,t}$  and  $v_t$  (positively and negatively, respectively), relative to the prices that would prevail in the absence of model uncertainty. ‘Risks’ that are not priced by an expected log utility agent will appear to be priced by the robust agent, as in the case of  $w_{v,t+1}$ , not because they are risks in the Epstein-Zin sense, but because they encode a fear of misspecification and capture histories under which those misspecifications would be most concerning.

Finally, it is worth contrasting the worst case process with the risk neutral process under

---

<sup>9</sup>In addition to the time varying influence of  $v_t$  on the agent’s welfare we also observe an influence on the steady state *via* the third term in the parentheses in the definition of  $F_0$ .

Epstein-Zin.<sup>10</sup> The deviation of the worst case dynamics from the benchmark is somewhat different from that of the Epstein-Zin risk neutral measure, since the worst case omits the part of the twist in measure arising from the log utility component of the stochastic discount factor,  $\Lambda_{t+1}^r$ . However, in the absence of additional phenomena such as habits, this contribution is small. Setting this aside, we are again faced with an issue of re-interpretation, as under Epstein-Zin the agent applies time varying discount rates driven by volatility, over and above any changes in the risk free rate. Under the risk neutral measure, this will manifest as a volatility component in expected consumption growth. Alternatively, from the robustness perspective, this component has a natural interpretation as LRR in the mind of the agent, detached from any attitudes towards risk.

#### 4.1.4 Discussion: Time varying volatility and uncertainty

It is worth contrasting the nature of the time variation in our agent’s ‘pessimism’ with that discussed in Drechsler (2013).<sup>11</sup> In our framework it is not the case that as volatility ( $v_t$ ) varies over time our agent’s *uncertainty* varies. Although we are working with a recursive formulation of the robust problem, the sequence problem (2) makes clear that the agent is uncertain about the worst case distribution over sequences implied by her benchmark and that this degree of uncertainty is fixed. Since in our recursive setup we are representing the worst case distribution over sequences in terms of worst case conditional distributions (combined with the law of motion (1)), the *manifestation* of this uncertainty in terms of distorted conditional means varies over time with  $v_t$ . Nevertheless, uncertainty is fixed.

In Drechsler (2013), ‘uncertainty’ is allowed to vary over time with volatility in the sense that the penalty for distortions to benchmark conditional distributions is linked to the level of volatility - when volatility is high, the marginal penalization for entropic deviation is reduced *in addition to* the agent’s exposure to misspecification being greater. Although it seems very plausible that times of high volatility are somehow associated with higher uncertainty, this connection must be modeled carefully if one is to claim a fully structural interpretation of the separate influence of these two factors. Indeed, the ambiguity literature is in some sense defined by the explicitly separate treatment of ‘risk’ and ‘uncertainty’ and we pursue that approach here.

Drechsler adds further structure to the worst case distribution in partitioning components of the state into elements whose dynamics are uncertain and those whose dynamics are

---

<sup>10</sup>We thank an anonymous referee for encouraging us to elaborate on this point.

<sup>11</sup>See also Kleshchelski and Vincent (2008) and Xu, Wu, and Li (2010) who analyze the effects of stochastic volatility in consumption on asset prices in a continuous time setting with a robust agent. Our model is entirely in discrete time. The tools we present can be used in many discrete time representative agent frameworks, a standard workhorse of modern macroeconomics.

fully trusted. A similar approach is also found in Liu, Pan, and Wang (2004) where model uncertainty is focused only on the jump - and not diffusion - components of their payoff process in a model aimed primarily at explaining stylized option-pricing facts. In contrast, we only have one parameter,  $\theta$  that controls the nature of the worst case in relation to the benchmark. Once one allows the agent to envisage (absolutely continuous) distributions that deviate from the benchmark, penalizing the entropic distance according to  $\theta$ , our control over the the worst case is gone -  $\theta$  has been ‘used up’.<sup>12</sup>

#### 4.1.5 Tentative quantification

We undertake a stylized quantitative analysis of this simple model, which is naturally comparable to that of Barillas, Hansen, and Sargent (2009) as we augment their model with stochastic volatility. Using nondurable plus services consumption data for the U.S. from 1948:Q2 to 2013:Q4 we estimate the endowment process parameters by Bayesian methods. We use the posterior means from the estimation for our parameterization, listed in the upper panel of table 1, which indicate substantial and persistent variation in the volatility of shocks over time. We calibrate  $\beta$  and the degree of ambiguity aversion to attain an annual risk free rate of 2.59% and two values of the unconditional quarterly market price of risk (0.25 and 0.375) to span a plausible range.<sup>13</sup>

To interpret how the agent trades off her concerns for misspecifications against their plausibility, we will follow the standard approach of the literature in connecting  $\theta$  to detection error probabilities (Hansen and Sargent (2008)). Detection error probabilities (DEPs) indicate whether, with a limited amount of data, an agent could accurately distinguish between the worst case and benchmark models using likelihood ratio tests. That is, we ask what the probability of mis-identifying the data generating process would be if one were running likelihood ratio tests on data generated under the benchmark and the worst case. The larger is  $\theta$ , the closer the benchmark will be to the worst case, due to the more rapid offsetting effect of the entropy penalty in the minimization problem described in section 3.2. Thus, the higher will be the detection error probabilities. It is this connection between  $\theta$  and the

---

<sup>12</sup>There are other important differences between these papers’ approaches and ours. In particular, under a given model/belief, the agent in Drechsler (2013) has Epstein-Zin preferences, whereas in our case our agent would have expected log utility conditional on a model. Additionally, Liu, Pan, and Wang (2004) Liu, Pan, and Wang (2004) also adopt a different (more general) regularization penalty than relative entropy.

<sup>13</sup>The targeted value of the risk free rate is equal to the median of the short sample simulations carried out by Beeler and Campbell (2012) under the Bansal and Yaron (2004) calibration of a LRR model. As in that work, we calculate the annual risk free rate by rolling over the one-period risk free rate, in this case from a quarterly frequency. In Barillas, Hansen, and Sargent (2009) 0.25 is approximately the value required to attain the Hansen-Jaganathan bounds though Sharpe ratios are commonly reported in the literature substantially higher than this, based on broadening the assets considered beyond simply holding the stock market.

detection error probabilities that allows calibration of the agent’s robustness. Further details of our DEP calculations are included in the appendix.

In the lower panel of table 1 we report DEPs associated with sample sizes of 100 and 250 quarters, meant to be illustrative of the sort of sample lengths commonly employed in macro-finance applications. We also report the implied risk aversion coefficient admitted by taking the alternative interpretation of the agent’s preferences as being Epstein-Zin with unit EIS. As one can see from the table one would require an extremely implausible degree of risk aversion to attain the desired prices of risk. In comparison, the detection error probabilities seem perhaps more reasonable, especially for a market price of risk (MPR) of 0.25. This is the essence of the story told in Barillas, Hansen, and Sargent (2009) who argue that a little robustness can substitute for a lot of risk aversion. Nevertheless, the detection error probabilities in this simple example are rather low (though in line with those of Barillas, Hansen, and Sargent (2009)) and suggestive of an extreme degree of robustness, even if it is less extreme than the implied risk aversion. Despite this we will analyze the properties of the worst case under these calibrations as the model is meant to be illustrative.<sup>14</sup>

In table 2 we show moments of the endowment process under the benchmark and worst case. In panel A we observe the dependence of the worst case distribution over  $w_{g,t+1}$  on  $v_t$ , as captured in the greater mean distortions at higher values of current volatility. In panels B and C we also show unconditional moments of consumption growth at yearly aggregation and observation. As expected, the worst case features lower growth and elevated volatility on average. Skewness in consumption growth is enhanced and the autocorrelations illustrate additional persistence under the worst case, reflecting the agent’s aversion to models with a small but persistent long run component in consumption growth - the hallmark of the canonical Long-Run Risk framework. The autocorrelations under the worst case are small but due to their presence at long lags, they result in substantially increased power at low frequencies. Pursuing this further, Dew-Becker (2013) has argued that an important measure of risk in the long run is what is typically termed the long run standard deviation (LRSD) of consumption growth:

$$\sigma \left( \Delta E_{t+1} \sum_{j=0}^{\infty} \log g_{t+1+j} \right) = \sqrt{\sum_{j=-\infty}^{\infty} \gamma_j} \quad (36)$$

where  $\sigma(\cdot)$  is to be understood as yielding the standard deviation,  $\Delta E_{t+1}$  represents the change in expectations from  $t$  to  $t + 1$  and  $\gamma_j$  is the  $j^{th}$  autocovariance of consumption

---

<sup>14</sup>Barillas, Hansen, and Sargent (2009) obtain a MPR of approximately 0.22 with a 5% detection error probability under a sample length of 235 quarters. Anderson, Hansen, and Sargent (2003) suggest that DEPs of approximately 10% might be thought plausible with a sample length of 200 quarters.

growth. Thus, the LRSD is the square root of the spectral density of consumption growth at frequency 0 and encodes *all* auto-covariances, rather than an arbitrarily selected subset.<sup>15</sup> We take the LRSD as an index of Long-Run Risk. In panel B of table 2 we scale the LRSD by the unconditional standard deviation (USD) of consumption, so as to focus on dynamics and retain comparability as the USD varies, as it does when comparing benchmark and worst case. We see clearly the concentration of power at low frequencies under the worst case.

## 4.2 A benchmark with explicit disasters

In the previous section we showed that our agent was informed by a worst case model featuring unconditional skewness in consumption growth. We interpreted this as a fear of disasters. In this section we pursue a more natural interpretation of disasters, in the sense of conditional skewness due to a ‘jump’.

### 4.2.1 Benchmark model

Much as we developed our robust perspective of Long-Run Risk around a benchmark that borrowed from Barillas, Hansen, and Sargent (2009), we here draw upon the framework laid out in Wachter (2013).<sup>16</sup> We will abstract from the presence of stochastic volatility in the ‘non-jump’ innovation and allow for both constant and time varying intensity,  $h_t$ , of a ‘jump’ component. Thus, our benchmark model is

$$\log g_{t+1} = G_0 + w_{z,t+1} + \bar{v}^{0.5} w_{g,t+1} \quad (37)$$

$$w_{g,t+1} \sim N(0, 1) \quad (38)$$

$$w_{z,t+1} | j_{t+1} \sim N(j_{t+1}\chi, j_{t+1}\delta^2) \quad (39)$$

$$j_{t+1} | h_t \sim \frac{e^{-h_t} h_t^{j_{t+1}}}{j_{t+1}!} \quad (40)$$

$$h_{t+1} | z_{t+1} \sim \Gamma(\delta_h + z_{t+1}, c_h) \quad (41)$$

$$z_{t+1} | h_t \sim \frac{e^{-\zeta_t} \zeta_t^{z_{t+1}}}{z_{t+1}!} \quad (42)$$

$$(43)$$

Following Backus, Chernov, and Zin (2014) we model disasters by using a Poisson-mixture

---

<sup>15</sup>The spectral density at angular frequency  $\omega$  is defined as  $f(\omega) \equiv \sum_{j=-\infty}^{\infty} \gamma_j \cos(\omega j)$ .

<sup>16</sup>Wachter’s model is in fact deeply connected to ours in that she asserts an EIS of unity within an Epstein-Zin framework - the special case with an algebraic equivalence with the preferences of a robust agent. Thus, we provide an interesting robustness reinterpretation of a closely related setup, somewhat reminiscent of Barillas, Hansen, and Sargent (2009)’s reinterpretation of Tallarini (2000).

of Normal distributions for the ‘jump’ component,  $w_{z,t+1}$ . The intensity process for  $w_{z,t+1}$ , denoted  $h_t$ , is an Autoregressive Gamma (ARG) process (though we will also consider a constant intensity case). Conditional on its own lagged value it follows a Poisson-mixture of Gamma distributions. If we set  $c_h = \frac{\sigma_h^2}{2}$  and  $\delta_h = (1 - \varphi_h) \frac{\bar{h}}{c_h}$  (thereby defining  $\bar{h}$  and  $\sigma_h$ ) then we have  $h_t = (1 - \varphi_h) \bar{h} + \varphi_h h_{t-1} + w_{h,t}$  where  $w_{h,t}$  is a Martingale difference sequence and  $\zeta_t \equiv \frac{\varphi_h h_{t-1}}{c_h}$ .

The conditional cumulant generating functions associated with  $w_{z,t+1}$  and  $w_{h,t+1}$  are

$$k_{t,w_z}(s) = h_t \left( e^{s\chi + \frac{(s\delta)^2}{2}} - 1 \right) \quad (44)$$

$$k_{t,w_h}(s) = s\varphi_h \left( \frac{1}{1 - sc_h} - 1 \right) h_t - \delta_h (sc_h + \log(1 - sc_h)) \quad (45)$$

Again we obtain an exponential affine approximation to  $u_t$  in terms of the state

$$u_t = \exp\{F_0 + F_h h_t\} \quad (46)$$

$$F_0 = \beta \left( G_0 + F_0 - \frac{1}{\theta} \bar{v} - \frac{\delta_h}{\alpha} \log \left( 1 + \frac{1}{\theta} F_h c_h \right) \right) \quad (47)$$

$$F_h = -\beta\theta \left( e^{-\frac{\chi}{\theta} + \frac{(\frac{\delta}{\theta})^2}{2}} - 1 + \frac{-\frac{\varphi_h}{\theta} F_h}{1 + \frac{F_h c_h}{\theta}} \right) \quad (48)$$

This implies a stochastic discount factor comprising ‘risk’ and ‘uncertainty’ components as before

$$\log \Lambda_{t+1}^r = k + \lambda_g^r w_{g,t+1} + \lambda_z^r w_{z,t+1} \quad (49)$$

$$\log \Lambda_{t+1}^u = \Xi_t + \lambda_g^u w_{g,t+1} + \lambda_z^u w_{z,t+1} + \lambda_h^u w_{h,t+1} \quad (50)$$

where  $k$  and  $\Xi_t$  are both known as of time  $t$  and the risk/uncertainty prices are given by

$$\lambda_g^r = -\bar{v}^{0.5} \quad (51)$$

$$\lambda_z^r = -1 \quad (52)$$

$$\lambda_g^u = -\frac{1}{\theta} \bar{v}^{0.5} \quad (53)$$

$$\lambda_z^u = -\frac{1}{\theta} \quad (54)$$

$$\lambda_h^u = -\frac{1}{\theta} F_h \quad (55)$$

Parameters of the consumption growth process are listed in table 3. We set  $\bar{h}$  to imply a steady state expected number of jumps in 100 years to be 2.5, informed by the calibration

used by Wachter (2013). The persistence parameter,  $\varphi_h$  is also based on Wachter’s parameterization.<sup>17</sup> The mean of  $w_{z,t+1}$  given a single jump,  $\chi$ , is set to  $-0.3$  and  $\delta$  is set to  $0.1$ . This parameterization was used in Backus, Chernov, and Zin (2014) as an approximation to the multinomial distribution for consumption declines in the case of disasters, used by Wachter (2013) and Barro and Ursua (2008). In the absence of disasters, the trend growth and volatility of the Gaussian innovation are calibrated to yield mean and standard deviation of annual consumption growth of  $1.80\%$  and  $1.99\%$  respectively.

We choose  $\delta_h$  and  $c_h$  to ensure the aforementioned steady state jump intensity while also allowing substantial volatility in equity returns.<sup>18</sup> Following Wachter, we take dividends to be levered consumption ( $D_t = C_t^\phi$ ) and allow for a  $40\%$  probability of default on government debt in the case of a disaster, in which case the proportional reduction in the promised face payoff is equal to the realized disaster size,  $w_{z,t}$ . We consider more elaborate asset pricing moments in this section in contrast to our simple heteroscedastic baseline as the earlier framework was simply too restrictive to attempt to hit a broad set of calibration targets.

#### 4.2.2 Constant disaster intensity

In the constant intensity case, where we choose  $\theta$  and  $\beta$  again to hit unconditional *MPR* targets as well as a desired expected return on government debt. We ensure an average return of  $1.36\%$  conditional on no disasters occurring and approximately  $1.05\%$  unconditionally, similarly to Wachter (2013). Note that the detection error probabilities are dramatically higher, and thus more plausible, than in our Long-Run Risk frameworks - a *MPR* of  $0.375$  ( $0.25$ ) is associated with a *DEP* of  $12\%$  ( $19\%$ ) given a sample period of  $250$  quarters. Intuitively, this captures the fact that rare disasters are a powerful phenomenon to interact with ambiguity. Under the benchmark the expected number of jumps in a century is  $2.5$  while for an *MPR* of  $0.375$  ( $0.25$ ) the number is approximately  $9$  ( $7$ ). In addition, for both parameterizations there is a small distortion to  $\chi$  (the parameter controlling the mean of the disaster shock conditional on the number of jumps,  $j_{t+1}$ ) such that  $\tilde{\chi}$  is increased in magnitude by about  $10\%$  to approximately  $-0.33$  and  $-0.34$  respectively. This would be highly undesirable for the agent given the likely drop in consumption in the case of a disaster, but the frequency is not so much higher that she can distinguish confidently between the benchmark and worst case if, say, she observes two jumps in  $25$  years (corresponding to our shorter *DEP* sample period).

---

<sup>17</sup> $\varphi_h = 1 - \kappa/12$  where  $\kappa$  is the persistence parameters in Wachter’s continuous time intensity process.

<sup>18</sup>Backus, Chernov, and Zin (2014) map Wachter’s calibration into an approximate discrete time version using an AR process for intensity but dramatically reduce the volatility and persistence of the intensity process to avoid unacceptably high probability of the process going negative. Given our ARG specification of  $h_t$ , which is necessarily positive, we can handle a more volatile intensity process featuring right skew.

The worst case distribution concentrates the damage of a misspecification in a way that is efficient, in the sense of trading the pain of the distortion relative to the benchmark against the offsetting entropy penalty. Rare events are ideal for this since they are sufficiently infrequent to be difficult to characterize, as well as being very painful. Since the presence of jumps allows us to generate a high MPR with far higher  $\theta$  than in the purely Gaussian cases considered above, there is a concomitant reduction in the distortion to the Gaussian shock (the mean shift to  $w_{g,t+1}$  is  $-\theta^{-1}\bar{v}^{0.5}$ ). In our LRR setups above the DEPs were rather low because the worst case represented the agent’s fear of misspecification by large mean shifts to shocks observed period after period, thereby substantially reducing average growth over any short span of time. This rendered the benchmark and worst case easily distinguishable even in small samples. In the case of jumps, there is a much smaller reduction in typical period-to-period growth - the worst case instead emphasizes dramatic but occasional declines.

Since part of the point of this paper is to acknowledge the natural uncertainty around phenomena such as disasters, it is worth exploring how some of our results might change as we vary our parameterization. We choose to do this in a particular way: we fix our MPR and return on government debt targets and then vary  $\bar{h}$  to imply a range of expected jumps in a century, allowing our preference parameters,  $\theta$  and  $\beta$  to adjust as necessary. In figure 1 we show the implied CRRA under the aforementioned Epstein-Zin reinterpretation, the implied DEP and the associated distorted value of  $\chi$  (recalling that the benchmark value is  $-0.30$ ). We also show the worst case jump frequency. Though there is some sampling variability some interesting patterns emerge.<sup>19</sup> While the required ‘risk aversion’ naturally increases as the benchmark probability of disaster increases - especially at the lower range before flattening out somewhat - the implied DEP decreases. That is, the plausibility of the risk aversion required to hit desired asset-pricing targets deteriorates as  $\bar{h}$  declines (and begins to approach levels that are typically though implausible - see Mehra (2003)) but the plausibility of the ambiguity aversion parameterization increases.<sup>20</sup> A simple implication of this is that the ‘advantage’ of robustness over risk aversion is greater at lower  $\bar{h}$  so that reducing the probability of a jump from our primary calibration to lower levels than in Wachter (2013) that some (such as Backus, Chernov, and Zin (2014)) prefer, would actually be in our favor.

It is important to note that  $\theta$  does decrease as  $\bar{h}$  declines (recall CRRA is  $1 - \alpha$  and  $\theta = -\frac{1}{\alpha}$  is the mapping between the Epstein-Zin and robustness interpretations). However, unlike  $\alpha$ ,  $\theta$  is not naturally thought of as a ‘deep parameter’ whose interpretation is invariant the rest of the model. It is context specific: one robustness model with a given value for  $\theta$

<sup>19</sup>There is some randomness due to the simulation methods used to calibrate  $\theta$  and  $\beta$  - even with 2 million periods - and DEP calculations also are stochastic.

<sup>20</sup>We do not claim this is a general result - or even that it might apply at much higher - but less relevant - levels of  $\bar{h}$  than we have explored.

may imply very different DEP from another model with the same value of  $\theta$ . In contrast, under the risk aversion interpretation one can appeal to Arrow-Pratt experiments and the like so that the same value between models has the same interpretation. What is typically used to judge the plausibility of preferences of the agent from an ambiguity perspective is the DEP, as advocated in Hansen and Sargent (2008), and this index of plausibility is used to compare across models. In a sense, the DEP is the deep parameter and  $\theta$  is determined by it and the model.

In our case it appears that there is an especially powerful interaction between ambiguity and disasters. It is not simply the case that the DEPs for many of the values considered in the figure are more reasonable than the risk aversion values (though most of the latter are also defensible) and much more reasonable than the DEPs in our simple heteroscedastic baseline, but they also move inversely with  $\bar{h}$ . On reflection, however, it is not necessarily the case that ‘ambiguity’ (as opposed to risk) should decrease as  $\bar{h}$  declines, due to the intuitive connection of uncertainty with rarity. In addition, due to the lesser frequency of the jumps the agent can plausibly worry about more extreme distributions, conditional on a jump. Indeed, in the panel showing the more extreme distortions to  $\chi$  as the jump becomes less frequent, we see evidence of this effect.

### 4.2.3 Time varying disaster intensity

Finally, we consider the case of time varying intensity of disasters. We have chosen  $\theta$  and  $\beta$  to hit an unconditional (allowing for disaster realizations) *MPR* target of 0.25 and the same targeted value of the mean return on government debt (conditional on no disasters) of 1.36% as in the constant intensity case. Detection error probabilities for 100 and 250 quarters are 14.7% and 7.3%, respectively. When evaluated at the worst case unconditional mean of  $h_t$  the ‘steady state’ expected number of jumps in 100 years is 14.7 (again in comparison with 2.5 under the benchmark).

We report asset pricing moments, unconditionally and conditioning on the non-occurrence of disasters, which Wachter (2013) emphasizes for comparability with the post-war experience in which there have been no episodes of the sort of disaster described by Barro (2006). As in Wachter (2013), our equity returns are volatile. We can provide insight into the reasons for the volatility of returns by appealing to the worst case in the mind of the agent.

Under the worst case,  $w_{h,t+1}$  does not appear to belong to a known distribution, despite the tractability we obtain with our exponential affine approximations. Nevertheless, as shown

in the appendix, we can express the conditional mean and variance of  $w_{h,t+1}$  as

$$\tilde{E}_t[w_{h,t+1}] = \tilde{k}_{t,w_h}^{(1)}(0) \quad (56)$$

$$= a_0 + a_h h_t \quad (57)$$

$$\tilde{\sigma}_t^2(w_{h,t+1}) = \tilde{k}_{t,w_h}^{(2)}(0) \quad (58)$$

$$= f_0 + f_h h_t \quad (59)$$

which implies that we can think of  $h_{t+1}$  as following an adjusted process

$$h_{t+1} = (1 - \tilde{\varphi}_h) \tilde{h} + \tilde{\varphi}_h h_t + \tilde{w}_{h,t+1} \quad (60)$$

$$\tilde{\varphi}_h \equiv \varphi_h + a_h \quad (61)$$

$$= \frac{\varphi_h}{(1 - \lambda_h^u c_h)^2} \quad (62)$$

$$\tilde{h} \equiv \frac{\delta_h c_h + a_0}{1 - \tilde{\varphi}_h} \quad (63)$$

We have that  $\tilde{h} > \bar{h}$  and  $1 > \tilde{\varphi}_h > \varphi_h$  so that, as before, the jump arrival rate is pessimistically distorted upwards but also, now that it is allowed to vary, is more persistent than under the benchmark. Indeed, under our calibration, the half life of a shock to  $h_t$  is approximately 21 years under the worst case, in comparison with 8.5 years under the benchmark.<sup>21</sup>

Again, we show in the appendix that  $w_{z,t+1}$  is distributed under the worst case as a Poisson-mixture of normals with inflated arrival rate  $\hat{h}_t \equiv h_t e^{\lambda_z^u \chi + \frac{(\lambda_z^u \delta)^2}{2}}$  and inflated mean on arrival,  $\tilde{\chi} > \chi$ . Consequently, when innovations to the intensity process strike, they are exaggerated and over-extrapolated (relative to the true model) when projected into the future under the worst case. Therefore, the agent acts as if facing a more volatile news-flow for future consumption and, thus, dividend growth.

This extrapolative behavior leads to excess sensitivity and underpins volatility in returns and, apparently, much of the predictability evidence as well, shown in table 4. In the mind of the agent, shocks to expected cashflows are exaggerated at long horizons. Thus, following a negative (positive) shock to  $h_t$  she will be prepared to pay more (less) for the claim and, in equilibrium, the price dividend ratio should rise (fall). Nevertheless, since we assert the true model is the benchmark, an econometrician would attribute a measure of the variation in worst case *cashflow* expectations to variation in expected returns as the source of the volatility and predictability of returns as there is not as much predictability in cashflows as implied

---

<sup>21</sup>By half life we mean the  $j$  required for the effect on the expectation in  $t + j$  to be half of the effect on the expectation at the first horizon:  $\frac{\log 0.5}{\log \varphi_h}$ .

under the worst case. Thus, in both our stochastic volatility baseline and under time varying disaster probabilities, we can to some extent speak to the literature on over-extrapolation by agents (see Cecchetti, Lam, and Mark (1997) and Fuster, Hebert, and Laibson (2011), for example). In our LRR case the extra persistence in consumption growth under the worst case comes from the connection of mean growth to its volatility, which in turn is persistent. So, agents overextrapolate from a higher-moment shock rather than a cashflow shock, as it were. In the time varying disaster risk case it is shocks to the risk of disaster that are over-extrapolated.

## **5 Conclusion**

Attributing a desire for robustness to Knightian uncertainty leads a robust agent to act as if guided by a worst case distribution. We suggest novel methods of characterizing and drawing from this distribution and show that it will feature a small persistent component when the agent faces persistent heteroscedasticity and exaggerated jumps (in terms of frequency and size) if the agent perceives the true data generating process to exhibit some degree of jump risk. This allows the model to match important asset pricing facts without taking a firm stance on whether LRR exists or whether extant disaster calibrations are correct.

## References

- 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.
- BACKUS, D., M. CHERNOV, AND I. MARTIN (2011): “Disasters Implied by Equity Index Options,” *The Journal of Finance*, 66(6), 1969–2012.
- BACKUS, D., M. CHERNOV, AND S. E. ZIN (2014): “Sources of Entropy in Representative Agent Models,” *The Journal of Finance*, 69(1), 51–99.
- BANSAL, R., AND A. YARON (2004): “Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles,” *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.
- BARRO, R. (2006): “Rare Disasters and Asset Markets in the Twentieth Century,” *The Quarterly Journal of Economics*, 121(3), 823–866.
- BARRO, R. J., AND J. F. URSUA (2008): “Consumption Disasters in the Twentieth Century,” *American Economic Review*, 98(2), 58–63.
- BEELER, J., AND J. Y. CAMPBELL (2012): “The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment,” *Critical Finance Review*, 1(1), 141–182.
- BIDDER, R., AND I. DEW-BECKER (2016): “Long-Run Risk Is the Worst-Case Scenario,” *American Economic Review*, 106(9), 2494–2527.
- BIDDER, R., AND M. SMITH (2012): “Robust animal spirits,” *Journal of Monetary Economics*, 59(8), 738–750.
- BIDDER, R. M., R. GIACOMINI, AND A. MCKENNA (2016): “Stress testing with misspecified models,” Federal Reserve Bank of San Francisco working paper.
- BOYARCHENKO, N. (2012): “Ambiguity shifts and the 2007-2008 financial crisis,” *Journal of Monetary Economics*, 59(5), 493 – 507, Carnegie-NYU-Rochester Conference Series on Public Policy - Robust Macroeconomic Policy at Carnegie Mellon University on November 11-12, 2011.
- BREUER, T., AND I. CSISZR (2013): “Systematic stress tests with entropic plausibility constraints,” *Journal of Banking and Finance*, 37(5), 1552 – 1559.
- CAMPBELL, J. Y., AND R. J. SHILLER (1989): “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies*, 1(3), 195–228.

- CECCHETTI, S. G., P. LAM, AND N. C. MARK (1997): “Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good To Be True?,” *American Economic Review*, 90, 787–805.
- COLLIN-DUFRESNE, P., M. JOHANNES, AND L. A. LOCHSTOER (2013): “Parameter Learning in General Equilibrium: The Asset Pricing Implications,” Working paper.
- CROCE, M. M., M. LETTAU, AND S. C. LUDVIGSON (2015): “Investor Information, Long-Run Risk, and the Term Structure of Equity,” *Review of Financial Studies*, 28(3), 706–742.
- DEW-BECKER, I. (2013): “Estimates of the volatility of the permanent component of consumption and their implications for asset pricing,” Duke University, Fuqua School of Business.
- DOLMAS, J. (2013): “Disastrous disappointments: asset-pricing with disaster risk and disappointment aversion,” Discussion paper.
- DRECHSLER, I. (2013): “Uncertainty, Time-Varying Fear, and Asset Prices,” *The Journal of Finance*, 68(5), 1843–1889.
- DRECHSLER, I., AND A. YARON (2011): “What’s Vol Got to Do with It,” *Review of Financial Studies*, 24(1), 1–45.
- EPSTEIN, L. G., AND M. SCHNEIDER (2008): “Ambiguity, Information Quality, and Asset Pricing,” *Journal of Finance*, 63(1), 197–228.
- FUSTER, A., B. HEBERT, AND D. LAIBSON (2011): “Natural Expectations, Macroeconomic Dynamics, and Asset Pricing,” Working Paper 17301, National Bureau of Economic Research.
- GABAIX, X. (2012): “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” *The Quarterly Journal of Economics*, 127(2), 645–700.
- GLASSERMAN, P., AND X. XU (2014): “Robust risk measurement and model risk,” *Quantitative Finance*, 14(1), 29–58.
- GOURIO, F. (2012): “Disaster Risk and Business Cycles,” *American Economic Review*, 102(6), 2734–2766.
- GREENWOOD, R., AND A. SHLEIFER (2014): “Expectations of Returns and Expected Returns,” *Review of Financial Studies*, 27(3), 714–746.
- HANSEN, L. P., J. C. HEATON, AND N. LI (2008): “Consumption Strikes Back? Measuring Long-Run Risk,” *Journal of Political Economy*, 116(2), 260–302.
- HANSEN, L. P., AND R. JAGANNATHAN (1991): “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy*, 99(2), pp. 225–262.

- HANSEN, L. P., AND T. J. SARGENT (2008): *Robustness*. Princeton University Press, Princeton, NJ.
- HANSEN, L. P., AND T. J. SARGENT (2010): “Fragile beliefs and the price of uncertainty,” *Quantitative Economics*, 1(1), 129–162.
- JU, N., AND J. MIAO (2012): “Ambiguity, Learning, and Asset Returns,” *Econometrica*, 80(2), 559–591.
- KLESHCHELSKI, I., AND N. VINCENT (2008): “Robust Equilibrium Yield Curves,” Working Paper, HEC Montreal.
- KNIGHT, F. H. (1921): *Risk, uncertainty and profit*. Houghton Mifflin Company, Boston, MA.
- LETTAU, M., AND J. A. WACHTER (2011): “The term structures of equity and interest rates,” *Journal of Financial Economics*, 101(1), 90 – 113.
- LIU, J., J. PAN, AND T. WANG (2004): “An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks,” *Review of Financial Studies*, 18(1), 131–164.
- MARAKANI, S. (2009): “Long run consumption risks: Are they there?,” .
- MEHRA, R. (2003): “The equity premium: Why is it a puzzle,” *Financial Analysts Journal*, 59(1).
- MEHRA, R., AND E. C. PRESCOTT (1985): “The equity premium: A puzzle,” *Journal of Monetary Economics*, 15(2), 145–161.
- NAKAMURA, E., J. STEINSSON, R. BARRO, AND J. URSA (2013): “Crises and Recoveries in an Empirical Model of Consumption Disasters,” *American Economic Journal: Macroeconomics*, 5(3), 35–74.
- RIETZ, T. A. (1988): “The equity risk premium a solution,” *Journal of Monetary Economics*, 22(1), 117–131.
- RUBIN, D. B. (1987): “A noniterative sampling importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm,” *Journal of the American Statistical Association*, 82(398).
- SCHORFHEIDE, F., D. SONG, AND A. YARON (2014): “Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach,” Working Paper 20303, National Bureau of Economic Research.
- SMITH, A. F. M., AND A. E. GELFAND (1992): “Bayesian Statistics Without Tears: A Sampling-Resampling Perspective,” *The American Statistician*, 46(2).

- STRZALECKI, T. (2011): “Axiomatic Foundations of Multiplier Preferences,” *Econometrica*, 79(1), 47–73.
- TALLARINI, T. D. (2000): “Risk-sensitive real business cycles,” *Journal of Monetary Economics*, 45(3), 507–532.
- WACHTER, J. A. (2013): “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?,” *The Journal of Finance*, 68(3), 987–1035.
- XU, W., C. WU, AND H. LI (2010): “Robust general equilibrium under stochastic volatility model,” *Finance Research Letters*, 7(4), 224 – 231.

Table 1: **Simple heteroscedastic baseline - parameterization**

Panel A: Endowment

Parameter	Description	Value	Parameter	Description	Value
$\bar{G}_0$	Trend growth	$4.66 \times 10^{-3}$	$\varphi_\nu$	Persistence of volatility	0.91
$\bar{\nu}$	Steady state volatility	$2.69 \times 10^{-5}$	$\sigma_\nu$	Variability of volatility	$5.45 \times 10^{-6}$

Panel B: Preferences

	Description	Low MPR	High MPR
MPR	Unconditional market price of risk	0.250	0.375
$\beta$	Time preference	0.9970	0.9965
$DEP (T = 100)$	Detection error probability (shorter sample)	11.0	3.1
$DEP (T = 250)$	Detection error probability (longer sample)	2.7	0.2
$CRRA (1 + \frac{1}{\theta})$	Risk aversion re-interpretation of $\theta$	45.5	65.0

*Notes:* In panel A the parameterization of the endowment process in equations (17) - (19) is taken to be the posterior mean from a Bayesian estimation using nondurable plus services consumption data for the U.S. from 1948:Q2 to 2013:Q4. In panel B the preference calibrations are set to attain a ‘Low MPR’ (0.25) and a ‘High MPR’ (0.375) together with an annual risk free rate of 2.59% in both cases, where preferences are characterized by recursion (7).

Table 2: Simple heteroscedastic baseline - properties of worst case

Panel A: Worst case innovation means

Calibration	$E[w_{g,t+1} v_t]$			$E[w_{v,t+1} v_t]$
	Low $v_t$	Median $v_t$	High $v_t$	$\forall v_t$
Low MPR	-0.18	-0.23	-0.27	0.06
High MPR	-0.25	-0.33	-0.40	0.12

Panel B: Unconditional moments of (annualized)  $\log(g_t)$

Model	Mean	SD	Skew	LRSD/SD
Benchmark	1.86	1.04	0.00	1.00
Worst Case (Low MPR)	1.32	1.13	-0.20	1.09
Worst Case (High MPR)	0.98	1.22	-0.24	1.16

Panel C: Autocorrelations of (annualized)  $\log(g_t)$

Model	(lag) 1	2	3	4	5
Benchmark	0.00	0.00	0.00	0.00	0.00
Worst Case (Low MPR)	0.03	0.02	0.01	0.01	0.01
Worst Case (High MPR)	0.05	0.04	0.03	0.02	0.01

*Notes:* The calibrations considered are to attain a ‘Low MPR’ (0.25) and a ‘High MPR’ (0.375) together with an annual risk free rate of 2.59% in both cases. In panel A we consider three values for  $v_t$  where low (high) is the 20<sup>th</sup> (80<sup>th</sup>) percentile under the benchmark unconditional distribution. Note that this is only relevant for  $w_{g,t+1}$ . In panel B we show moments of consumption growth under the benchmark model (invariant to preference parameterization) along with the analogous moments under the worst case distributions (dependent on the preference parameterization). Skew is the standardized third moment. LRSD/SD is the ratio of the long run and unconditional standard deviations, with the former as defined in the main text. In panel C we display the annual frequency autocorrelations of consumption growth aggregated to annual frequency. Again, the benchmark autocorrelations, but not the worst case correlations, are invariant to the preference parameterization.

Table 3: **Rare disasters - parameterization**

Panel A: Endowment and payoffs

Parameter	Description	Value	Parameter	Description	Value
$G_0$	Trend growth	$1.50 \times 10^{-3}$	$\varphi_h$	Jump intensity persistence (ARG only)	0.993
$\bar{\nu}$	Volatility (non-disaster shock)	$3.30 \times 10^{-5}$	$\delta_h$	Auxiliary parameter (ARG only)	1.788
$\chi$	Scaling of disaster shock mean	-0.3	$c_h$	Auxiliary parameter (ARG only)	$7.766 \times 10^{-6}$
$\delta$	Scaling of disaster shock SD	0.1	$q$	Government bond partial default	0.4
$\bar{h}$	Steady state jump intensity	$2.08 \times 10^{-3}$	$\phi$	Dividend leverage parameter	3.96

Panel B: Preferences

	Description	$h_t = \bar{h}$	$h_t = \bar{h}$	ARG $h_t$
		Low MPR	High MPR	Low MPR
MPR	Unconditional market price of risk	0.250	0.375	0.250
$\beta$	Time preference	0.9991	0.9987	0.9992
DEP ( $T = 100$ )	Detection error probability (shorter sample)	29.3	22.7	14.71
DEP ( $T = 250$ )	Detection error probability (longer sample)	19.3	11.6	7.28
CRRRA ( $1 + \frac{1}{\theta}$ )	Risk aversion re-interpretation of $\theta$	4.10	4.94	3.78

*Notes:* Panel A: We show the parameterization of the process defined by equations (37) - (42).  $\bar{h}$  implies a steady state expected number of jumps in 100 years to be 2.5. The persistence parameter,  $\varphi_h$  is based on Wachter’s (2013) parameterization ( $\varphi_h = 1 - \kappa/12$  where  $\kappa$  is the persistence parameter in her continuous time intensity process). The parameterization of  $\chi$  and  $\delta$  was used in Backus, Chernov, and Zin (2014) as an approximation to the multinomial distribution for consumption declines in the case of disasters, used by Wachter (2013) and Barro and Ursua (2008). In the absence of disasters, the trend growth and volatility of the Gaussian innovation are calibrated to yield mean and standard deviation of annual consumption growth of 1.80% and 1.99% respectively. Panel B: Preference calibrations are set to attain a ‘Low MPR’ (0.25) and a ‘High MPR’ (0.375) together with an annual risk free rate of 1.36% conditional on no disasters occurring. In the ARG intensity specification we only consider the ‘low MPR’ case. Preferences are characterized by the recursion (7).

Table 4: **Rare Disaster - ARG Intensity**

Panel A: Calibration

	Description	Unconditional	No disasters	Wachter (2013)
				No Disasters
$E[r^b]$	Average return on government debt	1.05	1.36	1.36
$SD(r^b)$	Standard deviation of return on government debt	3.38	0.92	2.00
$E[r^b - r^e]$	Average excess return	7.74	9.23	8.85
$SD(r^e)$	Standard deviation of equity return	21.05	17.66	17.66
$S.R.$	Sharpe ratio	0.38	0.52	0.49
$SD(\log g_t)$	Volatility of consumption growth	5.44	1.99	1.99
$SD(\log g_{d,t})$	Volatility of dividend growth	21.56	7.87	5.16

Panel B: Predictability

	Horizon (years)					
	1	2	4	6	8	10
Unconditional						
$\beta_1$	-0.07	-0.13	-0.25	-0.34	-0.44	-0.48
$R^2$	0.01	0.02	0.04	0.05	0.06	0.06
No Disasters						
$\beta_1$	-0.12	-0.24	-0.44	-0.62	-0.76	-0.86
$R^2$	0.07	0.14	0.24	0.32	0.37	0.40
Wachter (2013) - No Disasters						
$\beta_1$	-0.16	-0.30	-0.56	-0.77	-0.95	-1.10
$R^2$	0.13	0.24	0.41	0.52	0.59	0.63

*Notes:* Return series are obtained from aggregating monthly returns to annual. Excess returns are obtained by comparing the annualized gross returns on equity and government debt. We used 100,000 periods of simulations and dropped observations for jump periods when calculating the ‘conditional’ regressions. For the predictability regressions we estimate

$$\sum_{j=0}^k \log R_{t+j}^e - \log R_{t+j}^b = \beta_0 + \beta_1(p_t - d_t) + \varepsilon_t$$

where  $p_t - d_t$  is the log price dividend ratio, where the dividend is levered consumption,  $D_t = C_t^\phi$  (we do not calculate it as a 12-month moving average).

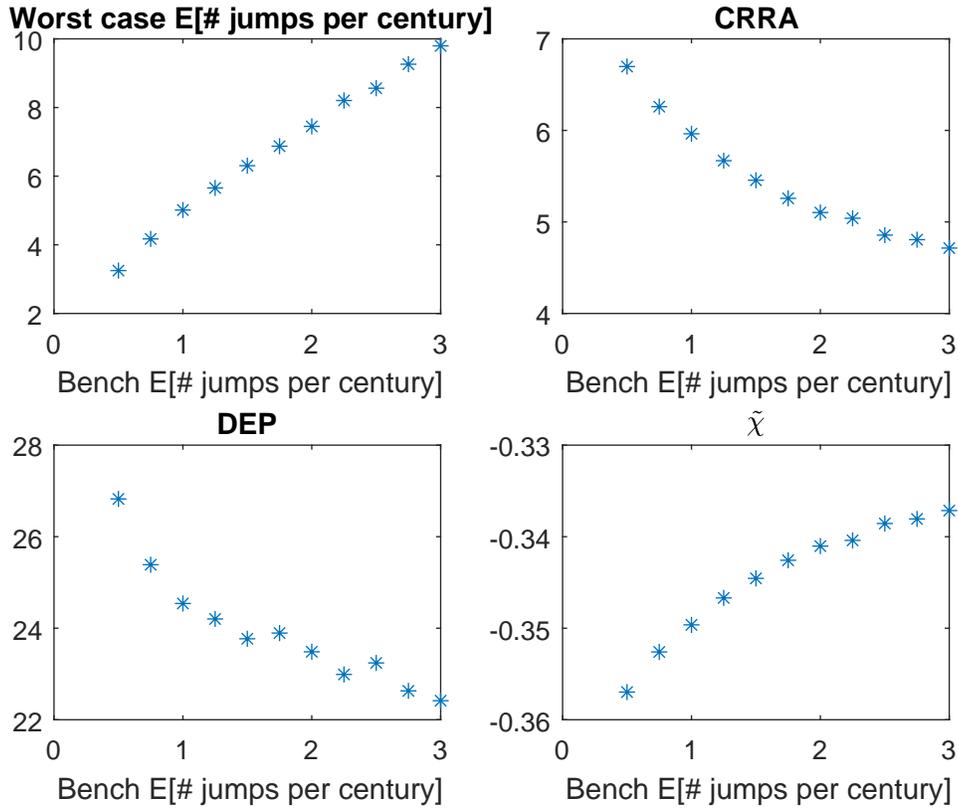


Figure 1: **Varying  $\bar{h}$ :** We vary the steady state jump intensity,  $\bar{h}$  (x-axes transform this into expected number of jumps in a century under the benchmark model) while choosing  $\theta$  and  $\beta$  to ensure an unconditional MPR of 0.375 and an average return on government debt (in the absence of disasters) of 1.36%. The remaining parameters are the same as in the constant intensity case in the main text, shown in table 3. The DEPs were calculated using 25 years of simulated data in the LR tests.