Learning, Rare Disasters, and Asset Prices*

Yang K. Lu[†] Michael Siemer[‡]

January 2016

Abstract

We incorporate joint learning about state and parameter into a consumption-based asset pricing model with rare disasters. Agents are uncertain whether a negative shock signals the onset of a disaster or how much long-term damage a disaster will cause and they update their beliefs over time. The interaction of state and parameter uncertainty increases the total amount of uncertainty and slows learning. Once the two types of uncertainty are both priced in asset prices, their joint effect enables our model to account for the level and volatility of U.S. equity returns without relying on exogenous variation in disaster risk or any realization of disaster shock in the data sample.

Keywords: rare disasters, Bayesian learning, equity premium puzzle, time-varying risk premia, return predictability

JEL: E21, G12, D83

*First draft: September 2011. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Board or the Federal Reserve System. This paper was previously circulated under the title "Asset Pricing with Learning about Disaster Risk." The authors are indebted to Francois Gourio and Robert G. King for their continuous support. The authors would like to thank Stelios Bekiros, Martin Ellison, George Evans, Albert Marcet, John M. Roberts, and Vladimir Yankov for useful comments and suggestions, as well as participants at numerous conferences and seminars. This research is supported by the Research Grants Council of Hong Kong, General Research Fund #643811, 2011-2013.

†Hong Kong University of Science and Technology. Address: Department of Economics, Clear Water Bay, Kowloon, Hong Kong; phone: +852 2358-7619; Email: yanglu@ust.hk; personal website: http://ihome.ust.hk/~yanglu/.

‡Board of Governors of the Federal Reserve System. Address: Division of Research and Statistics, Macroeconomic and Quantitative Studies, 20th & Constitution Ave NW, Washington DC 20551; phone: +1 (202) 912-7860; Email: michael.siemer@frb.gov; personal website: http://www.michael-siemer.com.

1 Introduction

The importance of economic uncertainty for the macroeconomy and asset pricing is well documented in the macrofinance literature.[1] Researchers have found many ways to measure uncertainty, yet little is known about what generates it. In this paper, we argue that when rare disasters are present in an economy, learning about the disaster risk is a fruitful source for time-varying economic uncertainty and accounts for a large fraction of volatility of equity prices if the uncertainty is priced properly.

Following the influential work by Rietz (1988) and Barro (2006), we define a rare disaster as an infrequent, large shock with a long-lasting negative effect on aggregate consumption. Examples of rare disasters include the Great Depression and World War II. Although rare disasters seldom occur, people often fear that the economy might be heading toward a disaster, which would create a great amount of uncertainty ahead.[2] A cross-country study by Nakamura et al. (2013) shows that the estimated probability that an economy is experiencing a disaster varies greatly over time. The same study also finds large standard deviations of the short- and long-run effects of disasters. A recent paper by Chen et al. (2015) argues that it is particularly difficult to estimate the probability of disasters and the average disaster size from data.

In light of these findings, we assume that the occurrence of a disaster is a hidden state and that a disaster unfolds over multiple periods during which consumption declines persistently.[3] We also assume that the parameter governing the long-term damage of rare disasters is unknown to agents. Yet, real-time consumption data are observable and are used by agents to update their beliefs regarding two aspects of rare disasters: (1) whether a decline in consumption is caused by

[1]See Bloom (2009); Gourio (2012); Drechsler and Yaron (2011), Fernández-Villaverde et al. (2011); Bloom et al. (2012); Schaal (2012); and Gourio et al. (2013), among others.

[2]For instance, during the financial crisis of 2007 to 2009, many commentators — including Nobel Prize-winning macroeconomists — warned that the U.S. economy could fall into another Great Depression. In response stock prices plummeted and volatility went through the roof. The nuclear accident in Japan in early 2011 provides another example. Within two days of the news breaking out, the Japanese stock market tumbled 22% as the country and, indeed, the world wondered whether the nuclear meltdown could be contained.

[3]This assumption is motivated by the empirical evidence in Nakamura et al. (2013). Many papers on rare disasters assume that the entire damage of a disaster occurs in a single period (Rietz (1988); Barro (2006); Farhi and Gabaix (2008); Guo (2009); Gabaix (2012); Wachter (2013)), but such an assumption has been criticized by many others as unrealistic (see, for example, Constantinides (2008)and Julliard and Ghosh (2012)).

a rare disaster or by a transitory shock, and (2) the long-term damage of this potential disaster to aggregate consumption.[4]

The interaction between state and parameter uncertainty gives rise to greater time variations in agents' perceived disaster risk than would exist with the two types of uncertainty separately. As the occurrence of a disaster is not directly observable and the effects of a disaster take time to unfold, the decline in consumption due to a transitory shock could be mistaken as a sign for the onset of a rare disaster. It turns out to be more difficult for agents to distinguish a transitory shock from a persistent disaster shock if the damage of a disaster is uncertain. Thus in our model, agents' beliefs about the occurrence of a disaster are more responsive to transitory shocks than in a model without parameter uncertainty. Moreover, agents facing parameter uncertainty use consumption data to update their beliefs about the long-term damage of a disaster, which leads to additional time variations in the perceived disaster risk in future consumption. As a result, joint learning about the hidden disaster state and the long-term damage of a disaster makes these learning effects more prominent in driving time variations in agents' beliefs.

To have asset prices fully reflect the time-varying perceived risk generated by joint learning, we develop a pricing approach that takes into account both state and parameter uncertainty evolving over time due to belief updating. In other words, asset prices are computed with agents' beliefs about both the state and the disaster parameter (a distribution) as state variables. We show that when agents have Epstein-Zin preferences (Epstein and Zin (1989)), our model is able to account for a large fraction of volatility of equity returns observed in data even in the absence of disaster realizations.

To highlight this feature of our model, we compute risk-free rates and equity returns using consumption data simulated from the consumption process estimated by Nakamura et al. (2013) but with all disaster shocks set to zero.[5] Our model yields an equity premium that matches the

[4] Evidence for time-varying beliefs on macroeconomic quantities is readily available, for example, in the Survey of Professional Forecasters (SPF). Even among professional forecasters, there is a significant time-varying dispersion in (consumption growth) forecasts.

[5] In this paper, we focus on understanding the learning framework and the pricing problem. There is ample research showing that learning matters for asset price movements (e.g., Johannes et al. (forthcoming)). We do not aspire to convince the readers that we have "the" learning model that explains U.S. equity prices but to show how the interaction of state and parameter uncertainty helps a consumption-based asset pricing model in matching the

observed one using U.S. return data from 1948 to 2008, a period during which the U.S. economy was *disaster-free*. More importantly, the volatility of model-implied equity returns is 17.13%, accounting for 92.5% of the observed volatility. Beyond matching the level and volatility of equity returns, our model also matches the volatility of risk-free rates in the data, whereas many other models on asset pricing tend to understate the variation in these rates.[6] Furthermore, following the common practice in the asset pricing literature, we test our model's performance in the predictive regression of future excess equity returns on the dividend price ratio. The results are largely consistent with the data.

To illustrate how joint learning about state and parameter uncertainty – coupled with the pricing approach consistent with joint learning – improves the model's performance in matching asset price moments, we compare our benchmark model to alternative models with one type of uncertainty absent either in learning or in pricing. When the simulated consumption process is free of disaster shocks, if the state is known, the parameter uncertainty about future disasters merely increases the level of equity return rather than its volatility. If the state is hidden but the damage of a disaster is known, learning about the true state happens too fast so that the priced state uncertainty alone accounts for only one-fifth of the return volatility in the benchmark model. If we allow joint learning but shut down time variations in parameter uncertainty when pricing assets – letting agents take their current beliefs about parameter as lasting forever – we find that the resulting equity returns are almost twice as volatile as those computed in the benchmark model, in which the pricing approach takes into account future updates of agents' belief about the parameter.

This paper is closely related to a fast-growing literature that introduces learning into a rare disaster model with the aim of endogenizing the time-varying disaster risk. Such a risk has proven crucial in explaining a wide array of asset pricing phenomena.[7] Most papers in that literature focus on one type of uncertainty and show that learning could be very helpful in fitting the persistence and volatility of asset returns.[8] For example, Koulovatianos and Wieland (2011) study learning

observed return moments.

[6]See, for example, Bansal and Yaron (2004); Benzoni et al. (2011); Ju and Miao (2012); Beeler and Campbell (2012); and Johannes et al. (forthcoming).

[7]Examples include Wachter (2013); Gabaix (2012); Gourio (2012) and Gourio et al. (2013).

[8]In fact, most papers in a broader literature on learning and asset pricing have only one type of uncertainty. Ex-

about the frequency of rare disasters. Gillman et al. (2014) assume that agents learn about the persistence of rare disasters. Johannes et al. (forthcoming) find that learning about the transition probabilities between normal and disaster states has more asset price impact compared to learning about the means and variances of shocks. Our paper complements the literature by arguing that the joint effect of state and parameter uncertainty in both learning and pricing can explain why equity returns in the United States were so volatile even when the economy was free of any disaster during the period from 1948 to 2008.

We are not the first to consider learning problems about multiple sources of uncertainty. Two pioneer learning papers, Lewis (1989) and Timmermann (2001), feature hidden structural breaks and parameter uncertainty. Fulop et al. (2014) and Johannes et al. (forthcoming) have three types of uncertainty in agents' learning. Bianchi (2013) and Orlik and Veldkamp (2014) study learning about hidden state in the presence of parameter uncertainty. Although all these papers find important interactions of different types of uncertainty for belief updating and therefore potentially for asset prices, it is computationally prohibitive to price more than one type of uncertainty. An extra assumption has to be imposed to make the pricing problem feasible.[9] Timmermann (2001) assumes that agents know the time when the fundamental process may have changed. Bianchi (2013), Fulop et al. (2014), and Johannes et al. (forthcoming) adopt anticipated utility pricing (following Kreps (1998) and Cogley and Sargent (2008)), which ignores parameter uncertainty by using mean parameter beliefs as the true value in pricing assets.[10]

We contribute to this literature methodologically by providing a feasible pricing approach that takes into account both state uncertainty and parameter uncertainty. More specifically, we let agents perform a likelihood-ratio test each period to detect the most recent change in state. This test is designed in the spirit of the Shiryaev-Roberts (SR) procedure that has been widely used

amples of models with only parameter uncertainty include Timmermann (1996); Weitzman (2007); Collin-Dufresne et al. (forthcoming); and Jagannathan and Liu (2015). Examples of models with only state uncertainty include Veronesi (1999); Veronesi (2004); Brandt et al. (2004); Chen and Pakos (2007); and Ghosh and Constantinides (2010).

[9]Lewis (1989) and Orlik and Veldkamp (2014) do not explore the asset pricing implications of their learning models.

[10]In one section of Johannes et al. (forthcoming), they study the pricing problem with parameter uncertainty but observed states.

for change-point detection problems in statistics literature, with the advantage that it is the most powerful test (i.e., it minimizes the probability of mistaking a recession for a disaster when the disaster does not occur) conditional on the size of the test (i.e., controlling for the probability of agents ignoring the possibility of a disaster while there is a disaster). This test helps to reduce an infinite-dimensional pricing problem to a feasible one with both beliefs about the hidden state and beliefs about the parameter as state variables. We show that in the context of our model, anticipated utility pricing yields equity returns that are almost twice as volatile as those implied by our pricing approach, which incorporates interaction of state and parameter uncertainty.[11]

2 The model

This section presents our benchmark asset pricing model with joint learning about state and parameter. After introducing the consumption process, we explain how agents update their beliefs and how they price assets.

2.1 Consumption process and information

We adopt the following process for consumption growth in which disasters affect long-run consumption:

$$c_t \equiv \Delta \log C_t = \mu + I_t \theta_\tau + \eta_t. \tag{2.1}$$

In this expression, c_t denotes the consumption growth at time t, C_t is the aggregate consumption at time t, and Δ denotes a first difference. η_t is an i.i.d. shock to consumption growth, which is normally distributed with mean zero and variance σ_η^2. The mean growth rate of consumption is μ in the normal state, denoted by $I_t = 0$, and is $\mu + \theta_\tau$ in the disaster state, denoted by $I_t = 1$. The

[11] In models with only parameter uncertainty, Cogley and Sargent (2008) find that, in the context of constant relative risk aversion (CRRA) preferences with low risk aversion, anticipated utility pricing provides a good approximation to a full Bayesian solution, whereas Collin-Dufresne et al. (forthcoming) shows that under recursive preferences (Epstein and Zin (1989)), whether the parameter uncertainty is priced makes a significant difference in asset returns.

disaster indicator I_t follows a two-state Markov chain with transition matrix Q defined as

$$Q_{ij} = \Pr\left(I_t = j | I_{t-1} = i\right) \text{ where } \sum_{j=0,1} Q_{ij} = 1,\, Q_{ii} > \frac{1}{2},\, i = 0, 1.$$

Because of the persistence built into the transition matrix, a disaster generally lasts for several periods, we call the consecutive periods with $I_t = 1$ a "disaster episode." θ_τ is the shift of mean growth rate when a disaster occurs. Notice that θ_τ has the subscript τ rather than t because we assume that θ_τ stays constant throughout a particular disaster episode (denoted by τ) and it is drawn randomly at the beginning of each disaster episode from a normal distribution $F(\theta)$ with mean μ_θ and variance σ_θ^2. Thus, θ_τ is a disaster-specific parameter that measures the severity of a particular disaster τ.

Agents in the benchmark model observe current and past realizations of consumption $c^t \equiv \{c_s\}_{s=0}^{t}$ but do not observe current and past states $\{I_s\}_{s=0}^{t}$ or disaster-specific parameters $\{\theta_s\}_{s=0}^{\tau}$. The uncertainty about the unobservable state is referred to as *state uncertainty* and the uncertainty about the unobservable parameter is referred to as *parameter uncertainty.*

The specification in Equation (2.1) is meant to capture two features of a disaster suggested by Nakamura et al. (2013):(1) a disaster typically lasts for several periods; and (2) each disaster is unique in terms of its long-term damage. Note that except for the time-varying drift in the consumption process, this specification is shared by many other models in the asset-pricing literature since the regime-switching property of the consumption growth is commonly regarded as important for understanding consumption-based asset pricing (for example, Cecchetti et al. (1990); Cecchetti et al. (1993); Cecchetti et al. (2000); Kandel and Stambaugh (1991); Ju and Miao (2012)). Our modeling of the time-varying drift follows Timmermann (2001), in which he represents structural breaks in the fundamental process as a Markov switching process with an expanding set of nonrecurring states.[12] Our model is a slight variation in the sense that all the disaster episodes are nonrecurring states but there is one recurring state – the normal state.

[12]Our specification is a simplified version of the empirical model used in Nakamura et al. (2013) as we abstract from the short-term effect of a disaster for the sake of simplicity. In the conclusion we briefly discuss the implications of adding back the short-term effect in our model.

2.2 Learning about state and parameter

In this subsection, we explain how agents update their beliefs when they have to learn the state I_t and the disaster-specific parameter θ_τ at the same time.

The presence of the two types of uncertainty distinguishes our model from the existing literature on learning, where agents typically have perfect information about either the state or the parameter. If I_t can be observed and the uncertainty is only about θ_τ, our model fits into the familiar framework of parameter learning. If, instead, the disaster-specific parameter θ_τ is observable and the uncertainty is only about the realization of I_t, our model reduces to a standard hidden Markov regime-switching model.

2.2.1 Definition of the learning trigger

In a standard model of parameter learning, the unknown parameters are constant over time and agents should use all past data to form their beliefs. In our benchmark model, parameter θ_τ is an independent draw from $F(\theta)$ each time a disaster starts and remains constant only through this particular disaster episode. Therefore, agents should only use the consumption data within the disaster episode to update their beliefs about θ_τ. Without directly observing the disaster state, I_t, and thus knowing when a disaster starts, agents are assumed to perform a statistical test each period to assess how likely the economy is in a disaster state and to decide when to start updating their beliefs about the severity of the disaster, θ_τ.[13] The result of this statistical test is represented by the learning trigger as defined below.

Definition. *The learning trigger, S_t, is an indicator that can take two values:* ***off*** *or* ***on****. When the learning trigger is* ***off*** *at the end of period t, agents ignore the small probability of an ongoing disaster and assume that $I_t = 0$. When the learning trigger is* ***on*** *at the end of period t, agents*

[13]See Kasa and Cho (2011) for another example of learning along this line. Economically, one could argue in favor of the learning trigger that in a world where learning is costly for agents, agents would only engage in learning activity if there was sufficient benefit to doing so. Technically, introducing a learning trigger keeps the model tractable since the state space would otherwise grow over time to be infinite-dimensional. In our model, the disaster state, I_t, is not directly observable to agents. Thus, without a learning trigger, agents in each period t face the uncertainty that a new disaster may have started in period $t, t-1, t-2, \ldots$. Such a model without a learning trigger would lead to an infinite number of beliefs about θ_τ and render the model intractable.

entertain the possibility of being in a disaster episode and use the current consumption c_t to update their beliefs about parameter θ_τ.

2.2.2 The test that triggers learning: the uniformly most powerful test

Now we specify the statistical test that triggers learning. We use a likelihood ratio test with the null hypothesis being $I_t = 1$ against the alternative hypothesis, $I_t = 0$:

$$H_0 : I_t = 1 \text{ and } H_1 : I_t = 0. \tag{2.2}$$

If we reject the null hypothesis, the learning trigger S_t is set to off; otherwise, it is left on. Denote $S^t \equiv \{S_s\}_{s=0}^{t}$. Conditional on S^{t-1}, we construct the test statistic, $\lambda(c_t)$, using the data c_t:

$$\begin{aligned} \lambda(c_t) &= \frac{L\left(c_t|H_0, c^{t-1}, S^{t-1}\right)}{L\left(c_t|H_1, c^{t-1}, S^{t-1}\right)} \\ &= \int_\theta \exp\left(\frac{\theta_\tau}{\sigma_\eta^2}c_t - \frac{\theta_\tau^2 + 2\theta_\tau\mu}{2\sigma_\eta^2}\right) \Pr\left(\theta_\tau|c^{t-1}, S^{t-1}\right) d\theta_\tau, \end{aligned} \tag{2.3}$$

where $L\left(c_t|\cdot\right)$ is the conditional likelihood function. If $S_{t-1} = on$ then $\Pr\left(\theta_\tau|c^{t-1}, S^{t-1}\right)$ is agents' beliefs about θ_τ inherited from the previous period; if $S_{t-1} = off$ then $\Pr\left(\theta_\tau|c^{t-1}, S^{t-1}\right)$ is the unconditional density $f\left(\theta_\tau\right)$ from the normal distribution $F\left(\theta\right)$.

We reject the null hypothesis if the test statistic $\lambda\left(c_t\right)$ is below a cutoff ϕ. According to the Neyman-Pearson lemma, this test is the uniformly most powerful test given a size α that is determined by ϕ:

$$\alpha = \Pr\left(\lambda(c_t) < \phi|H_0\right). \tag{2.4}$$

Therefore, conditional on a given probability of a false alarm (i.e., agents turn off the learning trigger while there is a disaster), this statistical test maximizes the probability of turning off the learning trigger in the absence of a disaster. This test is designed in the spirit of the Shiryaev-Roberts (SR) procedure that has been widely used for change-point detection problems in statistics literature.[14] The SR procedure minimizes the delay of detecting a parameter change in the statisti-

[14]See Shiryaev (1963); Roberts (1966); Pollak (1985); and Pollak and Tartakovsky (2008).

cal behavior of a random process for a given false alarm probability. However, the SR procedure is designed for problems in which distributions both before and after the change are known to agents. In our model, the mean of consumption growth in the disaster state is unknown and agents have to decide which observations should be incorporated to update their beliefs about the mean. We thus modify the SR procedure to suit our model.

2.2.3 Discussion

The learning trigger is set up such that agents in our model will investigate the severity of a disaster only when there is a statistically significant chance that the current consumption data are generated by a disaster shock. The real-world counterpart of the abnormal data is an event that marks the start of a turmoil. Examples include Black Tuesday in October 1929 as the start of the Great Depression, the collapse of Bear Stearns and Lehman Brothers in the fall of 2008 as the start of the Great Recession, the Tunisian revolution in December 2010 as the start of the Arab Spring, and the earthquake and tsunami in March 2011 as the start of the nuclear crisis in Japan. Among these examples, a period of turmoil sometimes had a negative long-run impact on consumption (i.e., it is a disaster) but sometimes it did not. Nonetheless, a disaster usually has a clearly marked start, which we assume to be detectable by agents using a statistical test.

The basic intuition for the "learning trigger" is that agents are not always suspicious that the economy is in a disaster state unless they observe some evidence of it. In theory, a Bayesian learner should always consider the possibility of being in a disaster and learn about its long-term damage regardless how small the possibility is. However, we find it unreasonable to have agents paranoid even when the economy is booming. Moreover, although our agents may miss some disasters due to the presence of the learning trigger, the likelihood ratio test guarantees that such a mistake is only made when the disaster's long-term effect is not too large, which limits the welfare loss caused by the learning trigger.[15]

We choose the likelihood ratio test rather than the reversed ordered Cusum test commonly used

[15] It would be ideal if we could compute and compare agents' welfare with and without the learning trigger. However, computing the welfare without the learning trigger when both states and parameters are unobservable is infeasible since the state space becomes infinite-dimensional (see Johannes et al. (forthcoming)).

in detecting a structural break (or a change point) because consumption is observed in a rather low frequency (annually here) compared to the financial data (monthly). The low frequency of the consumption data limits the use of the cusum test as it requires a minimal number of observations to obtain a reliable estimate before the test can be performed. By then, the disaster may have already ended and the test result would not be relevant for forecasting future consumption growth.

2.2.4 Learning with the learning trigger

After each realization of c_t, agents' beliefs about state I_t are updated using Bayes' rule:

$$\Pr\left(I_t = 1|S^{t-1}, c^t\right) = \frac{\Pr\left(c_t|S^{t-1}, I_t = 1, c^{t-1}\right)\Pr\left(I_t = 1|S^{t-1}, c^{t-1}\right)}{\Pr\left(c_t|S^{t-1}, c^{t-1}\right)}. \tag{2.5}$$

The likelihood function of c_t is obtained by integrating out the unknown parameter θ_τ using its conditional distribution, $\Pr\left(\theta_\tau|S^{t-1}, c^{t-1}\right)$, specified in Section 2.2.2. The prior belief about the state I_t is:

$$\Pr\left(I_t = 1|S_{t-1} = off, S^{t-2}, c^{t-1}\right) = Q_{01} \tag{2.6}$$

by the definition of the learning trigger and

$$\begin{aligned} & \Pr\left(I_t = 1|S_{t-1} = on, S^{t-2}, c^{t-1}\right) \\ = \; & Q_{01}\Pr\left(I_{t-1} = 0|S_{t-1} = on, S^{t-2}, c^{t-1}\right) + Q_{11}\Pr\left(I_{t-1} = 1|S_{t-1} = on, S^{t-2}, c^{t-1}\right). \end{aligned} \tag{2.7}$$

The realization of c_t is also used to obtain the value of S_t conditional on S^{t-1} according to the statistical test described in Section 2.2.2. Depending on the value of S_t, agents decide whether to update their beliefs about parameter θ_τ. If S_t is on, the beliefs about parameter θ_τ are updated using Bayes' rule:

$$\Pr\left(\theta_\tau|S_t = on, S^{t-1}, c^t\right) = \frac{\Pr\left(c_t|\theta_\tau, I_t = 1\right)\Pr\left(\theta_\tau|S^{t-1}, c^{t-1}\right)}{\Pr\left(c_t|S^{t-1}, I_t = 1, c^{t-1}\right)}. \tag{2.8}$$

If S_t is off, the beliefs about parameter θ_τ are neither updated nor recorded since they are no longer relevant for future disasters.

Also depending on the value of S_t, the updated beliefs about state I_t are recorded in the following way:

$$\Pr\left(I_t = 1|S_t = on, S^{t-1}, c^t\right) = \Pr\left(I_t = 1|S^{t-1}, c^t\right); \tag{2.9}$$

$$\Pr\left(I_t = 1|S_t = off, S^{t-1}, c^t\right) = 0. \tag{2.10}$$

2.3 Asset pricing

We study a representative-agent endowment economy with two assets: a risk-free bond and an equity that pays aggregate dividends D_t each period. We use R^f_{t+1} and R^e_{t+1} to denote their gross returns from period t to period $t+1$, respectively. Agents in our model are assumed to have Epstein-Zin preferences, which are defined recursively as

$$U_t = \left\{(1-\beta)\, C_t^{1-1/\psi} + \beta\left[E_t\left(U_{t+1}^{1-\gamma}\right)\right]^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}, \tag{2.11}$$

where C_t is consumption at period t, U_t is the utility at period t, β is the time discount factor, ψ is the inter-temporal elasticity of substitution (IES), and γ is the coefficient of relative risk aversion. These preferences imply the stochastic discount factor M_{t+1}:

$$M_{t+1} = \beta^{\epsilon}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{PC_{t+1}+1}{PC_t}\right)^{\epsilon-1}, \tag{2.12}$$

where $\epsilon = (1-\gamma)(1-1/\psi)^{-1}$ and PC_t is the price-dividend ratio at period t for an asset that delivers aggregate consumption as its dividend each period. Using this stochastic discount factor, PC_t can be obtained using the following recursion:

$$PC_t = \beta\left\{E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma+1}(PC_{t+1}+1)^{\epsilon}\right]\right\}^{1/\epsilon}. \tag{2.13}$$

We now turn to specifying the dividend process of the equity. It is common in the literature (for example, Ju and Miao (2012); Johannes et al. (forthcoming)) to model dividends and consumption separately since the aggregate dividend is much more volatile than aggregate consumption in the data. However, to maintain the basic feature of an endowment economy, the mean of the long-run dividend growth is usually adjusted to be equal to that of the long-run levered consumption growth (Bansal and Yaron 2004; Ju and Miao (2012)). We thus follow the literature to model the dividend process as

$$\frac{D_{t+1}}{D_t} = \left(\frac{C_{t+1}}{C_t}\right)^{\lambda} exp(g_d + \sigma_d \varepsilon_t). \tag{2.14}$$

where λ is the leverage ratio, g_d helps to match the long-run dividend and consumption growth, $\varepsilon_t \sim N(0,1)$ is the dividend shock, and σ_d is used to match the volatility of dividends in the data.[16] With this dividend process and the stochastic discount factor M_{t+1}, the price-dividend ratio for the equity, denoted by PD_t, can be obtained using

$$PD_t = E_t\left[\beta^{\epsilon}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma+\lambda}\left(\frac{PC_{t+1}+1}{PC_t}\right)^{\epsilon-1}(PD_{t+1}+1)\right]\exp\left(g_d + \frac{1}{2}\sigma_d^2\right). \tag{2.15}$$

If agents can observe θ_t and I_t perfectly, both price-dividend ratios PC_t and PD_t are functions of (θ_t, I_t). In our model, however, neither θ_t nor I_t is directly observable, so we replace them with the corresponding agents' beliefs, $\Pr(\theta_\tau|S^t, c^t)$ and $\pi_t \equiv \Pr(I_t = 1|S^t, c^t)$. Denote $PD\left[\Pr(\theta_\tau|S^t, c^t), \pi_t\right]$ as the price-dividend ratio function (PDR function hereafter) for the equity under imperfect information. It satisfies the following recursion:

$$\begin{aligned} & PD\left[\Pr\left(\theta_\tau|S^t, c^t\right), \pi_t\right]\exp\left(-g_d - \frac{1}{2}\sigma_d^2\right) \\ = \quad & \beta^{\epsilon}E_t\left[\exp\left[(\lambda-\gamma)\,c_{t+1}\right]\left(\frac{PC_{t+1}+1}{PC_t}\right)^{\epsilon-1}\left[PD\left[\Pr\left(\theta_\tau|S^{t+1}, c^{t+1}\right), \pi_{t+1}\right]+1\right]\right], \end{aligned} \tag{2.16}$$

where $PC_t = PC\left[\Pr(\theta_\tau|S^t, c^t), \pi_t\right]$ and $PC_{t+1} = PC\left[\Pr(\theta_\tau|S^{t+1}, c^{t+1}), \pi_{t+1}\right]$ are the PDR functions for the unlevered consumption claim under imperfect information that are obtained using the

[16] Abel (1999) shows that using the power parameter λ is a convenient way to model leverage. $\lambda = 1$ corresponds to an unlevered equity, and values of λ larger than 1 correspond to levered equities.

recursion (2.13).[17]

In addition to replacing (θ_t, I_t) with the agents' beliefs, imperfect information about θ_t and I_t also changes the expectation operator E_t. In particular, conditional on the information set at period t, the distribution of future $c_{t+1} = \mu + I_{t+1}\theta_\tau + \eta_{t+1}$ is determined by $\Pr(\theta_\tau | S^t, c^t)$ and π_t due to the persistence of θ_τ and I_t. Each future realization of c_{t+1} implies a set of updated state variables, $\Pr(\theta_\tau | S^{t+1}, c^t, c_{t+1})$ and $\Pr(I_{t+1} = 1 | S^{t+1}, c^t, c_{t+1})$, which are in turn associated with a particular future PDR. The expectation is taken by averaging across all possible future c_{t+1} and the associated PDRs. A similar methodology applies to the computation of the risk-free rate.

2.4 Calibration

Table 1 reports the parameter values we use in our numerical exercises. One period is one year in

Table 1: Calibration parameters

Parameter	Symbol	Value
Mean of Consumption Growth	μ	0.022
Std. Dev. of Consumption Growth Shock	σ_η	0.018
Mean of Disaster Shock	μ_θ	-0.024
Std. Dev. of Disaster Shock	σ_θ	0.049
Prob. to Enter a Disaster	Q_{01}	0.028
Prob. to Exit a Disaster	Q_{10}	0.165
Leverage Ratio	λ	2
Mean Adjustment of Dividend Growth	g_d	-0.0278
Std. Dev. of Dividend Growth Shock	σ_d	0.1219
Discount Factor	β	0.974
Risk Aversion	γ	8
IES	Ψ	2
Size of Test	α	0.1

our calibration. The parameters governing consumption are set equal to their posterior means estimated by Nakamura et al. (2013) whenever possible, so that the priors of our agents are disciplined by the empirical findings about the cross-country consumption data over the last 100

[17] In the computation, we replace $\Pr(\theta_\tau | S^t, c^t)$ by its mean and variance since it is a normal distribution.

years.[18]

However, our specification of the consumption process differs from the empirical model used by Nakamura et al. (2013) in two ways. First, we omit the transitory disaster shocks. Second, we assume that the permanent disaster shock is drawn at the onset of a disaster and remains constant throughout (i.e. it is perfectly correlated over time within each disaster) while it is i.i.d. in Nakamura et al. (2013). To match the overall variance of consumption growth through a typical disaster with only permanent disaster shocks in Nakamura et al. (2013), we reduce the standard deviation of the disaster shock in our calibration to $\sigma_\theta = 0.049$.[19] Notice that omitting transitory disaster shocks in Nakamura et al. (2013) does not invalidate the suitability of their estimates to our model since the effect of transitory shocks disappears over the long run. In other words, our agents are endowed with the same knowledge about long-run consumption growth as the econometricians.

The leverage ratio, λ, is set to 2, which is a conservative level compared with other models in the literature.[20] As discussed in the previous subsection, g_d in the dividend process (2.14) is chosen so that the long-run dividend growth is equal to the long-run consumption growth.[21] Using the parameter values of the consumption process, we thus set g_d to be -0.0278. The standard deviation of the dividend shock, σ_d, is used to match the volatility of dividends in the data, which is 0.129 for our annual sample from 1948 to 2008. We thus set $\sigma_d = 0.121$. The preference parameters are standard with a rather low risk aversion coefficient of 8 and the IES equal to 2.[22] We set the time discount factor, β, equal to 0.974 to match the average risk-free rate in the data.

The only free parameter left in our learning model is the size of the test that determines the state of the learning trigger. In the literature on statistical tests, it is common to set the test size

[18] The parameters that govern the distribution of disaster shocks imply that disasters, on average, decrease consumption in the long run, but it is possible that disasters can have positive long-run effects. According to Nakamura et al. (2013), crises can, for example, lead to structural change that benefits the country in the long run.

[19] Our model can also match the asset pricing moments similarly well when setting $\sigma_\theta = 0.121$, as in Nakamura et al. (2013). The only change required is a small reduction in the level of risk aversion γ.

[20] There is no consensus in the literature about the level of leverage. Typically the parameter value ranges from 1.5 to 4 (Nakamura et al. (2013), Gourio (2012), Bansal and Yaron (2004)).

[21] Note that with our specification of the consumption process, long-run consumption growth takes into account the impact of disasters.

[22] Bansal and Yaron (2004) rely on a risk aversion coefficient of 10, while the model of Mehra and Prescott (1985) requires a risk aversion coefficient of about 40 to match the equity premium. The Campbell and Cochrane (1999) model implies a time-varying local risk aversion coefficient larger than 30 in simulations.

to be 10%, 5%, or 1%. In our application, a larger size implies a higher probability for agents to mistakenly turn off the learning trigger, which would cause them to learn less frequently. We use a conservative 10% in our calibration so that the results can be viewed as the lower bound of how much learning can affect asset returns. Reducing the test size can only strengthen the learning effect. The appendix provides a robustness check of our results to changes in the size of the test.

3 Results

One novel contribution of our model is the presence of both parameter uncertainty and state uncertainty in learning and asset pricing. To gain a better understanding of how the two sources of uncertainty affect learning and asset prices, we compute and compare asset returns under four model variants.

The first model is our *benchmark* model in which agents learn about the hidden state I_t and the unknown parameter θ_τ. When pricing assets, agents take into account the future updates of their beliefs so that both state and parameter uncertainty are priced. In other words, the PDR function is computed using the recursion shown in Equation (2.16).

In the second model, the *parameter uncertainty* model, we turn off state uncertainty by assuming that agents have perfect information about the current and past states of the economy $(I_s)_{s=0}^{t}$ but view parameter θ_τ as an i.i.d. draw from distribution $F(\theta)$ each period. This model therefore does not feature learning but has the maximal amount of parameter uncertainty.[23] It is designed to quantify the effect of parameter uncertainty on asset returns.

The third model is labeled the *state uncertainty* model since we turn off parameter uncertainty and assume that agents learn about disaster state I_t conditional on perfect knowledge about θ_τ. This model demonstrates the effect of state uncertainty on asset returns and is one variant of the well-known hidden Markov regime-switching model.

Finally, to illustrate the importance of having both state and parameter uncertainty priced in the benchmark model, we contrast it with a fourth model, the *anticipated-utility-pricing* model.

[23] Learning about the disaster parameter reduces the amount of parameter uncertainty over time.

The label of the model comes from a common pricing approach used in the literature on asset pricing with parameter uncertainty.[24] The anticipated-utility-pricing model shares all elements with the benchmark model except that agents' beliefs about θ_τ are assumed to be fixed when they price assets, i.e., we replace $\Pr(\theta_\tau|S^{t+1}, c^{t+1})$ on the right-hand side of the recursion (2.16) by $\Pr(\theta_\tau|S^t, c^t)$.

We group the results into two sets. The first set shows the dynamics of agents' beliefs and asset returns implied by these four models throughout a typical disaster. We use this set of results to illustrate how negative shocks on consumption growth drive time variations in beliefs and asset returns. The intuition gained here paves the way to understand the quantitative results shown in the second set, in which we report the return moments and the predictive regressions computed using consumption data simulated with only transitory shocks to the economy.

3.1 A Disaster Realization

In this subsection, we investigate the dynamics of beliefs and asset returns with a sample disaster episode starting at period 5 and lasting for 6 periods (years). The long-term damage on consumption from this particular disaster is set to be 4% each period during the disaster, which implies a 24% total drop of consumption in the long run. All other shocks, η_t, are set to zero.

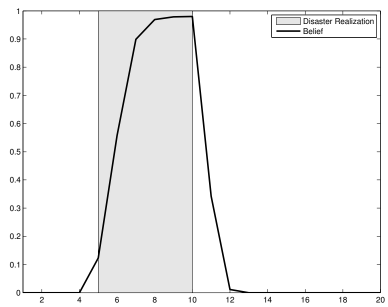
Figure 1 plots the evolution of agents' posterior beliefs that the disaster state $I_t = 1$ in the benchmark model (Panel a),[25] in the parameter uncertainty model (Panel b), and in the state uncertainty model (Panel c). Since the state is observable in the parameter uncertainty model, the beliefs are equal to zero when $I_t = 0$ and one when $I_t = 1$ in Panel b. Learning about the disaster state occurs in both Panel a and Panel c, but the speed of learning differs because of the presence of parameter uncertainty in the benchmark model. Panel d shows the difference in beliefs by subtracting the beliefs in Panel a from the beliefs in Panel c. The difference is first positive and

[24] Anticipated-utility has been widely used in the literature, including Kreps (1998), Cogley and Sargent (2008), Johannes et al. (forthcoming), and Piazzesi and Schneider (2010). It is important to clarify that these authors hold the mean belief constant and do not consider the entire distribution. However, this modification does not qualitatively change our results.

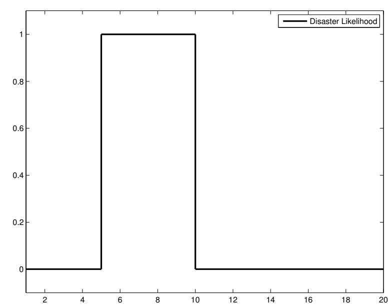
[25] Agents' beliefs in the anticipated-utility-pricing model are identical to those in the benchmark model because these two models only differ in the computation of returns.

then negative, indicating that learning about the true state is significantly slowed by the presence of parameter uncertainty. Also notice that at the early stage of the disaster, when parameter uncertainty is most pronounced, the difference in belief is the largest. As more observations of consumption growth accumulate over time, parameter uncertainty diminishes and, in turn, the difference in beliefs becomes smaller.

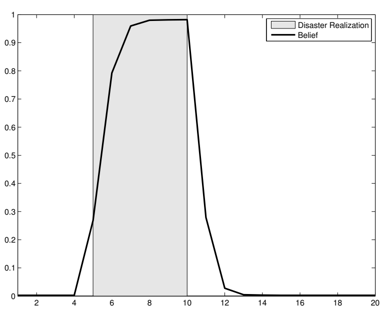
Figure 1: A Disaster Realization, Part 1



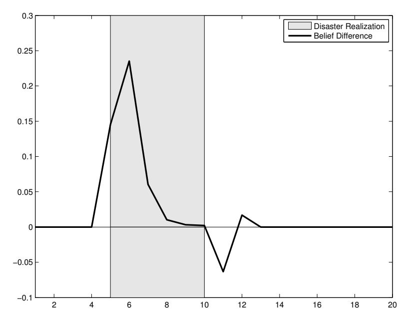
(a) Benchmark



(b) Parameter Uncertainty



(c) State Uncertainty

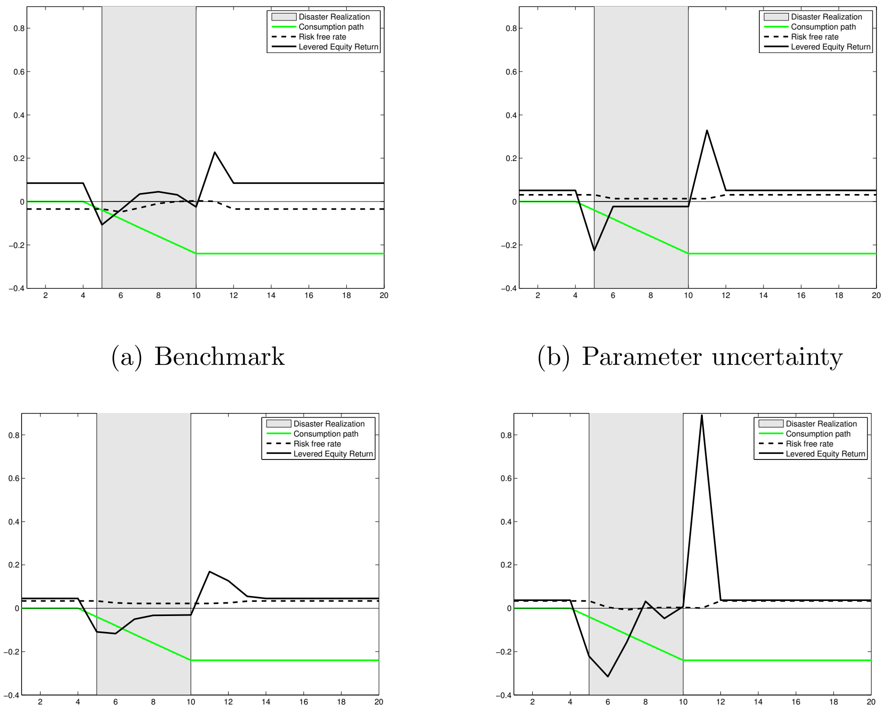


(d) Difference between panel (a) and (c)

Figure 2 shows asset returns in the four models. The black solid line is the equity returns and the black dashed line is the risk-free rates. The green line is the de-trended $\log C_t$.

Let us first consider the simplest case: the parameter uncertainty model (Figure 2, Panel b). Agents observe that the true state changes at period 5 and again at period 10. Hence, at the onset of a disaster, the stock market crashes as the expected future consumption growth decreases. The risk-free rate does not change upon impact, but it subsequently drops due to higher demand for the safer asset. After the disaster, the prospect of future consumption growth improves, which leads

Figure 2: A Disaster Realization, Part 2



(a) Benchmark

(b) Parameter uncertainty

(c) State Uncertainty

(d) Anticipated-utility-pricing

to a stock market boom. The dynamics resemble those in Figure 7 of Nakamura et al. (2013).

Next we turn to the state uncertainty model shown in Figure 2, Panel c. Compared with the parameter uncertainty model, both the crash and the boom become more gradual: movements in equity returns are smaller but more persistent. This change is driven by agents' learning about state I_t. As shown in Panel c of Figure 1, at the onset of a disaster, agents' posterior beliefs that the economy is in a disaster are as low as 25%. The presence of state uncertainty prevents the equity return from dropping as deeply upon impact as it does in the model with observable state. When the state uncertainty disseminates after worse data on consumption are revealed, the equity return continues to drop before it reverts gradually. Similarly, when the disaster episode ends at period 10, it takes agents a couple of periods to be sure that the economy has come back to the normal state. The rebound of the equity return is thus more gradual than it is in the parameter uncertainty model.

Now we come back to our benchmark model (Figure 2, Panel a) in which both state and parameter uncertainty are present. Compared with the model with only state uncertainty, the equity premium in normal state ($I_t = 0$) is higher, the stock market crash and boom are larger in magnitude, and the equity return and the risk-free rate are more volatile during the disaster episode. The presence of parameter uncertainty in future consumption growth increases the equity premium and makes the equity return drop more when agents suspect the start of a disaster. Learning about the parameter during the disaster reduces parameter uncertainty over time, which causes the extra time-variations exhibited by the equity return and the risk-free rate. After the disaster ends, the resolution of both state and parameter uncertainty produces the more pronounced boom in the stock market.

Finally, we compare asset returns in the anticipated-utility-pricing model (Figure 2, Panel d) with those in the benchmark model (Figure 2, Panel a). The primary difference between the two panels is the size of stock market crashes and booms, which are significantly larger in the model with anticipated-utility pricing. The busts are larger because when agents first observe a large negative shock on consumption growth and think that θ_τ is likely to be very negative, they take their current beliefs as lasting forever in pricing assets in the anticipated-utility-pricing model, whereas in the benchmark model, they take into account that their current pessimistic view can be altered in the future if more favorable data are observed. Consequently, the equity return in the anticipated-utility-pricing model responds more aggressively at the onset of the disaster, resulting in a more severe crash in the stock market. Similarly, when the negative shock disappears after period 10, the revision of agents' beliefs about θ_τ in the anticipated-utility-pricing model yields a larger impact on the equity return than it does in the benchmark model. As a result, the stock market boom is more pronounced. Also notice that the equity premium in the normal state is much lower and the risk-free rate during the disaster is less volatile in Panel d than their counterparts in Panel a. The equity premium is lower in the normal state because the anticipated-utility-pricing model only prices state uncertainty and ignores the pricing of the risks generated by parameter learning. Thus, the equity premium in the normal state is close to that in the parameter uncertainty model and the risk free rate behaves similarly to its counterpart in the state uncertainty model.

3.2 Quantitative Results

In this subsection, we compute asset returns implied by our benchmark model and the other three model variants using simulated consumption data. The consumption data are simulated using the data generating process specified in equation (2.1) with no disaster realization throughout (I_t is set to zero at all periods).[26]

The aim of this exercise is to compare moments of the model-implied returns to those of U.S. data from 1948 to 2008. Our return data are annual returns of the U.S. one-month Treasury bill and U.S. major stock indexes (NYSE/AMEX/NASDAQ), obtained from the Center for Research in Security Prices (CRSP) database. We use the annualized monthly Consumer Price Index (CPI) data from Bureau of Labor Statistics (BLS) to deflate nominal returns. Arguably, the U.S. economy did not suffer from any macroeconomic disaster during the period from 1948 to 2008.[27] Thus, this exercise highlights the ability of our benchmark model to generate a high equity premium and reasonable return variations even without any realization of disasters, a feature that distinguishes our model from many others in the literature on rare disasters.

3.2.1 Moments

Table 2 reports the means and standard deviations of risk-free rates and equity returns, expressed in percentages. R^f denotes the risk-free rate, and R^e denotes the equity return. $E(\cdot)$ is the mean of returns, and $\sigma(\cdot)$ is the standard deviation of returns. Panel I reports moments of the actual data. Panel II reports the corresponding model-implied return moments. We first explain the results of equity returns and then discuss the risk-free rates.

Row 1 in Table 2 shows that the benchmark model implies an equity premium similar to its data

[26] Since there is no disaster realization in the simulation, the disaster parameter θ_τ is never drawn. In computing the state uncertainty model, we assume that agents take the true θ_τ at its prior mean value. Each simulated consumption path has 60 periods to match the length of the U.S. return data. We repeat the simulation 10,000 times and take the averages of asset return moments across simulations.

[27] The posterior probability that the U.S. economy was in a disaster is nearly zero from 1948 to 2008, according to Figure 4 in Nakamura et al. (2013).

Table 2: Asset pricing moments

Moments	$E(R^f)$	$\sigma(R^f)$	$E(R^e)$	$\sigma(R^e)$	$E(R^e) - E(R^f)$
			Panel I: Data		
U.S. data from 1948 to 2008	1.14	2.35	8.54	18.50	7.40
			Panel II: Models		
1. Benchmark	1.13	3.15	9.20	17.13	8.07
2. Parameter uncertainty	3.07	0.00	5.96	13.61	2.89
3. State Uncertainty	3.20	0.26	5.43	14.32	2.23
4. Anticipated-utility-pricing	3.42	0.34	7.15	29.11	3.73

The data are computed using annual returns of the U.S. one-month Treasury-bill and U.S. major stock indexes (NYSE/AMEX/NASDAQ) from 1948 to 2009, obtained from CRSP. All nominal returns are deflated using the annualized monthly CPI data from BLS.

counterpart and accounts for 92.5% of the observed volatility of equity returns. Although there is no realized disaster shock throughout the simulated consumption samples, the presence of both state and parameter uncertainty makes it easy for agents to temporarily confuse a transitory negative shock with a disaster shock. The temporary confusion often triggers joint learning about the disaster state and the severity of a potential disaster. State uncertainty and parameter uncertainty interact with each other, which enhances the learning effects and generates significant time-varying disaster risk endogenously. As the benchmark model prices the joint effects of state and parameter uncertainty, the time-varying disaster risks are fully reflected in asset returns, which enables the model to generate high equity premium and high volatility of equity returns without relying on the realization of disaster shocks in the consumption process.

Row 2 in Table 2 reports the results from the parameter uncertainty model. Recall that agents are assumed to know the current state of the economy perfectly while they are still exposed to the disaster risk for the next period. In addition, the disaster risk incorporates time-invariant parameter uncertainty in the sense that parameter θ_τ is viewed as a random draw from $F(\theta)$ each period. Therefore, the equity premium in this model reflects the price of the disaster risk with parameter uncertainty.[28] However, without any realization of disasters throughout the simulated consumption path, the time-invariant state and the time-invariant parameter uncertainty imply that the PDR of the equity remains constant. As a result, all the variation in equity returns comes from shocks to the dividend process ($\sigma_d \varepsilon_t$).

Row 3 in Table 2 shows the results from the state uncertainty model. Compared to the parameter uncertainty model, the volatility of equity returns is larger because of the time-varying state uncertainty. However, because the state uncertainty model assumes away parameter uncertainty, the equity premium is lower than its counterpart in the parameter uncertainty model. This result reveals that a standard hidden Markov regime-switching model typically requires either high risk aversion or a large mean shift in consumption growth to match excess equity returns.

Row 4 in Table 2 shows the results from the anticipated-utility-pricing model. Compared to the benchmark model, the model with anticipated-utility-pricing agents yields a significantly higher equity return volatility and a much lower equity premium. To understand these results, we borrow the intuition obtained in Section 3.1 on the return dynamics during a disaster episode. First, since anticipated-utility-pricing agents view their current beliefs about parameter θ_τ as lasting forever, the equity price responds more aggressively to changes in the current beliefs than it does when agents understand that their future beliefs about parameter θ_τ will be updated with new consumption data. The equity return is thus more volatile in the former case. Second, when agents are learning from the consumption data simulated without any disaster shock, the posterior means of their beliefs about parameter θ_τ are higher than its unconditional mean most of the time. Hence, anticipated-utility-pricing agents often price assets with the beliefs that the current disaster is a mild one, whereas agents in the benchmark model are aware that the disaster could have much worse long-term damage than the current consumption data suggest. Ignoring such risks associated with belief-updating about the disaster parameter explains why the anticipated-

[28]In this sense, the parameter uncertainty model resembles the model used in Nakamura et al. (2013).

utility-pricing model has a much lower equity premium than the benchmark model.

We now compare the risk-free rates across various models. A notable feature of our benchmark model is that it implies much lower and more volatile risk-free rates than the other three models. Once again, this feature can be attributed to the fact that agents take into account their belief-updating in the future when they price assets. As discussed before, an awareness of future changes in the posterior mean of beliefs about θ_τ implies a higher risk in future consumption. The higher risk not only generates higher equity returns but also increases the demand for the risk-free asset and lowers the risk-free returns. Furthermore, this risk varies with how long agents have been actively learning. The longer the learning trigger is activated (i.e., $S_t = on$), the less parameter uncertainty is present because of the accumulated consumption data. When agents are more confident about their estimate of θ_τ, the posterior mean of their future beliefs about θ_τ will be less sensitive to new data, which reduces agents' perceived risk in future consumption stemming from their future belief-updating. This mechanism creates an additional time dependence of risk-free rates and is only at work when both state uncertainty and parameter uncertainty are present in learning and pricing. As a result, risk-free rates are much more volatile in our benchmark model.

Note that the U.S. data also feature low and relatively volatile risk-free rates. Although we choose the value of the time discount factor β to match the first moment of risk-free rates in the data, the fact that the volatility of our model-implied risk-free rates is similar to its data counterpart provides additional support for our benchmark model.

The contrast between the anticipated-utility-pricing model and the benchmark model in Panel II in Table 2 shows that the finding of Cogley and Sargent (2008) does not hold in the context of our asset pricing model. Cogley and Sargent find that using the exact Bayesian approach yields similar asset pricing results as using the anticipated utility framework based on Kreps (1998) that neglects parameter uncertainty. However, in their conclusion, they anticipate our finding by acknowledging that "anticipated-utility modeling may be problematic for applications in finance when high risk aversion is assumed." Furthermore, the excellent approximation by the anticipated-utility approach in Cogley and Sargent (2008) is achieved when agents have CRRA utility. When agents have Epstein-Zin preferences, our results demonstrate that parameter uncertainty can play an important role in asset pricing. Similar results are also obtained by Collin-Dufresne et al. (forthcoming).

3.2.2 Return predictability

Since Campbell and Shiller (1988) and Fama and French (1988), predictive regressions that use dividend-price ratio as a predictor for future excess returns have become popular in testing models in the asset-pricing literature. In this subsection, we run predictive regressions of excess equity returns on lagged dividend price ratios using both the data and the model-implied asset returns of the four models:

$$\ln R^e_{t\to t+k} - \ln R^f_{t\to t+k} = \alpha_k + \beta_k \ln\left(D_t/P_t\right) + \nu_{t+k}, \tag{3.1}$$

where the left-hand side is the future excess return on equity, k denotes the forecasting horizon, and D_t/P_t is the dividend price ratio. Following the common practice in the literature, we run predictive regressions over horizons ranging from one to five years.

Table 3 reports the slope coefficients β_k and R^2 from the regressions. The regression results from the data shown in Panel I confirm the findings in the literature that the dividend price ratio has significant predictive power over future excess returns. Moreover, both the estimate of β_k and the value of R^2 increase with the forecasting horizon.

The results from the four models are shown in Panel II of Table 3. Both the benchmark model and the anticipated-utility-pricing model have positive and significant β_k at each forecasting horizon, although the values of β_k are higher than their data counterparts. Moreover, β_k increases with the forecasting horizon, a pattern that is consistent with the data. These results confirm the finding of Timmermann (1996) and Timmermann (2001) that the learning effect on stock price dynamics is an intuitive candidate for explaining the predictability of excess returns. Timmermann (1996) offers an intuitive explanation for this finding:

> An estimated dividend growth rate which is above its true value implies a low dividend yield as investors use a large mark-up factor to form stock prices. Then future returns

Table 3: Excess return predictive regressions

Forecasting horizon k		1	2	3	4	5
				Panel I: Data		
U.S. data from 1948 to 2008	β	0.10***	0.18***	0.24***	0.29***	0.38***
	R^2	0.08	0.12	0.16	0.19	0.23
				Panel II: Models		
1. Benchmark	β	1.14***	1.38***	1.48***	1.53***	1.55***
	R^2	0.26	0.23	0.20	0.17	0.15
2. Parameter uncertainty	β	0.00	0.00	0.00	-0.01	0.00
	R^2	0.00	0.00	0.00	0.00	0.00
3. State Uncertainty	β	0.76	1.11	1.28	1.36	1.41
	R^2	0.03	0.04	0.04	0.05	0.05
4. Anticipated-utility-pricing	β	0.87***	0.95***	1.00***	1.02***	1.03***
	R^2	0.32	0.29	0.27	0.25	0.23

* significant at the 90% level, ** 95% level, *** 99% level.

> will tend to be low since the current yield is low and because the estimated growth rate of dividends can be expected to decline to its true value, leading to lower than expected capital gains along the adjustment path. (p. 524)

Hence, the predictive power of the dividend-price ratio stems from equity prices reacting more strongly to changes in dividend growth as a result of agents' learning. Consistent with this view, the R^2 values in the anticipated-utility-pricing model are higher than those in the benchmark model since equity prices are more responsive to changes in dividend growth for the anticipated-utility-pricing agents. Relative to the data, the R^2 values in the anticipated-utility-pricing model are too high at all horizons, whereas the R^2 values in our benchmark model are closer to what the data suggest. None of the models can reproduce the pattern in the data that the predictive power measured by the R^2 value monotonically increases with the forecasting horizon. The dividend-price ratio has no predictive power in the model with only parameter uncertainty since the ratio is constant in the absence of disaster realizations. The dividend-price ratio in the state uncertainty model exhibits some predictive power, with insignificant β_k and rather low R^2 values.

4 Conclusion

In this paper, we incorporate joint learning about state and parameter in a rare disaster model and study its implications on asset prices in an endowment economy. The interaction of state and parameter uncertainty in learning and asset pricing enables our model to generate high equity premium and sizable volatility in equity returns without relying on the actual occurrence of disasters in the economy or exogenous variations in disaster probability, a great improvement over the existing literature on rare disasters. We also show that having both state learning and parameter learning as priced risks is essential for the model to perform well in matching means and volatilities of the U.S. equity returns and risk-free rates from 1948 to 2008, a period during which the U.S. economy did not suffer from any macroeconomic disaster.

In our benchmark model, only aggregate consumption is observable by agents. Adding other observable variables such as investment, or output, is likely to further improve our model's performance in matching asset pricing facts because agents' beliefs about future consumption will be more volatile.

In modeling the consumption process, we abstract from the short-term effect of a disaster and focus only on its long-term effect. However, empirical evidence in Nakamura et al. (2013) shows that the short-term damage of a disaster is on average twice as large as its long-term damage. Extending the current model to incorporate the short-term effect of a disaster could be interesting since not only would asset prices be more responsive to changes in agents' beliefs, but agents' beliefs could also be more volatile when the persistence of a disaster shock is closer to that of a normal-time shock. We leave this extension for future research.

References

Abel, Andrew B., "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics*, 1999, *43* (1), 3–33.

Bansal, R. and A. Yaron, "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 2004, *59* (4), 1481–1509.

Barro, Robert J., "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 2006, *121*, 823–866.

Beeler, Jason and John Y Campbell, "The long-run risks model and aggregate asset prices: an empirical assessment," *Critical Finance Review*, 2012, (1), 141–182.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S Goldstein, "Explaining asset pricing puzzles associated with the 1987 market crash," *Journal of Financial Economics*, 2011, *101* (3), 552–573.

Bianchi, Daniele, "Parameter Learning, Uncertainty Shocks and the Variance Risk Premium," 2013.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J Terry, "Really uncertain business cycles," Technical Report, National Bureau of Economic Research 2012.

Bloom, Nick, "The Impact of Uncertainty Shocks," *Econometrica*, 2009, *77* (3), 623–685.

Brandt, M.W., Q. Zeng, and L. Zhang, "Equilibrium Stock Return Dynamics under Alternative Rules of Learning about Hidden States," *Journal of Economic Dynamics and Control*, 2004, *28* (10), 1925–1954.

Campbell, John Y. and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, April 1999, *107* (2), 205–251.

Cecchetti, Stephen G, Pok-Sang Lam, and Nelson Mark, “Mean reversion in equilibrium asset prices,” 1990.

_ , _ , and _ , “The equity premium and the risk-free rate: matching the moments,” *Journal of Monetary Economics*, 1993, *31* (1), 21–45.

_ , _ , and _ , “Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?,” *American Economic Review*, 2000, *90*, 787–805.

Chen, H. and M. Pakos, “Rational Overreaction and Underreaction in Fixed Income and Equity Markets: The Role of Time-Varying Timing Premium,” *GSIA Working Papers*, 2007.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, “Measuring the 'Dark Matter' in Asset Pricing Models,” *Working paper*, 2015.

Cogley, Timothy and Thomas J Sargent, “Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making,” *International Economic Review*, 2008, *49* (1), 185–221.

Collin-Dufresne, P., M. Johannes, and L. Lochstoer, “Parameter Learning in General Equilibrium: The Asset Pricing Implications,” *American Economic Review*, forthcoming.

Constantinides, George M, “Comment on Barro and Ursua,” *Brookings Papers on Economic Activity*, 2008.

Drechsler, Itamar and Amir Yaron, “What's vol got to do with it,” *Review of Financial Studies*, 2011, *24* (1), 1–45.

Epstein, Larry G. and Stanley Zin, “Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 1989, *57*, 937–969.

Farhi, Emmanuel and Xavier Gabaix, “Rare Disasters and Exchange Rates,” *NBER Working Paper 13805*, 2008.

Fernández-Villaverde, J., P. Guerrón-Quintana, J.F. Rubio-Ramirez, and M. Uribe, "Risk matters: The real effects of volatility shocks," *American Economic Review*, 2011, *101* (6), 2530–61.

Fulop, Andras, Junye Li, and Jun Yu, "Self-Exciting Jumps, Learning, and Asset Pricing Implications," *Review of Financial Studies*, 2014, p. hhu078.

Gabaix, Xavier, "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 2012, *127* (2), 645–700.

Ghosh, Anisha and George M Constantinides, "The Predictability of Returns with Regime Shifts in Consumption and Dividend Growth," *NBER Working Paper 16183*, 2010.

Gillman, Max, Michal Kejak, and Michal Pakos, "Learning about Disaster Risk: Joint Implications for Consumption and Asset Prices," *CERGE-EI Working Paper Series*, 2014, (507).

Gourio, Francois, "Disaster Risk and Business Cycles," *American Economic Review*, 2012, *102* (6), 2734–66.

_ , Michael Siemer, and Adrien Verdelhan, "International Risk Cycles," *Journal of International Economics*, 2013, *89* (2), 471–484.

Guo, Kai, "Exchange Rates and Asset Prices in An Open Economy with Rare Disasters," *Working Paper*, 2009.

Jagannathan, Ravi and Binying Liu, "Dividend Dynamics, Learning, and Expected Stock Index Returns," *NBER Working Paper*, 2015.

Johannes, M., L. Lochstoer, and Y. Mou, "Learning about Consumption Dynamics," *Journal of Finance*, forthcoming.

Ju, Nengjiu and Jianjun Miao, "Ambiguity, Learning, and Asset Returns," *Econometrica*, 2012, *80* (2), 559–591.

Julliard, Christian and Anisha Ghosh, "Can Rare Events Explain the Equity Premium Puzzle?," *Review of Financial Studies*, 2012, *25* (10), 3037–3076.

Kandel, Shmuel and Robert F Stambaugh, "Asset returns and intertemporal preferences," *Journal of Monetary Economics*, 1991, *27* (1), 39–71.

Kasa, Kenneth and In-Koo Cho, "Learning and Model Validation," *Society for Economic Dynamics 2011 Meeting Papers*, 2011, (1086).

Koulovatianos, Christos and Volker Wieland, "Asset pricing under rational learning about rare disasters," *CEPR Working Paper*, 2011.

Kreps, David M, "Anticipated Utility and Dynamic Choice," *Econometric Society Monographs*, 1998, *29*, 242–274.

Lewis, K.K., "Can Learning Affect Exchange-Rate Behavior?: The Case of the Dollar in the early 1980's," *Journal of Monetary Economics*, 1989, *23* (1), 79–100.

Mehra, Rajnish and Edward Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Nakamura, Emi, Jon Steinsson, Robert J. Barro, and Jose F. Ursua, "Crises and Recoveries in an Empirical Model of Consumption Disasters," *American Economic Journal: Macroeconomics*, 2013, *5* (3), 35–74.

Orlik, Anna and Laura Veldkamp, "Understanding uncertainty shocks and the role of black swans," *NBER Working Paper*, 2014.

Piazzesi, M and M Schneider, "Trend and Cycle in Bond Risk Premia," *Working Paper Stanford University*, 2010.

Pollak, Moshe, "Optimal detection of a change in distribution," *The Annals of Statistics*, 1985, pp. 206–227.

— **and Alexander G Tartakovsky**, "Exact optimality of the Shiryaev-Roberts procedure for detecting changes in distributions," in "Information Theory and Its Applications, 2008. ISITA 2008. International Symposium on" IEEE 2008, pp. 1–6.

Rietz, Thomas A., "The Equity Risk Premium: A Solution," *Journal of Monetary Economics*, 1988, *22*, 117–131.

Roberts, SW, "A comparison of some control chart procedures," *Technometrics*, 1966, *8* (3), 411–430.

Schaal, Edouard, "Uncertainty, productivity and unemployment during the Great Recession," *Manuscript, Federal Reserve Bank of Minneapolis*, 2012.

Shiryaev, Albert N, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, 1963, *8* (1), 22–46.

Timmermann, A., "Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning," *The Review of Economic Studies*, 1996, *63* (4), 523–557.

Timmermann, Allan, "Structural Breaks, Incomplete Information, and Stock Prices," *Journal of Business & Economic Statistics*, 2001, *19* (3), 299–314.

Veronesi, P., "Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model," *Review of Financial Studies*, 1999, *12* (5), 975–1007.

— , "The Peso Problem Hypothesis and Stock Market Returns," *Journal of Economic Dynamics and Control*, 2004, *28* (4), 707–725.

Wachter, Jessica A., "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?," *Journal of Finance*, 2013, *68* (3), 987–1035.

Weitzman, M.L., "Subjective Expectations and Asset-Return Puzzles," *The American Economic Review*, 2007, *97* (4), 1102–1130.

Appendix

A Robustness

This section presents robustness checks for asset pricing moments and predictive regressions.

A.1 Asset pricing moments

Table 4 shows the results of robustness checks on asset pricing moments. The results from the benchmark model are shown as a reference in the first row. We first examine how the value of relative risk aversion affects asset return moments. In the benchmark calibration, the relative risk aversion is set to be 8. Here we report the results computed using a higher value of relative risk aversion ($\gamma = 9$) in the second row and the results with a lower value of relative risk aversion ($\gamma = 7$) in the third row. Intuitively, an increase (decrease) in risk aversion makes asset prices more (less) responsive to changes in agents' beliefs. As a result, both the equity premium and the equity return volatility increase (decrease).

Next, we consider the cases of leverage ratio higher ($\lambda = 2.5$) and a lower($\lambda = 1.5$) than in the benchmark calibration ($\lambda = 2$). Because a change in leverage ratio only affects levered equity returns, the risk-free rates remain unchanged. An increase in leverage ratio amplifies the effect of consumption shocks on equity prices, therefore resulting in a higher equity return, a higher equity premium and a higher equity return volatility.

Although the results are found to be sensitive to both the relative risk aversion γ and the leverage ratio λ, the values of γ and λ in the calibration of the main text are both set at conservative levels ($\gamma = 8$ and $\lambda = 2$) by the standard of the literature. It shows that with the help of learning, our simple benchmark model can match the data moments with reasonable parameter values even in the absence of disaster realizations.

Table 4: Robustness - asset pricing moments

	$E(R_f)$	$\sigma(R_f)$	$E(R^e_{lev})$	$\sigma(R^e_{lev})$	$E(R^e_{lev}) - E(R_f)$
Benchmark	1.13	3.15	9.20	17.13	8.07
Higher risk aversion $\gamma = 9$	0.40	3.94	10.08	17.67	9.60
Lower risk aversion $\gamma = 7$	1.94	2.14	8.04	16.35	6.08
Higher leverage ratio $\lambda = 2.5$	1.13	3.15	10.82	18.92	9.70
Lower leverage ratio $\lambda = 1.5$	1.13	3.15	7.33	15.28	6.19

A.2 Predictive regressions

Let us now turn to the predictive regressions. Table 5 shows that the basic patterns of the results from our benchmark model do not change significantly in the robustness checks: the estimates for β_k are all positive and significant, the β_k value increases with the forecasting horizon, and the R^2 value decreases with the forecasting horizon.

Table 5: Robustness – excess return predictive regressions

Forecasting horizon k		1	2	3	4	5
Benchmark	β	1.14***	1.38***	1.48***	1.53***	1.55***
	R^2	0.26	0.23	0.20	0.17	0.15
High risk aversion	β	1.23***	1.48***	1.59***	1.64***	1.66***
$\gamma = 9$	R^2	0.31	0.28	0.24	0.20	0.18
Low risk aversion	β	1.05***	1.28***	1.39***	1.44***	1.46***
$\gamma = 7$	R^2	0.20	0.17	0.15	0.13	0.12
Higher leverage ratio	β	1.07***	1.30***	1.40***	1.45***	1.47***
$\lambda = 2.5$	R^2	0.30	0.28	0.25	0.22	0.19
Lower leverage ratio	β	0.32***	1.61***	1.75***	1.81***	1.83***
$\lambda = 1.5$						
	R^2	0.20	0.17	0.15	0.13	0.11

* significant at the 90% level, ** 95% level, *** 99% level.