Learning About Rare Economic Disasters and Asset Prices*

Du Du
Hong Kong University of Science and Technology

Redouane Elkamhi
University of Toronto

November, 2012

Abstract

This article proposes a disaster-type model in which investors learn about disaster probabilities as well as subsequent economic recoveries. Investors not only learn from news, but also from the absence of news. No news during good economic times is perceived as good news, while no news during bad economic times is perceived as bad news. Investors' rational reactions toward disasters and recoveries help reconcile various asset pricing puzzles. Our model is solely based on US consumption experience, does not require jumps in consumption levels, and is consistent with empirically low consumption autocorrelation. It generates time-varying disaster and recovery intensities helpful to explain equity volatility. Lastly, it differentiates jump from diffusive risk premiums which are instrumental in explaining option market regularities.

JEL code: G12, G13, G14

Key words: learning, economic disaster, recovery, jump risk

*Email: dudu@ust.hk; Tel: (852)-2335-5049; Fax: (852)-2358-1749. Email: redouane.elkamhi@rotman.utoronto.ca. We thank Hengjie Ai, Peter Christoffersen, Sudipto Dasgupta, Alexander David, Jan Ericsson, Xuewen Liu, Raunaq Pungaliya, Mark Seasholes, Pietro Veronesi, Pengfei Wang, Kuo-chiang Wei, Lu Yang, Chu Zhang, and seminar participants at Korea University, SKK, Vienna, Amsterdam, Paris and Atlanta for helpful discussions and comments.

I Introduction

Disaster research within the paradigm of Rietz-Barro hypothesis has significantly improved our understanding of asset pricing puzzles.[1] A common feature of most of this literature is that investors do not learn through time about disaster probabilities. In this study, we investigate the implications of learning about the probability of economic disasters on the risk premium embedded in bonds, stocks, and options markets. We show that allowing for learning about disasters reconciles various asset pricing regularities in a model that does not require jumps in consumption level and that is solely based on US consumption experience.[2]

We propose a model that builds upon the Rietz-Barro framework while not being subject to the common criticism of traditional disaster models (TDMs) (See, for instance, Constantinides (2008); Mehra and Donaldson (2008)). Consumption in our model is continuous and disasters are treated as structural changes. This modeling choice has many advantages. First, TDMs do not allow for a recovery period which takes the form of a switch from the disaster phase to the normal phase in our model.[3] Second, TDMs assume that the entire drop in consumption upon disaster occurs over a single time period. By modeling a disaster as a separate economic phase, consumption declines are unfolded over several years. Third, in TDMs, stock market crashes are associated with large jumps in the consumption level. In contrast, consumptions are smooth in the data. Our model allows for stock market crashes to occur occasionally while consumption follows a continuous path.

Our study starts with the observation that economic disasters are infrequent by their nature and it is natural that investors forget about them after an extended period of normal

[1] See Barro (2006, 2009) and Nakamura, Steinsson, Barro, and Ursua (2011) for implication of disasters on the risk free rate and equity premium. In subsequent studies, Wachter (2011) focuses on excess volatility for the aggregate stock; Gabaix (2011) tackles ten puzzles in financial economics; Gourio (2011) develops a model able to explain stylized facts in the credit market; Du and Elkamhi (2011) presents a systematic study on the pricing of options and defaultable bonds.

[2] Longstaff and Piazzesi (2004) illustrate that disasters based solely on the US experience can explain around a third (2.6%) of the documented equity premium. Consequently, the more recent traditional disaster literature relies extensively on the international drops in consumption assembled in Barro and Ursua (2008) to motivate the choice of large magnitude in consumption drops (a mean around 30%, almost three times the drop in US consumption during the Great Depression in the US).

[3] Nakamura, Steinsson, Barro, and Ursua (2011) provide an empirical study of recoveries from disasters and conclude a mean net disaster magnitude of 15% based on international evidence. Using a theoretical framework, Gourio (2008) shows that recoveries reduce the implied equity premium when the intertemporal elasticity of substitution is close to that used in the long-run risk literature (e.g., Bansal and Yaron (2004)).

times. But when disasters strike, investors react by showing excessive concern about the possibility of future disasters. In an extended version of the model, we also allow for learning about the intensity of economic recovery. Upon recovery, investors become more optimistic and update their assessment of the posterior intensity of economic recovery upward.

During each economic phase, our model maintains comparable asymmetric investors' reactions as in Veronesi (1999). In addition, the present study introduces a novel learning mechanism. Investors not only interpret the arrival of news but also interpret the absence of news in an asymmetric way. More explicitly, the model captures the idea that during the normal phase, *no news is perceived as good news.* In other words, in the absence of disasters, investors update their posterior assessment of disaster probabilities downward. Further, the longer investors experience normal economic conditions, the larger the increase in asset prices. However, when a disaster strikes, two compounding effects take place. First, investors become paranoid about future disasters' likelihood and dramatically increase their posterior belief about the disaster intensity. Second, in the extended model, *no news is perceived as bad news* during the disaster phase. The longer the disaster period persists, the more pessimistic investors become concerning the potential economic recovery. Consequently, they continuously and negatively update their posterior assessment of the recovery rate, resulting in a large and persistent drop in asset prices.[4]

Under recursive utility considered in this article, investors' learning as described above adds to the price of disaster risk and leads to more dramatic stock market crashes. First, preference for earlier resolution of intertemporal risks implies that bad news about consumption growth raises the agent's marginal utility. Learning strengthens this effect since the agent becomes more concerned about disasters in the future after a recent strike. Second, pessimism about the future leads to further depression of the aggregate equity prices beyond that implied by the worsening economic condition. The first effect creates the strong precautionary saving motive which helps resolve the risk-free rate puzzle (e.g., Weil, 1989). By combining the two effects, our model generates the extra learning-induced pre-

[4] With few exceptions (e.g., Collin-Dufresne, Goldstein, and Helwege (2008, CDGH); Benzoni, Collin-Dufresne, and Goldstein (2011, BCDG)) in which the agent's belief about future jumps is updated after the arrival of a Poisson process, the learning literature focuses largely on learning about drifts. Examples include David (1997), Veronesi (1999), and Veronesi (2000), and later extensions by David and Veronesi (2002), Pastor and Veronesi (2006), David (2008) and Ai (2011). All these models feature the continuous arrival of news and continuous belief updates. In our framework, news arrives in a discrete manner but investors also continously update their beliefs in the absence of news.

mium that helps to resolve the equity premium puzzle (e.g., Mehra and Prescott, 1985)[5].

As an important advantage, learning about disaster intensity endogenously generates time-varying disaster rates under the filtered measure. This is a desirable property in matching the second moment of assets returns (e.g., Wachter (2011); Gabaix (2011)).[6] We show that our model can deliver a reasonable match of empirically observed volatility of asset returns. The model-implied equity return volatility is monotonically decreasing in the posterior probability of the high disaster-intensity state (denoted by π). This effect, taken together with the model implied increasing and convex asset prices function, provides a rationalization of the "leverage" effect observed in the data.

The performance of the model for both bond and equity markets sheds light on an interesting finding concerning the long-run risk literature. Recursive utility when combined with learning can explain asset prices without large persistent variation in consumption moments. Aggregate consumption follows a random walk in our setup, and the model-implied first-order autocorrelation is around 0.28. This magnitude is far less than 0.5 considered in the long-run risk literature, and it is consistent with the recent evidence by Beeler and Campbell (2009, around 0.2). This level is also close to the 0.25 autocorrelation obtained by time-averaging a continuous-time random walk in Working (1960). Beeler and Campbell (2009) also report that higher-order consumption autocorrelations, particularly the third and fourth lags, are strikingly low relative to the predictions of the traditional long-run risk models. In our setup, higher order autocorrelations are all close to zero. We argue that, given the lack of strong evidence in support of persistent fluctuations in consumption growth and its higher moments, the learning mechanism emphasized in this paper contributes to the resolution of asset pricing puzzles while maintaining close ties with the observed consumption dynamics.

We next turn our attention to the model's implications on equity index options. The learning-induced compensation drives up the jump-risk premium as a fraction of the total premium, hence the potentiality to effectively address premiums implicit in option data. To put it differently, Liu, Pan, and Wang (2005, LPW) argue that for an equilibrium model

[5] Ai (2010) also uses recursive utility combined with learning. Unlike our paper, he adopts the learning channel as in Veronesi (1999, 2000). In his model, learning creates a positive covariance between the realized return and the expected return on the production technology which induces a higher diffusive risk premium. In this paper, learning creates a positive covariance between the expected arrivals of future events and their actual realizations, which induces higher compensations for jump risks. In addition to introducing a new learning channel, our model-implied jump premium helps jointly explain aggregate equity and its derivatives, the latter of which is not addressed in Ai (2010).

[6] Unlike the present setup, these models impose the time varying disaster intensity in an exogenous way.

to explain option premiums, it is important that the model differentiates the pricing of jump risks from that of diffusive risks. In our model, learning about economic disasters provides an additional channel that controls the pricing of disasters separately from that of diffusive shocks. Quantitively, our model generates implications of the average ATM premium (defined as the difference between ATM), the average return volatility, the average smirk premium (defined as the difference between 10% OTM and ATM volatility), as well as the average 10% OTM volatility that are all consistent with their empirical levels.

In a related research, Benzoni, Collin-Dufresne, and Goldstein (2011, BCDG)) also studies the impact of learning on option pricing. Their paper is the first that successfully explains the shift in the shape of the implied volatility smirk before and after the 1987 crash. Similar results can be generated by our setup if we treat 1987 event as a trigger of change in economic states. Relative to BCDG, our article distinguishes itself along two key dimensions. First, in contrast to the permanent jump assumed in BCDG, we allow for economic recovery which enriches our model dynamics supported by empirical evidences. Second, in our extended model, investors learn also about the intensity of economic recovery which generate novel asymmetric learning mechanism as discussed above.

In summary, accounting for investors learning about economic disasters provides an extension to disaster based research. The simple model we propose is able to simultaneously explain various regularities in different markets. The pricing performance of our model comes along with four realistic improvements. First, our model is not subject to some of the aforementioned criticism of the TDM. Second it is solely based on the US consumption experience and more importantly consistent with empirically observed consumption dynamics. Third, it endogenously generates time varying disaster and recovery intensities. Fourth, the model differentiates the pricing of jump risks from that of diffusive risks which is desirable for derivative pricing.

For completeness, we also provide an extension of the model in which investors update their beliefs about the intensity of economic recovery. Following the switch from the disaster phase to the normal phase, the aggregate stock market experiences an upward jump. In this setup, investors' reactions to the arrival of disaster is further strenghtened by their reactions to a period of no news during the disaster phase. More explicitly, getting paranoid upon the strike of disaster and growing more pessimistic waiting for economic recovery, when combined, create a strong positive covariance between assets' expected returns and their actual realizations. As a result, the general model better explains all the three markets through introducing novel and more realistic investors' reactions during the bad economic conditions.

In an influencial paper, Weitzman (2007) argues convincingly for the importance of learning about parameter uncertainty in explaining asset pricing puzzles. Without considering disasters, he shows that Bayesian updating of unknown structural parameters inevitably adds a permanent tail-thickening effect to posterior expectations. In this article, we emphasize learning about unknown parameters related to disasters, which we show not only provides an expected-utility solution for asset pricing puzzles but also uncovers novel investors' reactions to news as well as to the absence of news.

The rest of the paper is organized as follows. Section II and III present the setup of the base model and derive the model-implied pricing. Section IV describes the model mechanism. In Section V, we provide the calibration and discuss the quantitative results. Section VI extends the base model to the more general case where the agent also learns about the recovery rate. Section VII concludes. All technical details are in the Appendices.

II The setup

A. Preference and consumption process

Assume the existence of a representative agent whose preference is described by the stochastic differential utility (SDU) developed in Duffie and Epstein (1992), which is the continuous-time version of the recursive utility considered in Kreps and Porteus (1978) and Epstein and Zin (1989). Given the consumption process $\{C_u : u \geq 0\}$, the period-t utility of the agent, denoted by J_t, is defined recursively as

$$J_t = E_t \left[\int_t^{\infty} f\left(C_u, J_u\right) du \right]. \tag{2.1}$$

In the above equation, $f\left(.\right)$ denotes the normalized aggregator defined as

$$f\left(C_t, J_t\right) = \frac{\beta}{1 - 1/\psi} \frac{C_t^{1-1/\psi} - \left[\left(1-\gamma\right) J_t\right]^{\frac{1-1/\psi}{1-\gamma}}}{\left[\left(1-\gamma\right) J_t\right]^{\frac{1-1/\psi}{1-\gamma}-1}}, \tag{2.2}$$

where β, $\gamma\left(\neq 1\right)$, and $\psi\left(\neq 1\right)$ denote respectively, the subjective discount rate, the risk aversion, and the elasticity of intertemporal substitution.[7]

In the TDMs (e.g., Rietz (1988); Barro (2006)), economic disasters are treated as jumps in consumption levels. In this paper, we model disasters as an alternative economic phase. A disaster strikes when a switch from the normal phase (n) to the disaster phase (d) occurs. Conditional on a given phase $s_t \in \{n, d, r\}$ at period t, the aggregate consumption evolves according to

$$\frac{dC_t}{C_t} = \mu\left(s_t\right)dt + \sigma\left(s_t\right)dB_t, \tag{2.3}$$

where dB_t is a standard Brownian common to both phases. Upon the strike of a disaster, consumption level does not jump but its process experiences a structural change. We set $\mu\left(d\right) < 0 < \mu\left(n\right)$ and $\sigma\left(d\right) > \sigma\left(n\right)$ implying that relative to the normal phase with positive growth rates, consumption shrinks on average with higher volatilities during the disaster phase. As described in the Introduction, treating a rare economic disaster as a structural change rather than an instantaneous permanent jump is more realistic. Further, it is arguable that investors can easily distinguish a rare economic disaster phase from normal economic conditions, much more than they can distinguish recessions from expansions. Finally, we model another phase, referred to as the recovery phase under which $\mu\left(r\right) > \mu\left(n\right)$. This specification is consistent with empirical evidences (e.g., Nakamura, Steinsson, Barro, and Ursua (2011)) that a more rapid growth usually follows at the end of the disaster period. For simplicity, we assume $\sigma\left(r\right) = \sigma\left(n\right)$.

Denote by λ, ν, and ξ, respectively, the switching intensity from the normal phase to the disaster phase, i.e., disaster probability, the switching intensity from the disaster phase to the recovery phase, i.e., the recovery probability, and the switching intensity from the recovery phase to the normal phase. Variations in the first two moments of consumption growth introduce the intertemporal macroeconomic risk. By adopting the recursive utility,

[7] A recent paper by Chen, Joslin and Tran (2012) studies heterogeneity of investors on their assessment of the disaster risk. While their model generates interesting implications that even a small fraction of optimistic investors can generate high risk-sharing capacity which attenuates the magnitude of disaster risk, we choose the representative agent framework at the presence of learning for three reasons. First, beliefs of both optimistic and pessimistic investors converge following the strike of disaster. This is consistent with observations that sentiment becomes highly contagious during bad times. Second, due to the mechanism of "learning from no news", belief convergence also occurs during normal times. Specifically, after an extended period of time without disasters, pessimistic investors, who feel they might be wrong, would revise downward their assessment of disaster likelihood more than their optimistic counterparts. Third, our choice enables us to provide a study on the impact of learning about disaster risks on asset prices and uncover novel learning mechanism.

the representative agent is concerned about how quickly intertemporal risks are resolved, which is determined by the speed at which news about future consumption growth arrives. The rate of news arrival is governed by the rate at which the distribution of economic phases converges to their steady state, where the convergence rate is given by $\lambda + \nu + \xi$.

B. Learning about the disaster rate

Since disasters are rare events and their occurrences are infrequent by nature, it seems plausible to treat their arrival intensity λ as inaccurately observed. Specifically, assume that the agent knows λ only takes two possible values, λ^G and λ^B, but she does not observe which one is realized. Setting $\lambda^G < \lambda^B$, we interpret the two values as the good and the bad $\lambda-$regime. We use the word regime to differentiate from phase which denotes the economic condition. The intensity λ switches from one value to another according to the following Markov chain:

$$\Pr\left(\lambda_{t+dt} = \lambda^B | \lambda_t = \lambda^G\right) = \phi_G dt, \tag{2.4}$$

$$\Pr\left(\lambda_{t+dt} = \lambda^G | \lambda_t = \lambda^B\right) = \phi_B dt, \tag{2.5}$$

where (ϕ_G, ϕ_B) are known parameters. The implied long-run disaster rate is thus

$$\bar{\lambda} = \frac{\phi_B}{\phi_G + \phi_B}\lambda^G + \frac{\phi_G}{\phi_G + \phi_B}\lambda^B. \tag{2.6}$$

In the base model, we assume that the recovery rate ν is accurately observed. The more general case where the agent also learns about ν is discussed in Section VI.

Since the agent does not know the actual $\lambda-$regime, she forms a posterior estimation of the disaster rate based on her observations during normal times. Denote by π_t the agent's posterior assessment that λ is in the good regime at period t, i.e. $\pi_t \equiv \Pr\left(\lambda_t = \lambda^G | \mathcal{F}_t\right)$. Conditional on the normal phase and by Theorem 19.6 of Lipster and Shiryaev (2001), we obtain the following dynamics for π_t:

$$d\pi_t = \mu_\pi dt + \pi_t \frac{\lambda^G - \lambda_{pt}}{\lambda_{pt}} d\hat{M}_t(n,d), \tag{2.7}$$

where

$$\mu_\pi \equiv \pi_t\left(\lambda_{pt} - \lambda^G\right) + (1-\pi_t)\phi_B - \pi_t\phi_G; \tag{2.8}$$

$d\hat{M}_t(n,d)$ is the Poisson process governing the strike of a disaster under the filtered measure with intensity λ_{pt};

$$\lambda_{pt} = \pi_t \lambda^G + (1-\pi_t)\lambda^B = \lambda^G + (1-\pi_t)\left(\lambda^B - \lambda^G\right) \tag{2.9}$$

which denotes the posterior estimation of λ. $\lambda_{pt} > \lambda^G$ since $\lambda^B > \lambda^G$. In particular, $\lambda_{pt} = \bar{\lambda}$ when π_t is at its long-run average of

$$\bar{\pi} \equiv \frac{\phi_B}{\phi_G + \phi_B}. \tag{2.10}$$

To understand the agent's learning behavior about the arrival of disasters, let's start from $\pi_t = \bar{\pi}$ under which μ_π reduces to $\pi\left(\lambda_p - \lambda^G\right) > 0$. If no disaster occurs within $[t, t+dt]$, i.e., $d\hat{M}_t(n,d) = 0$, the agent revises upward the likelihood of the good $\lambda-$regime with a low disaster rate such that π_t drifts upward by $\mu_\pi dt$. As a result, the posterior estimation of the disaster rate, λ_{pt}, drifts downward. Intuitively, when nothing happens over the next instant of time, the agent becomes less aware of the potential disaster by assigning a lower probability to its occurrence in the future. In other words, investors during the normal phase interpret the absence of news as good news.[8]

By contrast, if a disaster strikes within $[t, t+dt]$, the agent substantially revises downward the good $\lambda-$regime probability by

$$\pi_t^+ - \pi_t = \pi_t \frac{\lambda^G - \lambda_{pt}}{\lambda_{pt}} = \frac{\left(\lambda^G - \lambda^B\right)}{\lambda_{pt}} \pi_t (1-\pi_t) < 0, \tag{2.11}$$

where we've used (2.9) for the second equality. Consequently, λ_p jumps upward by: $\lambda_{pt}^+ - \lambda_{pt} = \left(\pi_t^+ - \pi_t\right)\left(\lambda^G - \lambda^B\right) > 0$. For the ease of notation, in the following we drop the time subscript for the states when there is no necessity to emphasize their time dependences. Intuitively, upon an actual strike of disaster, the agent becomes more concerned about intertemporal consumption risks by upgrading her assessment about the arrival of future disasters. For any given π, the agent updates λ_p by a larger amount when the difference

[8] In an unreported exercise, we also consider an extreme case that $\phi^G >> \phi^B$ and an intial π_t (conditional on the normal phase) very close to 1. With π_t serving as the only state varable in the base model, this scenario represents an agent who is currently in a very good state of the economy. However, she believes that the transition probability from the good $\lambda-$regime to the bad $\lambda-$ regime is extremely plausible. While it seems difficult to make the case that this "contrarian investor" is the representative agent of the economy, it turns out that the drift in equation (2.8) becomes negative in this case. Intuitively, as time goes on without any disaster this "contrarian" investor becomes more worried about the disaster that is strongly expected. As a result, she updates upward her assessment of the disaster likelihood.

between the two $\lambda-$regimes, $\left|\lambda^G - \lambda^B\right|$, is larger. Since disaster rates have direct impacts on risk compensations, the above learning behavior eventually finds its way into asset prices in the form of learning-induced premiums.

III Asset pricing implications

A. Pricing kernel

At the presence of learning, π_t or equivalently λ_{pt} serves as the state for pricing. By homogeneity, the agent's utility must be separable in consumption and the state, hence,

$$J_t = \frac{C_t^{1-\gamma}}{1-\gamma} \left[\beta I\left(\pi_t, s_t\right)\right]^{\theta}, \tag{3.1}$$

where $s_t \in \{n, d, r\}$ denotes the economic phase;

$$\theta \equiv \frac{1-\gamma}{1-1/\psi}. \tag{3.2}$$

It can be shown that $I\left(\pi_t, s_t\right)$ in (3.1) captures the wealth-consumption ratio in the economy at period t.[9]

Following Duffie and Epstein (1992), we show that the pricing kernel is given by

$$\Lambda_t = \exp\left(-\int_0^t \left[\beta\theta + \frac{1-\theta}{I\left(\pi_u, s_u\right)}\right] du\right) C_t^{-\gamma} I\left(\pi_t, s_t\right)^{\theta-1}, \tag{3.3}$$

which loads on both the aggregate consumption and the wealth-consumption ratio $I\left(.\right)$. By Ito's lemma for regime switching models (e.g., Mao and Yuan (2006)), the pricing kernel conditional on the normal phase and the disaster phase follow respectively:

$$\frac{d\Lambda_t}{\Lambda_t} = -r\left(\pi_t, n\right) dt - \gamma\sigma\left(n\right) dB_t + \left[\left(\frac{I\left(\pi_t^+, d\right)}{I\left(\pi_t, n\right)}\right)^{\theta-1} - 1\right] \left[d\hat{M}_t\left(n, d\right) - \lambda_p dt\right] \tag{3.4}$$

$$\frac{d\Lambda_t}{\Lambda_t} = -r\left(\pi_t, d\right) dt - \gamma\sigma\left(d\right) dB_t + \left[\left(\frac{I\left(\pi_t, r\right)}{I\left(\pi_t, d\right)}\right)^{\theta-1} - 1\right] \left[dM\left(d, r\right) - \nu dt\right] \tag{3.5}$$

[9]Proof is available upon request.

$$\frac{d\Lambda_t}{\Lambda_t} = -r\left(\pi_t, r\right)dt - \gamma\sigma\left(r\right)dB_t + \left[\left(\frac{I\left(\pi_t, n\right)}{I\left(\pi_t, r\right)}\right)^{\theta-1} - 1\right]\left[dM\left(r, n\right) - \xi dt\right] \tag{3.6}$$

where $d\hat{M}_t\left(n, d\right)$, $dM\left(d, n\right)$, and $dM\left(r, n\right)$ are the Poisson processes modeling the disaster, the recovery, and the switching from the recovery phase to the normal phase respectively; $r\left(\pi_t, .\right)$ denotes the short term interest rate.[10]

There are two differences between (3.4) and (3.5). First, while the disaster rate λ is not observable and needs to be "filtered" by its posterior estimation λ_{pt}, the recovery rate v is accurately observed in the base model. Second, whereas π jumps downward upon the strike of a disaster, it experiences no change following the economic recovery. In fact, the agent stops updating π_t conditional on $s_t = d$ since the disaster phase has already been realized. Similar observations applies to the comparison between (3.4) and (3.6). Section VI provides discussions on the agent's learning during the disaster time in a more general setup.

In Appendix A.1, we derive the model-implied restrictions for the wealth-consumption ratios (W/C) denoted by $I\left(\pi, n\right)$, $I\left(\pi, d\right)$, and $I\left(\pi, r\right)$ conditional on the normal, the disaster phase, and the recovery phase respectively. We then provide in Appendix B.1 the numerical procedure that simultaneously solves them using the collocation method (e.g., Miranda and Fackler, 2002). Under the calibration that ψ is greater than one, $I\left(.\right)$ reacts positively to π. Across the three phases, $I\left(., r\right) > I\left(., n\right) > I\left(., d\right)$ for any given degree of uncertainty. Intuitively, aggregate wealth depresses relative to its claim during disaster times when consumption grows at the lowest speed and it rises relative to its claim during recovery times when consumption grows at the highest speed. Substituting for W/C ratio, we derive in Appendix A.1 the expressions of the short rates under both the disaster and normal phases, and they are reported in (A.13)–(A.15).

B. Aggregate stock

Following Abel (1990) we model the aggregate dividend as the levered consumption, i.e., $D = C^{\eta}$, where $\eta > 1$ denotes the leverage. By Ito's lemma, the implied dividend process conditional on $s_t \in \{n, d\}$ follows:

$$\frac{dD_t}{D_t} = \mu_D\left(s_t\right)dt + \sigma_D\left(s_t\right)dB_t, \tag{3.7}$$

[10] To show that pricing kernel implied from our setup has regular properties, we apply simulation studies and find that the first up to the fourth (non-central) moments of the implied pricing kernel are all finite.

where

$$\mu_D(s_t) = \eta\mu(s_t) + \frac{1}{2}\eta(\eta-1)\sigma(s_t)^2, \tag{3.8}$$

$$\sigma_D(s_t) = \eta\sigma(s_t). \tag{3.9}$$

Define $I^S(\pi_t, s_t) = \frac{P_t}{D_t}$, where P_t is the price of the aggregate stock. By Ito's lemma with jumps, the stock returns conditional on the normal phase and the disaster phase follow

$$\frac{dP_t}{P_t} = \left(\mu_D(n) + \frac{1}{I^S(\pi_t,n)}\frac{dI^S(\pi_t,n)}{d\pi}\mu_\pi\right)dt + \sigma_D(n)\,dB_t + \left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right)d\hat{M}(n,d), \tag{3.10}$$

$$\frac{dP_t}{P_t} = \mu_D(d)\,dt + \sigma_D(d)\,dB_t + \left(\frac{I^S(\pi,n)}{I^S(\pi,d)} - 1\right)dM(d,n), \tag{3.11}$$

respectively, where for (3.11) we've used the assumption that there is no learning during the disaster phase in our base model. Similar to $(I(\pi_t,n), I(\pi,d))$, we derive the restrictions for $\left(I^S(\pi_t,n), I^S(\pi,d)\right)$ in Appendix A.2 and provide in Appendix B.1 the numerical procedure for solving them.

From (3.10)–(3.11), the aggregate stock returns volatility conditional on the normal phase and the disaster phase follow:

$$volR(\pi,n) = \sqrt{\sigma_D(n)^2 + \lambda_p\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right)^2}. \tag{3.12}$$

$$volR(\pi,d) = \sqrt{\sigma_D(d)^2 + v\left(\frac{I^S(\pi,n)}{I^S(\pi,d)} - 1\right)^2}. \tag{3.13}$$

respectively. Combining (3.10)–(3.11) with the pricing kernel processes (3.4)–(3.4), we obtain the following instantaneous expected premiums for holding the aggregate equity:

$$EP(\pi,n) = \gamma\sigma(n)\sigma_D(n) - \lambda_p\left(\left(\frac{I(\pi_t^+,d)}{I(\pi,n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right), \tag{3.14}$$

$$EP(\pi,d) = \gamma\sigma(d)\sigma_D(d) - v\left(\left(\frac{I(\pi,n)}{I(\pi,d)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi,n)}{I^S(\pi,d)} - 1\right). \tag{3.15}$$

C. Equity index options

At period t, the equilibrium price of a put option written on the aggregate stock is by definition:

$$E_t\left[\frac{\Lambda_{t+\tau}}{\Lambda_t}\max(K-P_{t+\tau},0)\right],$$

where K denotes the the strike price; τ denotes the time to maturity. Starting from (π_t, s_t), we simulate a large number of realizations of terminal option payoffs and use their average as the model-implied option price conditional on (π_t, s_t). More details are given in Appendix B.3 where we also describe the computation of unconditional option prices. Denote by O_t either the conditional (at the given economic phase) or the unconditional price of a put option contract at period t which is characterized by (K, τ). The implied Black-Scholes volatility (B/S-vol) is computed as

$$\text{B/S-vol}_t = BSC^{-1}(\tau, K, O_t, r_{t,t+\tau}, dp_{t,t+\tau}),$$

where BSC^{-1} is the inverse of the Black-Scholes formula for the put option, inverted over the argument σ; $r_{t,t+\tau}$ and $dp_{t,t+\tau}$ are the interest rate and the dividend-price ratio over the period of $[t, t+\tau]$.

IV Model mechanism

Our base model features two components: intertemporal consumption risks due to the alternations of disaster and normal economic phases, and the agent's learning about the disaster intensity. In this section, we describe the mechanism in our base case setup.

A. Impacts of disaster modeled as structural changes

To facilitate the exposition of the model mechanism, we focus on moments conditional on the normal phase. We switch off the learning channel and study the impact of structural changes in consumption. When the disaster rate λ is accurately observed and conditional on $s_t = n$, the short-term rate r and the equity premium EP are given by

$$r^{no}(n) = \beta + \rho\mu(n) - \frac{1}{2}\gamma(1+\rho)\sigma(n)^2 - \lambda\frac{1}{\theta}\left[\left(\frac{I(d)}{I(n)}\right)^{\theta} - 1\right] - \lambda\left(\frac{I(d)}{I(n)}\right)^{\theta-1}\left[1 - \frac{I(d)}{I(n)}\right], \tag{4.1}$$

$$EP^{no}(n) = \gamma\sigma(n)\,\sigma_D(n) - \lambda\left(\left(\frac{I(d)}{I(n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(d)}{I^S(n)} - 1\right), \tag{4.2}$$

respectively, where $\rho = 1/\psi$. Different from the TDMs which feature jumps in consumption level, an economic disaster manifests itself in (4.1)–(4.2) through jumps in the W/C ratio $\frac{I(d)}{I(n)}$. Due to structural changes in the consumption process, W/C ratio jumps while consumption exhibits a continuous path.

When risk aversion and EIS parameters are both greater than one, $I(.)$ responds positively to expected consumption growth but negatively to consumption volatility, hence $\frac{I(d)}{I(n)} < 1$. This result, combined with the implied negative θ defined in (3.2), suggests that $\left(\frac{I(d)}{I(n)}\right)^{\theta}$ and $\left(\frac{I(d)}{I(n)}\right)^{\theta-1}$ are both greater than one. As a result, the last two terms in (4.1) are both negatives and capture the precautionary saving motives against intertemporal consumption risks.

Turning to (4.2), the first and second terms denote the usual compensation for diffusive and jump risks. In the second term, $\left(\frac{I(d)}{I(n)}\right)^{\theta-1} - 1$ denotes the jump size of the pricing kernel (see Eq. (3.4)), which is positive from the above discussion. Like consumption, aggregate dividend also follows a continuous path, and stock valuation declines sharply against dividend when a disaster strikes, which is captured by the negative $\frac{I^S(d)}{I^S(n)} - 1$. Taken together, the agent's marginal utility jumps upward exactly when the stock market crashes, hence the extra compensation in the form of jump-risk premium for holding the aggregate equity.[11]

In sum, modeling disaster as a structural change when combined with recursive utility provides a channel under which the change of economic phase is priced. The model also enables the separation of diffusive from jump risk premium, which is a desirable property for option pricing. However, the extra jump risk premium induced by disaster in this degenerated setup is of much lower magnitude to explain asset prices. In addition, the disaster intensity in this setup is constant and as a consequence contributes negligibly to the explanation of assets volatilities.

[11] It is straightforward to show that the pricing of disasters modeled as structural changes in consumption depends crucially on the choice of recursive utility. Setting $\gamma = \frac{1}{\psi}$, (2.1)–(2.2) degenerate to the usual power utility and θ degenerates to one. Consequently, expressions for r and EP conditional on $s_t = n$ degenerate to $r_t^{no,power}(n) = \beta + \gamma\mu(n) - \frac{1}{2}\gamma(1+\gamma)\sigma(n)^2$ and $EP_t^{no,power}(n) = \gamma\sigma(n)\,\sigma_D(n)$, which are both independent of jumps.

B. Impacts of learning about economic disasters

In this section we aim to decompose the equity premium and provide intuition on the importance of allowing for learning about rare disasters in explaining asset pricing regularities. We first rewrite $r_t(n)$ and $EP_t(n)$ in the base model with learning about disaster as follows:

$$r(\pi,n)=\beta+\rho\mu(n)-\frac{1}{2}\gamma(1+\rho)\sigma(n)^2-\lambda_p\frac{1}{\theta}\left[\left(\frac{I(\pi^+,d)}{I(\pi,n)}\right)^{\theta}-1\right]-\lambda_p\left(\frac{I(\pi^+,d)}{I(\pi_t,n)}\right)^{\theta-1}\left[1-\frac{I(\pi^+,d)}{I(\pi_t,n)}\right], \tag{4.3}$$

$$EP(\pi,n)=\gamma\sigma(n)\sigma_D(n)-\lambda_p\left(\left(\frac{I(\pi_t^+,d)}{I(\pi_t,n)}\right)^{\theta-1}-1\right)\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)}-1\right). \tag{4.4}$$

For a given π, we decompose jumps of the wealth-consumption ratio into two terms:

$$\frac{I(\pi^+,d)}{I(\pi_t,n)}=\frac{I(\pi,d)}{I(\pi,n)}\frac{I(\pi^+,d)}{I(\pi,d)}, \tag{4.5}$$

On the right hand side of (4.5), the first term $\frac{I(\pi,d)}{I(\pi,n)}$ is the counterpart of $\frac{I(d)}{I(n)}$ in the no-learning case, and the second term $\frac{I(\pi^+,d)}{I(\pi,d)}$ is due to learning, where $\pi^+=\pi\frac{\lambda^G}{\lambda_p}<\pi$. Since $I(.,d)$ is increasing in its first argument, as plotted in Figure 1, the second term is also less than one which adds to the downward jump of W/C ratio. By combining (4.5) with (4.3), learning about disasters strengthens the precautionary saving which works towards the resolution of the risk-free rate puzzle.

Using similar decompositions for the jump size of $I^S(.)$, we show that learning implies stock market crashes following the arrival of the disaster phase. On one hand, the agent who cares about intertemporal risks, further trades down aggregate equity relative to that suggested by the worsening economic condition because of her paranoia about future disasters. On the other hand, the extra downward jump in the W/C ratio translates into the extra upward jump in the pricing kernel as captured by

$$\left(\frac{I(\pi_t^+,d)}{I(\pi_t,n)}\right)^{\theta-1}-1=\left(\frac{I(\pi,d)}{I(\pi,n)}\frac{I(\pi^+,d)}{I(\pi,d)}\right)^{\theta-1}-1,$$

which delivers higher compensation for disaster risk. The two effects, which are both attributed to learning about disaster likelihood, work together toward the increase of the jump risk premium component, and thus toward the direction of the resolution of the equity premium puzzle.

In what follows we illustrate the extra jump risk premium due to learning. By rewriting (4.4) as

$$EP(\pi, n) = \gamma\sigma(n)\sigma_D(n) - \lambda_p\left(\left(\frac{I(\pi, d)}{I(\pi, n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi, d)}{I^S(\pi, n)} - 1\right) + EP^L(\pi, n), \quad (4.6)$$

where the second term is the corresponding counterpart from (4.2); the last term EP_t^L is referred to as the learning-induced premium, which is given by

$$EP^L(\pi, n) = \lambda_p\left[\left(\left(\frac{I(\pi, d)}{I(\pi, n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi, d)}{I^S(\pi, n)} - 1\right) - \left(\left(\frac{I(\pi_t^+, d)}{I(\pi_t, n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi_t^+, d)}{I^S(\pi_t, n)} - 1\right)\right]. \quad (4.7)$$

First, learning about disaster risk leads to more dramatic jumps in the stock market, as evidenced by

$$\frac{I^S(\pi_t^+, d)}{I^S(\pi_t, n)} - 1 < \frac{I^S(\pi, d)}{I^S(\pi, n)} - 1 < 0.$$

Second, stock market crashes are viewed as riskier at the presence of learning. This effect is captured by

$$\left(\frac{I(\pi_t^+, d)}{I(\pi_t, n)}\right)^{\theta-1} - 1 > \left(\frac{I(\pi, d)}{I(\pi, n)}\right)^{\theta-1} - 1 > 0.$$

As a result, $EP_t^L(n) > 0$, hence the extra positive portion of equity premium associated with jumps.

As suggested by Liu, Pan, and Wang (2005; LPW), this latter component has a large impact on the explanation of premiums and smirk in the option market relative to the TDMs. LPW add an additional layer of uncertainty aversion towards jumps, along the lines of robust control as in Anderson, Hansen, and Sargent (2000), and document that the implied rare-event premium due to this additional layer of uncertainty is essential to producing a large smirk premium. By resorting to imprecise knowledge about the arrival of disasters, the present paper provides an alternative mechanism that prices rare events separately from diffusions, hence its relevance to option pricing.

V Quantitative results

A. Calibration

Following numerous previous studies, we set $\mu(n) = \sigma(n) = 2\%$.[12] Since the Great Depression represents the prototype of economic disasters in the US, we calibrate $\mu(d)$ and $\sigma(d)$ to annual consumption data from 1929 to 1933 downloaded from Robert Shiller's website. The consumption process is different during the disaster phase in that the consumption growth rate shrinks on average with higher volatility. In particular, we find $\mu(d) = -3.6\%$ and $\sigma(d) = 5\%$.

In contrast to the TDMs, we rely on US experience to calibrate the parameters related to disasters. Disaster arrival rate in our model is inaccurately observed, and we set its long-run average $\bar{\lambda}$ at 1.7%. This level is much lower than the 3.6% used in recent TDMs and based on international evidences (e.g., Barro and Ursua (2008)). We next choose $\nu = 1/3$ implying that the disaster phase lasts on average for three years, which is consistent with the duration of disaster phase during the Great Depression. The switching intensities between the two $\lambda-$regimes are set at $\phi^G = 0.0025$ and $\phi^B = 0.025$ which is consistent with BCDG. With these values, the disaster rate in its stationary region points to the "good" regime with probability $\phi^B / \left(\phi^G + \phi^B\right) = 0.91$. For the base case calibration, we impose

$$\lambda^G = \frac{\bar{\lambda}}{2} \frac{\phi^G + \phi^B}{\phi^B}; \ \lambda^B = \frac{\bar{\lambda}}{2} \frac{\phi^G + \phi^B}{\phi^G}, \tag{5.1}$$

so that the long-run intensity given by (2.6) is indeed $\bar{\lambda}$. Under (5.1) and our choices of $\left(\phi^G, \phi^B\right)$, λ^G and λ^B are 0.00935 and 0.0935, respectively.[13]

We set the subjective time discount rate, β, to 2% which is equivalent to an annual discount factor of 0.98 in discrete time models. We choose the risk aversion parameter $\gamma = 5$, which falls within the range of 1 to 10 deemed reasonable by Mehra and Prescott (1985). For comparison, Barro (2006) considers $\gamma = 3$ and 4; Bansal and Yaron (2004)

[12] Since disasters are rare by nature, the average moment values are close to those during normal times. While the expected consumption growth rate is usually calibrated at around 2%, reports about consumption volatility are more diverse. For example, using quarterly data from 1947–2001, Menzly, Santos, and Veronesi (2005) report a volatility of 1%. Using older data, Campbell (1999) reports a volatility of 3.26%.

[13] We conduct an extensive comparative analysis with respect to the four parameters related to learning: ϕ^G, ϕ^B, λ^G and λ^B. We find that the price implications are largely robust to reasonable deviations from their base case calibrations. Specifically, we consider the impacts on the pricing kernel, the risk free rates, the equity premium, the implied volatilities, and the smirk premium. Details concerning all those results are available upon request.

considers $\gamma = 7.5$ and 10. We set $\psi = 2$ which is consistent with Bansal, Kiku, and Yaron (2007) who estimate ψ to be 2.43 with a standard deviation of 1.3. It is also in line with Vissing-Jorgensen (2002) who estimate ψ to be greater than 1. Finally, to allow for a conservative leverage, we set the leverage parameter $\eta = 3$ lower than the values used in Abel (1990) and Bansal and Yaron (2004).

B. The dynamics of investors' belief updates of disaster intensity

We start by illustrating the updates in belief concerning disaster intensity. The top two panels of Figure 1 plot the absolute jump size of π, $|\pi^+ - \pi|$, as well as the jump size of the filtered intensity, $\lambda_p^+ - \lambda_p$, as a function of π, where π denotes the posterior assessment that λ is in the good regime.

Both jump sizes are zero when $\pi = 0$ or 1 – obviously no belief update takes place if the investors knew the disaster rate for certain. The largest update of π_t and λ_{pt} in absolute value occurs roughly at $\pi_t = 0.77$. This result is different from the implications of models where investors learn about the drift in economic fundamentals (e.g., Veronesi (1999); Ai (2010)). In those models, investors continously update their assessment about π_t, which follows a diffusive process and takes the greatest variance when $\pi_t = 0.5$ – the state the agent is the most uncertain about the parameter regime.

By contrast, the jump in π_t in our setup is determined not only by uncertainty but also by λ_p. More explicitly, there are two components that interact to determine the magnitude of belief updates. The first is the uncertainty regarding the true disaster intensity and it reaches its maximum at $\pi = 0.5$. The second component, unique to learning about a counting process, is investors' assessment of the disaster rate λ_p whose expression is given by (2.9). We name this component *the element of surprise*. Intuitively, the strike of a disaster would be of a great surprise to investors when they have almost forgot about its possibility, i.e., when λ_p an instant before disaster is near λ^G. Note that at this state the uncertainty is low since the implied π is close to one. Next, consider the maximum uncertainty state at $\pi = 0.5$. This state corresponds to λ_p half way between the good regime $\left(\lambda^G\right)$ and the bad regime $\left(\lambda^B\right)$. If a disaster happens at that moment, the element of surprise is much lower than in the case where λ_p is close to λ^G.

The element of surprise is strictly increasing in π (since λ_p is decreasing in π), while the level of uncertainty is hump-shaped and reaches its maximum at a $\pi = 0.5$. With the increase of π starting from $\pi = 0$, both elements become stronger until π reaches 0.5.

With π higher than 0.5, the element of surprise continues to rise while at the same time the uncertainty decreases. As a result, the largest belief update takes place at π between 0.5 and 1. This novel disaster learning mechanism has an influence on the state dependences of both the equity premium and the volatility of asset returns as we will discuss later.

C. Wealth-consumption and price-dividend ratios

We next turn to the weath-consumption ratio, (W/C) and the price-dividend ratio, (P/D). To save space, we focus on their values conditional on the normal economic phase, and the two bottom panels of Figure 1 plot $I(\pi, n)$ and $I^S(\pi, n)$ as a function of π.

To understand their monotonicity, first notice that a positive shock on λ_p implies a higher disaster intensity which tends to substantially depress consumption. On one hand, the income effect makes the agent consume less today which raises the W/C ratio. On the other hand, the intertemporal substitution effect encourages the agent to borrow from the future which depresses the W/C. The substitution effect dominates if $\psi > 1$, hence, the negative (or positive) relation between W/C ratio and λ_p (or π). Since the leverage parameter η is constant in our economy, a shock to λ_p that depresses future consumptions would also depress future dividends. Using a similar logic for the price dividend ratio, the implied P/D ratio is also increasing in π when $\psi > 1$.

In addition to the monotonicity, W/C ratio is convex in π which is attributed to learning about rare disasters. First, the longer the normal economic phase persists, the more optimistic investors become concerning future disaster likelihood. As a result, the filtered π drifts upward which drives up the aggregate wealth. Second, consider the two scenarios: i) π moves up toward 0.5 and ii) π moves away from 0.5 toward 1. In the first scenario the impact of investors' optimism is dampened by the larger uncertainty, while in the second scenario investors' growing optimism and less uncertainty compound to drive up the W/C ratio.

Similarly, P/D ratio is both increasing and convex in π. In particular, the P/D ratio becomes more convex in comparaison to the W/C ratio at higher (better) π states. Intuitively, when investors are very optimistic (i.e., when π is close to one) and since dividend is the levered consumption, prices over-react faster with respect to the current dividend than it is for wealth with respect to consumption.

The above mechanism is similar only in spirit to that identified in Veronesi (1999). Veronesi emphasizes overreaction to bad news in the good π states. In contrast, investors in our model, during the normal phase and good π states, overreact after an extended period of no news (silence is informative) so that the price rises in a convex way. Also different from Veronesi (1999), investors react to the actual news of disaster by a sharp (discrete) update of their posterior π. As discussed above, this reaction takes its maximal value at π between 0.5 and 1.

Using P/D ratio as an example to illustrate the implied moment values, we report in Panel B of Table I both the P/D ratios computed at the long-run average of $\pi_t = \bar{\pi}$ (Column 2) and the P/D ratios averaged over the stationary distribution of π (Column 6). Consistent with the data, the P/D ratio is procyclical – $P/D\,(\bar{\pi}, n)$=83.4 is much higher than $P/D\,(\bar{\pi}, d)$=47.2. Furthermore, the average P/D is higher than $P/D\,(\bar{\pi}, .)$ reflecting that the P/D ratio is convex in π.

In comparison with the case of no learning about disasters (reported in Column 4, Panel B of the same table), the implied P/D is nearly an order of magnitude lower. The belief update upon a disaster strike results in a larger stock market jump as well as a higher compensation for jump risk. While the implied P/D ratios from the base model still look high compared with their empirical counterparts, we show in Section VI that the they can be further reduced in our extended setup where investors also learn about the economic recovery rate.

D. Short term rate and equity premium

In Panel B of Table I, Columns 3–4 report model implications about the short term interest rate r and the equity premium EP computed at $\pi_t = \bar{\pi}$. For comparison, we report in Panel A of the same table the implied r and EP from a reduced model without learning, where the accurately observed λ is set equal to $\bar{\lambda}$, the long-run average of the filtered disaster intensity λ_p evaluated at $\bar{\pi}$. Consistent with the data, r is procyclical while EP is countercyclical. Conditional on the normal phase, $r\,(\bar{\pi}, n)$ in our model is 1.66% which is much lower than that implied by the reduced (no-leaning) model. On the other hand, $EP\,(\bar{\pi}, n)$ is 6.99% which is more than triple that of its no-learning counterpart. As explained earlier, the agent's learning behavior strengthens the precautionary saving motive and generates extra compensation for jump risks, hence the low $r\,(.)$ and the high $EP\,(.)$.

To understand the different moment values conditional on the disaster phase, we use

$EP(:,d)$ as an example and write down its expressions without and with learning as follow:

$$EP^{no}(d) = \gamma\sigma(d)\,\sigma_D(d) - \nu\left(\left(\frac{I^{no}(n)}{I^{no}(d)}\right)^{\theta-1} - 1\right)\left(\frac{I^{no,S}(n)}{I^{no,S}(d)} - 1\right), \tag{5.2}$$

$$EP(\bar{\pi},d) = \gamma\sigma(d)\,\sigma_D(d) - \nu\left(\left(\frac{I(\pi_t,n)}{I(\pi_t,d)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi_t,n)}{I^S(\pi_t,d)} - 1\right). \tag{5.3}$$

When risk aversion and EIS parameters are both greater than one, the pricing kernel and stock price jump in opposite directions since $\theta < 0$. As a result, we obtain positive jump-risk premiums in both cases. By comparison, the implied premiums are much higher during disaster times than during normal times, which is attributed to the high recovery rate ν arising from the short-lived disaster phase. A similar analysis explains the implied negative $r(.,d)$. In the data, a negative real rate tends to emerge when the central bank cuts the nominal rate to near zero to fight against recessions, the extreme form of which is the disaster phase.

As reported in Table I, the implied $EP^{no}(d)$ without learning is higher than $EP(\bar{\pi},d)$. This result seems counter-intuitive at the first sight, and it is straightforward to understand considering the forward-looking evolution of disaster intensity in the two different cases. In both the learning and the no-learning case, equity premium is computed when the (posterior) disaster intensity λ_p equals its long run average $\bar{\lambda}$ in both cases. The no-learning case features a constant disaster intensity at $\bar{\lambda}$. In contrast, λ_p starts to drift downward from $\bar{\lambda}$ as soon as the economy recovers. Given that normal economic phase lasts much longer than economic disasters, the compensation for jump risk is determined by an average intensity lower than the initial value of $\bar{\lambda}$. In the extended model described later, investors also learn about recovery rates during disaster times. We show that the extra pessimism of investors in response to the persistence of disaster phase would outweigh the growing optimism after recovery. As a result, the extended model generates a higher $EP(\bar{\pi},d)$ than that implied from the no-learning case for any initial value.

Results reported so far are based on moment values evaluated at the long-run average of the state. Now, we examine the implied r and EP as a function of the state π which is plotted in Figure 2. To facilitate exposition, we focus on their values conditional on the normal phase. Another reason for this choice is that disaster time is more interesting to consider in the extended model in which investors learn also about the intensity of the economic recovery.

Starting from $\pi = 0$, the top left panel shows that $EP(\pi, n)$ rises first and then drops with its maximum obtained at $\pi = 0.82$. This pattern roughly replicates that of $|\pi^+ - \pi|$, the absolute jump size of the state, which is plotted in Figure 1. We further plot the decompositions of EP according to (4.6) in the top right panel of the same figure, where the solid line plots the learning-induced premium $EP_t^L(\pi, n)$ defined by (4.7); the dotted line plots the residual component, i.e., $EP^R(\pi, n) \equiv EP(\pi, n) - EP^L(\pi, n)$. $EP^R(\pi, n)$ is decreasing in π, or equivalently, increasing in the posterior estimation of the disaster rate λ_p. Intuitively, a higher λ_p implies a higher jump-risk premium demanded to hold the aggregate equity when we ignore learning. Unlike $EP^R(\pi, n)$, $EP^L(\pi, n)$ is hump-shaped: it equals zero when the agent has complete information about the two $\lambda-$regimes in which no belief update occurs; and it takes its maximum when π is between 0.5 and 1 at which the agent updates her belief the most. The qualitative difference between the two components and their comparable magnitudes highlights the impacts of learning on the pricing of the aggregate equity.

Turning to the short-term rate which is plotted in the bottom left panel of the same figure, $r(\pi, n)$ is initially decreasing and then becomes increasing in π with its minimum obtained at π between 0.5 and 1. This shape is driven by the learning-induced precautionary saving. While $EP(\pi, n)$ is concave in π, $r(\pi, n)$ is a convex function of π. As a result, taking averages over the stationary distribution region of π yields lower $EP(n)$ and higher $r(n)$ as compared to their values computed at $\pi = \bar{\pi}$. These results are reported in Column 7–8 in Panel B of Table I.

E. Stock return volatility and "leverage" effect

Conditional on the normal phase, the following two equations give the stock returns volatility implied from our base case model and from the reduced model without learning, respectively:

$$volR(\pi, n) = \sqrt{\sigma_D(n)^2 + \lambda_p \left(\frac{I^S(\pi_t^+, d)}{I^S(\pi_t, n)} - 1 \right)^2}, \tag{5.4}$$

$$volR^{no}(n) = \sqrt{\sigma_D(n)^2 + \lambda \left(\frac{I^{S,no}(d)}{I^{S,no}(n)} - 1 \right)^2}. \tag{5.5}$$

Evaluated at $\lambda_p = \lambda = \bar{\lambda}$, return volatility implied from (5.4) is higher than $volR^{no}(n)$ which is due to the higher jump size in $I^S(.)$ at the presence of learning (Section IV.B).

Stock returns become even more volatile at higher λ_p implying that investors are more concerned about the potential future disasters. As argued in Wachter (2011) and Gabaix (2011), time-varying disaster intensity is a desirable property in matching the stock return volatility. Our model endogenously generates this property through learning.

We next examine the state dependence of $volR(\pi, n)$ which is plotted in the top left panel of Figure 3. To understand the underlying economics, we plot together in the top right panel of the same figure the state dependence of the disaster intensity λ_p and the squared jump size $\left(\frac{I^S(\pi_t^+, d)}{I^S(\pi_t, n)} - 1\right)^2$. First, λ_p is decreasing in π meaning less concern about the potential disasters at the good π regime. Second, jump size is hump-shaped taking its maximum value at a π between 0.5 and 1 which mimicks that of $|\pi^+ - \pi|$. The two impacts largely cancel out each other at low π states, and work in the same direction at high π states. As a result, the implied $volR(\pi, n)$ first exhibits no clear monotonicity and then becomes decreasing in π.

Taking together that $volR(\pi, n)$ is mostly decreasing in π while $I^S(\pi, n)$ is increasing in π (bottom right panel of Figure 1), our model provides a theoretical support for the empirically observed "leverage" effect, i.e., low price level tends to be associated with high volatility. The "leverage" effect is also discussed in Veronesi (1999) where investors learn about the unknown drift of the aggregate dividend, and the implied return volatility is hump shaped in π. In particular, volatility in his model significantly depresses as π drops to 0.5 from one, resulting in a reduction in uncertainty. In our setup, the reduction in uncertainty is largely offset by the associated higher disaster intensity (the element of surprise). Also different from Veronesi (1999) which implicitly considers one economic phase, return volatility in our model is countercyclical – higher during the disaster phase than during the normal phase (see the last column in Panel B of Table I)

F. Premiums implicit in equity index options

Unlike many other financial variables, time series of option data are relatively short starting from 1980s, and the empirically documented volatility smirks are typically based on observations during normal times. Given the dynamic of consumption and GDP documented Barro and Ursua (2008), the US economy did not experience any disaster since the Great Depression. To facilitate comparison with the data, we focus then on option prices conditional on the normal phase in the following analysis.

Panel A of Table II reports option pricing implications from the reduced model without learning. For comparison, we also report the return volatility in the same panel. The implied volatility of ATM options (7.48%) is lower than the return volatility (8.97%) which is counterfactual: empirically ATM options are priced with a premium with respect to the equity volatility. On the other hand, implied volatility from the 10% OTM put options (18.6%) is 10.4% higher than that from ATMs implying a pronounced volatility smirk in the cross sectional plot of option-implied volatilities against moneyness. The reason is well understood in the literature: as one moves from ATMs to deep OTMs, we are looking at a sequence of assets that are increasingly sensitive to jumps, hence the higher compensation for the jump risk in the form of a higher implied volatility.

Turning to our base model, Panel B of Table II reports implications evaluated at the long-run average of the state, i.e., $\pi = \bar{\pi}$. First, the implied ATM volatility is 12.8% representing a 2% premium relative to the return volatility of 10.6%, which is consistent with the data (e.g., Du (2010)). Second, the smirk premium, measured as the difference between the 10% OTM and the ATM volatility, is 13.8% which is higher than that implied by the reduced model.[14] This result highlights the mechanism that learning-induced compensation drives up the jump-risk premium as a fraction of the total premium. Since premium implicit in option pricing largely reflects the jump risk (e.g, Pan, (2002)), our model generates a more pronounced volatility smirk attributed to investors' learning behavior than that suggested by the no-learning case.

We next examine the state dependence of option implied premiums. The bottom left panel plots the difference between ATM volatility and $volR$ referred to as ATM premium; the bottom right panel plots the smirk premium. The state dependence of both premiums resembles that of $EP(\pi, n)$ plotted in Figure 2 which is attributed to learning about disaster intensity. Specifically, upon the strike of a disaster, investors update their belief the most when π is between 0.5 and 1 at which uncertainty and the element of surprise achieves the best tradeoff (Section V.B). Consequently, the model generates the highest jump-risk premium leading to the highest option implied premiums.

Finally, Panel C of Table II reports the unconditional moments associated with option pricing, computed as averages of moment values over the stationary region of π during the normal phase. Since the ATM premium and the smirk premium in our model are both concave functions of π, as plotted in Figure 3, their unconditional levels drop relative to their

[14] While the implied 13.8% premium at $\bar{\pi}$ looks high, we show toward the end of this subsection that the unconditional smirk premium, ATM premium and 10% OTM volatility are all close to their empirical levels.

values conditional on $\pi = \bar{\pi}$. In particular, the implied unconditional ATM premium, smirk premium, and 10% OTM volatility are 1.8%, 12.5%, and 24.3%, which are all comparable to their empirical levels at 2.4%, 10%, and 25%, respectively.

VI The general model

A. Learning about recovery

In this subsection, we consider the general model where the recovery rate ν is also inaccurately observed. Again, we make the simplifying assumption that ν only takes two possible values: ν^G and ν^B. Setting $\nu^G > \nu^B$, we interpret the two possible values as the good and the bad regime for the economic recovery. ν switches between the two possible regimes according to the following Markov chain:

$$\Pr\left(\nu_{t+dt} = \nu^B | \nu_t = \nu^G\right) = \psi_G \tag{6.1}$$

$$\Pr\left(\nu_{t+dt} = \nu^G | \nu_t = \nu^B\right) = \psi_B \tag{6.2}$$

where ψ_G and ψ_B are known parameters.

Denote by π_t^ν the investors' posterior belief that ν is in the good regime at period t, i.e.,

$$\pi_t^\nu \equiv \Pr\left(\nu_t = \nu^G | \mathcal{F}_t\right).$$

Conditional on the disaster phase and by Theorem 19.6 of Lipster and Shiryaev (2001), π_t^ν evolves according to:

$$d\pi_t^\nu = \mu_\pi^\nu dt + \pi_t^\nu \frac{\nu^G - \nu_{pt}}{\nu_{pt}} d\hat{M}_t\left(d, n\right), \tag{6.3}$$

where

$$\mu_\pi^\nu \equiv \pi^\nu\left(\nu_{pt} - \nu^G\right) + \left(1 - \pi^\nu\right)\psi_B - \pi^\nu \psi_G; \tag{6.4}$$

$d\hat{M}_t\left(d, n\right)$ is the Poisson process governing the start of recovery under the filtered measure with intensity ν_{pt};

$$\nu_{pt} \equiv \pi_t^\nu \nu^G + \left(1 - \pi_t^\nu\right)\nu^B = \nu^B + \pi_t^\nu\left(\nu^G - \nu^B\right) \tag{6.5}$$

denoting the posterior estimation of ν which is increasing in π_t^ν. Either π_t^ν or ν_{pt} can serve as an additional state in the general model.

To understand the agent's learning about recovery, we start from the long-run average

$$\pi_t^\nu = \bar{\pi}^\nu = \frac{\psi_B}{\psi_G + \psi_B}, \tag{6.6}$$

under which μ_π^ν in (6.4) reduces to $\pi_t^\nu \left(\nu_p - \nu^G\right) < 0$. If no economic recovery starts over the interval $[t, t + dt]$, the agent becomes less certain about recovery, and thus revises downward the likelihood of the good $\nu-$regime. By contrast, if an economic recovery takes place within $[t, t + dt]$, the agent substantially revises upward the good $\nu-$regime by

$$\pi^{\nu,+} - \pi^\nu = \pi^\nu \frac{\nu^G - \nu_p}{\nu_p} = \frac{\left(\nu^G - \nu^B\right)}{\nu_p} \pi^\nu \left(1 - \pi^\nu\right) > 0, \tag{6.7}$$

The associated upward jump in ν_{pt} is

$$\nu_{pt}^+ - \nu_{pt} = \left(\pi^{\nu,+} - \pi^\nu\right)\left(\nu^G - \nu^B\right) > 0.$$

Intuitively, upon the current recovery, the agent becomes confident that the economy would likely recover following the strike of future disasters.

The top left panel of Figure 4 illustrates the jump size $\pi^{\nu,+} - \pi^\nu$ as a function of π^ν, where π^ν denotes the posterior assessment that recovery rate is in good regime. Different from learning about the disaster rate, $\pi^{\nu,+} - \pi^\nu$ take its maximum when π^ν is around 0.35. To see the reason, first note that uncertainty about the two $\nu-$regimes is hump-shaped in π^ν and takes the maximum at $\pi^\nu = 0.5$. Second, the implied jump size is also determined by ν_p referred to as the element of surprise. Intuitively, recovery would be least surprising when investors are optimistic, i.e., when υ_p is near $\upsilon^G \left(> \upsilon^B\right)$. Taken together, both uncertainty and surprise become stronger with the decrease of π^ν starting from $\pi^\nu = 1$ until π^ν reaches 0.5. For π^ν lower than 0.5, the element of surprise continues to rise while uncertainty decreases at the same time. As a result, the largest belief update takes place at π^ν between 0 and 0.5.

In contrast to the base model, the mechanism of learning in the general model is alive during both the disaster and the normal economic phases. On one hand and during the disaster phase, investors form a posterior belief about the arrival of recovery according to (6.3)–(6.4) and stop updating their beliefs about the strike of disasters. The longer the disaster lasts, the more pessimistic investors become so that they negatively update the filtered recovery rate. However, upon the start of recovery, investors gain more confidence about the ability of the economy to bounce back, and as a result they substantially revise upward the intensity of economic recovery. On the other hand and during the normal phase, investors form a posterior estimation of disaster likelihood according to (2.7)–(2.8), and stop updating their beliefs about future recoveries. In this case, the longer the good economic condition persists, the more optimistic investors become, resulting in a negative update of the filtered disaster intensity. This negative trend persists until the actual strike of an economic disaster, at which investors become paranoid and substantially revise upward the probabillty of future disasters.

B. Model implications

Using a similar procedure as the ones used in the base model, we numerically solve the general model and the details are in Appendix C. To generate quantitative implications, we set the long-run average of the recovery rate $\bar{\nu}$ equals to 1/3 which is consistent with its base level. Next, we set $\psi^G = 0.025$ and $\psi^B = \psi^G/3$ which implies that the recovery rate in its stationary region points to the "good" regime with probability $\psi^B / \left(\psi^G + \psi^B\right) = 0.25$. Next, we impose

$$\nu^G = \frac{\bar{\nu}}{2} \frac{\psi^G + \psi^B}{\psi^B}; \ \nu^B = \frac{\bar{\nu}}{2} \frac{\psi^G + \psi^B}{\psi^G} \tag{6.8}$$

so that the long-run intensity is indeed $\bar{\nu}$. Given our calibrations of $\left(\psi^G, \psi^B\right)$, ν^G and ν^B are 0.66 and 0.22, respectively. Thus, conditional on the disaster phase, the recovery takes 1.5 years on average when $\nu = \nu^G$, and takes 4.5 years on average when $\nu = \nu^B$. Empirically, the Great Depression lasts from 1929–1933 which is close to the duration indicated by ν^B. If we treat the most recent economic recession identified by NBER as another disaster phase, then its duration fits that suggested by ν^G.[15] Finally, all other parameters remain at their base case levels as described in Section V.A.

Conditional on phase $s \in \{n, d\}$, we denote by $I^S(\pi, \pi^v, s)$ the price-dividend ratio. π and π^v are the two states characterizing uncertainty about both disaster and recovery rates, respectively. Given our model calibration, the top right panel of Figure 4 (solid line) plots $I^S(\pi, \pi^v, d)$ as a function of π^v, where π is fixed at its long run average $\bar{\pi}$. First, the longer the disaster phase persists, the more pessimistic the investors become concerning

[15] As in the base model, we conduct extensive comparative analysis with respect to the four parameters related to learning about recovery: ν^G, ν^B, ψ^G and ψ^B. We find model implications are robust to reasonable deviations from their reported calibrations.

the likelihood of recovery. As a result, the filtered π^v drifts downward which is associated with a decreasing price. Second, consider the two scenarios: i) π^v drops toward 0.5 and ii) π^v drops from 0.5 to 0. In the first scenario, investors' pessimism compounds with the increased uncertainty to drive down the P/D ratio, while in the second scenario the impact of investors' pessimism is dampened by less uncertainty. As a result, $I^S(\bar{\pi}, \pi^v, d)$ is increasing and convex in π^v during the disaster phase.

The above mechanism is novel in the literature – investors update their beliefs about economic recovery in the absence of any news. In addition, this mechanism induces both a dramatic price drop at the start of bad times when π^v is relatively high, and moderate drops at later stage of the disaster time when π^v becomes relatively low. This result is consistent with the empirical observations that stock market crashes at the start of bear markets, and becomes relatively stable toward the end of the bear market.

Column 2–5 in Panel A of Table III report model implications about the P/D ratio, short term interest rate r, equity premium EP, and return volatility $volR$ computed at $(\pi_t, \pi_t^v) = (\bar{\pi}, \bar{\pi}^v)$. By allowing investors to learn about recovery rate, the growing pessimism during the disaster phase compounds the impact of getting paranoid upon the strike of disasters. As a result, compared to the base model, the P/D ratio depresses further, investors demand higher compensation to hold the aggregate equity, and short term rate further decreases.

Empirically, the S&P500 index reaches its lowest level in June 1932 towards the end of the Great Depression, and it bounces back by more than three times over the next four years producing an average return of 32.5% on an annual basis. If we use subsequent realized returns to proxy for the expected returns, the observation lends support to the 37.4% expected equity premium during the disaster phase that is implied from the general model (Table III).

Columns 6–9 in the same panel report the corresponding unconditional moments (the unconditional moment refers to an average over the stationary distribution of π and π^v). In the last row of the table we take the weighted average with respect to the economic phases. The implied r, EP, and $volR$ are 1.49%, 7.73%, and 13.1%, respectively. These numbers are reasonable matches to their empirical levels except for the P/D ratio. The implied P/D ratio (51.3) is still high relative to its long-term norm (25–35), but it matches the observed levels since the mid 1990s (e.g., Lettau, Ludigson, and Wachter, (2007)).

Turning to the state dependences, the bottom left two panels of Figure 4 plot EP and r as a function of π^v during the disaster phase, where π is fixed at its long run average. From the bottom left panel, $EP(\bar{\pi}, \pi^v, d)$ rises first and then drops with the increase of π^v, which reflects the hump shape of $\pi^{v,+} - \pi^{v}$. EP reaches its maximum at π^{v} higher than 0.5. To understand the state dependence, we plot together in the top right panel of the same figure the model-implied $I^{S}(\bar{\pi}, \pi^{v}, d)$ (solid line) and $I^{S}(\bar{\pi}, \pi^{v,+}, d)$ (dashed line). While $\pi^{v,+} - \pi^{v}$ takes the maximum at around $\pi^{v} = 0.35$, the convexity of $I^{S}(.)$ as the function of π^{v} implies that $I^{S}(\bar{\pi}, \pi^{v,+}, d) - I^{S}(\bar{\pi}, \pi^{v}, d)$ takes the maximum at $\pi^{v} > 0.5$. The shape of $EP(\bar{\pi}, \pi^{v}, d)$ mimicks that of the $I^{S}(.)$, the stock price jump size which controls the quantity of the jump risk. The inverted pattern is found in $r(\bar{\pi}, \pi^{v}, d)$: this time learning creates the stronger precautionary saving motives which drives down the short term rate.

As is the case in the base model, we report in Panel B & C of Table III the option implied volatilities conditional on the normal economic phase. For comparison, we also report the corresponding return volatilities. Compared to that implied from the base model, learning about recovery during disaster time does not change much the implied volatilities during the normal time.[16] The implied ATM premium, smirk premium, and 10% OTM volatility are (4%, 15.1%, 28.1%) when evaluated at $(\pi_{t}, \pi_{t}^{v}) = (\bar{\pi}, \bar{\pi}^{v})$ (Panel B of Table III), and (2.9%, 12.4%, 25.2%) after taking the average with respect to the state distributions (Panel C of Table III). Specifically, the unconditional values deliver reasonable matches to the their empirical levels at 2.4%, 10%, and 25%, respectively.

VII Conclusions

This article proposes a model within the paradigm of Rietz-Barro hypothesis that is able to explain various asset pricing regularities in the equity, bond and options markets. Investors learn about disaster probabilities in the base model and also learn about the recovery rate in the general model. In the absence of information concerning economic disasters, investors decrease their posterior assessment of disaster probabilities. The longer good economic conditions persist, the larger the increase in asset prices. However, upon the strike of a disaster, two compounding effects take place. First, investors become paranoid about future disaster likelihood and instantaneously revise upward their posterior beliefs about the disaster intensity. Second, the longer the disaster period persists the more pessimistic investors become concerning the potential economic recovery. As a result, a large and persistent price correction takes place as well as an increase in volatilities during this phase.

[16] From Panel A of Table III, we see that learning about recovery significantly increases the return volatility compared to the implication from the base model during the disaster phase.

Our model features no jumps in the level of consumption. The model calibration is solely based on US experience. It generates time-varying disaster and recovery intensities which are useful to explain asset volatilities and "leverage" effects. Our framework differentiates the pricing of jump risks from that of diffusive risks which turns out to have a large impact in explaining option market regularities. Finally, we find that a recursive utility when combined with learning about rare disasters is able to explain asset pricing puzzles with reasonably low consumption moments autocorrelation.

This paper focuses on particular types of events: the strike of disasters and subsequent recoveries. By changing the calibration of the consumption process and the switching intensity, our framework can analyse the impact of investors learning about business cycle on financial markets. Another likely and more interesting direction to pursue is to allow investors to learn about the likelihood of economic disasters and recoveries in the economy from the observed equity market crashes. Doing so can shed light on the importance of the feedback effect from financial markets to aggregate economic variables.

Appendix

A. Derivations for the base model

A.1. Aggregate wealth and short term rates

This subsection derives the model-implied restrictions for the W/C ratio, $I(\pi, n)$, $I(\pi, d)$, and $I(\pi, r)$ which are conditional on the normal and the disaster phase, respectively. By Ito's lemma, the value function J_t given by (3.1) follows:

$$\frac{dJ_t}{J_t} = (1-\gamma)\mu(n) - \frac{1}{2}\gamma(1-\gamma)\sigma(n)^2 + \theta\frac{1}{I(\pi,n)}\frac{dI(\pi,n)}{d\pi}\mu_\pi + \left[\left(\frac{I(\pi^+,d)}{I(\pi,n)}\right)^\theta - 1\right] d\hat{M}_t(n,d) \tag{A.1}$$

$$\frac{dJ_t}{J_t} = (1-\gamma)\mu(d) - \frac{1}{2}\gamma(1-\gamma)\sigma(d)^2 + \left[\left(\frac{I(\pi,r)}{I(\pi,d)}\right)^\theta - 1\right] dN_t(d,r), \tag{A.2}$$

$$\frac{dJ_t}{J_t} = (1-\gamma)\mu(r) - \frac{1}{2}\gamma(1-\gamma)\sigma(r)^2 + \left[\left(\frac{I(\pi,n)}{I(\pi,r)}\right)^\theta - 1\right] dN_t(r,n), \tag{A.3}$$

conditional on $s_t = n$, $s_t = d$, and $s_t = r$, respectively, where μ_π is the drift of $d\pi$ defined by (2.8); we've used that the agent stops updating π_t conditional on the disaster phase and the recovery phase. On the other hand, the differential form for recursive utility (defined by (2.1)) is:

$$dJ_t = -f\left(C_t, J_t\right)dt + dH_t, \tag{A.4}$$

where H_t is some martingale. By combining the aggregator with the value function, we can rewrite $f\left(.\right)$ as

$$f\left(C_t, J_t\right) = \theta \frac{J}{I\left(\pi, s\right)} - \beta\theta J. \tag{A.5}$$

Taking expectation on both sides of (A.4),

$$E_t\left(dJ_t\right) + f\left(C_t, J_t\right) = 0, \tag{A.6}$$

where $E_t\left(.\right)$ is with respect to the filtered measure. In (A.6), substituting for $E_t\left(dJ_t\right)$ from (A.1)–(A.2) and substituting for $f\left(C_t, J_t\right)$ from (A.5), we obtain

$$\frac{1}{I\left(\pi,n\right)} = \beta + \left(\rho-1\right)\mu\left(n\right) + \frac{1}{2}\gamma\left(1-\rho\right)\sigma\left(n\right)^2 - \frac{1}{I\left(\pi,n\right)}\frac{dI\left(\pi,n\right)}{\partial\pi}\mu_\pi - \lambda_p\frac{1}{\theta}\left[\left(\frac{I\left(\pi^+,d\right)}{I\left(\pi,n\right)}\right)^\theta - 1\right] \tag{A.7}$$

$$\frac{1}{I\left(\pi,d\right)} = \beta + \left(\rho-1\right)\mu\left(d\right) + \frac{1}{2}\gamma\left(1-\gamma\right)\sigma\left(d\right)^2 - v\frac{1}{\theta}\left[\left(\frac{I\left(\pi,r\right)}{I\left(\pi,d\right)}\right)^\theta - 1\right]. \tag{A.8}$$

$$\frac{1}{I\left(\pi,r\right)} = \beta + \left(\rho-1\right)\mu\left(r\right) + \frac{1}{2}\gamma\left(1-\gamma\right)\sigma\left(r\right)^2 - \xi\frac{1}{\theta}\left[\left(\frac{I\left(\pi,n\right)}{I\left(\pi,r\right)}\right)^\theta - 1\right]. \tag{A.9}$$

for the normal phase, the disaster phase, and the recovery phase, respectively.

By applying Ito's lemma with jumps to (3.3) conditional on each of the two economic phases, we obtain the expression for the short-term interest rate:

$$r_t\left(\pi,n\right) = \beta\theta + \frac{1-\theta}{I\left(\pi,n\right)} + \gamma\mu - \frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(n\right)^2 - \left(\theta-1\right)\frac{1}{I\left(\pi,n\right)}\frac{dI\left(\pi,n\right)}{d\pi}\mu_\pi - \lambda_p\left[\left(\frac{I\left(\pi^+,d\right)}{I\left(\pi,n\right)}\right)^{\theta-1} - 1\right], \tag{A.10}$$

$$r_t\left(\pi,d\right) = \beta\theta + \frac{1-\theta}{I\left(\pi,d\right)} + \gamma\mu - \frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(d\right)^2 - \nu\left[\left(\frac{I\left(\pi,r\right)}{I\left(\pi,d\right)}\right)^{\theta-1} - 1\right], \tag{A.11}$$

$$r_t\left(\pi,r\right) = \beta\theta + \frac{1-\theta}{I\left(\pi,r\right)} + \gamma\mu - \frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(r\right)^2 - \xi\left[\left(\frac{I\left(\pi,n\right)}{I\left(\pi,r\right)}\right)^{\theta-1} - 1\right], \tag{A.12}$$

where we've used that π stops evolving conditional on the disaster and the recovery phase.

In (A.10) and conditional on the normal phase, substituting for $\frac{1}{I}$ from (A.7) yields:

$$r\left(\pi_t,n\right)=\beta+\rho\mu\left(n\right)-\frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(n\right)^2-\lambda_p\frac{1}{\theta}\left[\left(\frac{I\left(\pi_t^+,d\right)}{I\left(\pi_t,n\right)}\right)^{\theta}-1\right]-\lambda_{pt}\left(\frac{I\left(\pi_t^+,d\right)}{I\left(\pi_t,n\right)}\right)^{\theta-1}\left[1-\frac{I\left(\pi_t^+,d\right)}{I\left(\pi_t,n\right)}\right], \tag{A.13}$$

where we've used the definition of θ given by (3.2. By a similar procedure, we obtain the short term rate conditional on the disaster and the recovery phase:

$$r\left(\pi_t,d\right)=\beta+\rho\mu\left(d\right)-\frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(d\right)^2-v\frac{1}{\theta}\left[\left(\frac{I\left(\pi_t,r\right)}{I\left(\pi_t,d\right)}\right)^{\theta}-1\right]-v\left(\frac{I\left(\pi_t,r\right)}{I\left(\pi_t,d\right)}\right)^{\theta-1}\left[1-\frac{I\left(\pi_t,r\right)}{I\left(\pi_t,d\right)}\right]. \tag{A.14}$$

$$r\left(\pi_t,r\right)=\beta+\rho\mu\left(r\right)-\frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(r\right)^2-v\frac{1}{\theta}\left[\left(\frac{I\left(\pi_t,n\right)}{I\left(\pi_t,r\right)}\right)^{\theta}-1\right]-\xi\left(\frac{I\left(\pi_t,n\right)}{I\left(\pi_t,r\right)}\right)^{\theta-1}\left[1-\frac{I\left(\pi_t,n\right)}{I\left(\pi_t,r\right)}\right]. \tag{A.15}$$

A.2. Aggregate stock

This subsection derives the model-implied restrictions for the P/D ratio, $I^S\left(\pi_t,n\right)$ and $I^S\left(\pi_t,d\right)$, which is conditional on the normal and the disaster phase, respectively. The point is to pursue the fundamental asset pricing relationship (e.g., chapter 1 of Cochrane, 2005), i.e.,

$$EP_t=E_t\left(\frac{dS_t}{S_t}\right)/dt+\frac{D_t}{S_t}-r_t, \tag{A.16}$$

where r_t is the short-term rate given by (A.13)–(A.14); EP_t is the equity premium given by (3.14)–(3.15); $\frac{D_t}{S_t}=\frac{1}{I^S(\pi_t,s_t)}$ which denotes the dividend yield conditional on $s_t\in\{n,d\}$.

From the stock return processes (3.10)–(3.11),

$$E_t\left(\frac{dS_t}{S_t}|s_t=n\right)=\mu_D\left(n\right)+\frac{1}{I^S\left(\pi_t,n\right)}\frac{dI^S\left(\pi_t,n\right)}{d\pi}\mu_\pi+\lambda_p\left(\frac{I^S\left(\pi_t^+,d\right)}{I^S\left(\pi_t,n\right)}-1\right), \tag{A.17}$$

$$E_t\left(\frac{dS_t}{S_t}|s_t=d\right)=\mu_D\left(d\right)+\nu\left(\frac{I^S\left(\pi,n\right)}{I^S\left(\pi,d\right)}-1\right). \tag{A.18}$$

In (A.16) and conditional on $s_t=n$, substituting for EP and $E_t\left(\frac{dS_t}{S_t}\right)$ from (3.14) and

(A.17) yields

$$\begin{aligned} & r_t(n) + \gamma\sigma(n)\sigma_D(n) - \lambda_p\left(\left(\left(\frac{I(\pi^+,d)}{I(\pi_t,n)}\right)^{\theta-1} - 1\right)\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right)\right) \\ = \ & \mu_D(n) + \frac{1}{I^S(\pi_t,n)}\frac{dI^S(\pi_t,n)}{d\pi}\mu_\pi + \lambda_p\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right) + \frac{1}{I^S(\pi_t,n)}. \end{aligned} \tag{A.19}$$

Next, substituting for $r_t(n)$ from (A.13) in the above equation yields:

$$\begin{aligned} \frac{1}{I^S(\pi_t,n)} = \ & \beta + \rho\mu(n) - \mu_D(n) - \frac{1}{2}\gamma(1+\rho)\sigma(n)^2 + \gamma\sigma(n)\sigma_D(n) - \frac{1}{I^S(\pi_t,n)}\frac{dI^S(\pi_t,n)}{d\pi}\mu_\pi \\ & -\lambda_p\frac{1}{\theta}\left[\left(\frac{I(\pi_t^+,d)}{I(\pi_t,n)}\right)^\theta - 1\right] - \lambda_p\left(\frac{I(\pi_t^+,d)}{I(\pi_t,n)}\right)^{\theta-1}\left(\frac{I^S(\pi^+,d)}{I^S(\pi_t,n)} - \frac{I(\pi^+,d)}{I(\pi_t,n)}\right), \end{aligned} \tag{A.20}$$

By a similar procedure, we obtain the restriction for $I^S(\pi_t,d)$:

$$\begin{aligned} \frac{1}{I^S(\pi_t,d)} = \ & \beta + \rho\mu(d) - \mu_D(d) - \frac{1}{2}\gamma(1+\rho)\sigma(d)^2 + \gamma\sigma(d)\sigma_D(d) \\ & -v\frac{1}{\theta}\left[\left(\frac{I(\pi_t,d)}{I(\pi_t,n)}\right)^\theta - 1\right] - v\left(\frac{I(\pi_t,n)}{I(\pi_t,d)}\right)^{\theta-1}\left(\frac{I^S(\pi_t,n)}{I^S(\pi_t,d)} - \frac{I(\pi_t,n)}{I(\pi_t,d)}\right). \end{aligned} \tag{A.21}$$

B. Numerical procedures for the base model

B.1. Boundary conditions for $I(.)$ and $I^S(.)$

Unlike (A.8), (A.7) is a differential equation driven by the agent's belief update during the normal phase. To obtain the boundary condition for (A.7), we set

$$\mu_\pi = -\pi\left(\lambda^G - \bar{\lambda}\right) + (1-\pi)\phi_B - \pi\phi_G = 0$$

which has two roots. Denote by π^0 the root that lies between zero and one, at which (A.7) degenerates to

$$\frac{1}{I(\pi^0,n)} = \beta + (\rho-1)\mu(n) + \frac{1}{2}\gamma(1-\rho)\sigma(n)^2 - \lambda_p^0\frac{1}{\theta}\left[\left(\frac{I(\pi_0^+,d)}{I(\pi_0,d)}\right)^\theta - 1\right]. \tag{A.22}$$

In (A.22),

$$\pi^{0}=\frac{\left(\lambda^{G}-\lambda^{B}+\phi_{B}+\phi_{G}\right)-\sqrt{\left(\lambda^{G}-\lambda^{B}+\phi_{B}+\phi_{G}\right)^{2}-4\left(\lambda^{G}-\lambda^{B}\right)\phi_{B}}}{2\left(\lambda^{G}-\lambda^{B}\right)}; \tag{A.23}$$

$$\lambda_{p}^{0}\equiv\pi_{0}\lambda^{G}+\left(1-\pi_{0}\right)\lambda^{B} \tag{A.24}$$

which denotes the posterior estimation of disaster intensity at π_0;

$$\pi^{0,+}\equiv\pi^{0}\frac{\lambda^{G}}{\lambda_{p}^{0}} \tag{A.25}$$

which denotes the updated value of π following the strike of a disaster at the state π^0. (A.7)–(A.8) together with the boundary condition (A.22) fully characterize the model-implied W/C ratios.

Like that for $I(.)$, (A.21) is an algebraic equation, whereas (A.20) is a differential equation driven by the agent's belief update during the normal phase. Again we set $\pi=\pi^0$ in (A.20) to obtain its boundary condition of

$$\begin{aligned}\frac{1}{I^{S}\left(\pi_{0},n\right)} &= \beta+\rho\mu\left(n\right)-\mu_{D}\left(n\right)-\frac{1}{2}\gamma\left(1+\rho\right)\sigma\left(n\right)^{2}+\gamma\sigma\left(n\right)\sigma_{D}\left(n\right) \\ &\quad -\lambda_{p}^{0}\frac{1}{\theta}\left[\left(\frac{I\left(\pi_{0}^{+},d\right)}{I\left(\pi_{0},n\right)}\right)^{\theta}-1\right]-\lambda_{p}^{0}\left(\frac{I\left(\pi_{0}^{+},d\right)}{I\left(\pi_{0},n\right)}\right)^{\theta-1}\left(\frac{I^{S}\left(\pi_{0}^{+},d\right)}{I^{S}\left(\pi_{0},n\right)}-\frac{I\left(\pi_{0}^{+},d\right)}{I\left(\pi0,n\right)}\right),\end{aligned} \tag{A.26}$$

where $\pi^{0,+}$ and λ_p^0 are defined by (A.25) and (A.24). (A.20)–(A.21) together with the boundary condition (A.26) fully characterize the model-implied P/D ratios.

B.2. Solving $I(.)$ and $I^S(.)$

We use three steps to formulate the approximate solutions to $\left(I(\pi,n),I(\pi,d)\right)$ using the collocation method (e.g., Miranda and Fackler, 2002). The procedures for solving $\left(I^S(\pi,n),I^S(\pi,d)\right)$ are very similar, and we omit the details to save space.

Step 1: choice of basis functions. We choose the Chebyshev polynomials defined on $[-1,1]$ which are given by

$$q_{0}\left(z\right)=1,q_{1}\left(z\right)=z,q_{2}\left(z\right)=2z^{2}-1,$$

and computed recursively by

$$q_{j}\left(z\right)=2zq_{j-1}\left(z\right)-q_{j-2}\left(z\right)$$

for polynomials with orders higher than two. For the more general domains $[a,b]$, these polynomials become $p_j(x)$ under the transformation

$$x = a + \frac{(z+1)(b-a)}{2}.$$

When using Chebyshev polynomials up to the mth order to approximate some unknown functional, numerical analysis and empirical experience both suggest that polynomial approximants over the interval $[a,b]$ should be constructured by interpolating the unknown function at the Chebyshev nodes defined by

$$x_i = \frac{a+b}{2} + \frac{b-a}{2}\cos\left(\frac{n-i+0.5}{n}\Pi\right) \text{ for } i = 1, ..., m, \tag{A.27}$$

where Π denotes the ratio of a circle's circumference to its diameter. For the probability variable π in our model, $a = 0$, and $b = 1$.

Step 2: residual operator. Conditional on the normal phase, let $\mathcal{L}^n$ be the residual operator associated with (A.7) and its boundary condition (A.22), i.e.,

$$\begin{aligned}\mathcal{L}^n(I(\pi,n)) &= \beta + (\rho-1)\mu(n) + \frac{1}{2}\gamma(1-\rho)\sigma(n)^2 - \frac{1}{I(\pi,n)}\frac{dI(\pi,n)}{\partial\pi}\mu_\pi \\ &\quad -\lambda_p\frac{1}{\theta}\left[\left(\frac{I(\pi^+,d)}{I(\pi,n)}\right)^\theta - 1\right] - \frac{1}{I(\pi,n)},\end{aligned} \tag{A.28}$$

$$\mathcal{L}^n(I(\pi_0,n)) = \beta + (\rho-1)\mu + \frac{1}{2}\gamma(1-\rho)\sigma^2 - \lambda_p^0\frac{1}{\theta}\left[\left(\frac{I(\pi_0^+,d)}{I(\pi_0,n)}\right)^\theta - 1\right] - \frac{1}{I(\pi_0,n)}. \tag{A.29}$$

Similarly, let $\mathcal{L}^d(I(\pi,d))$ be the residual operator associated with (A.8) conditional on the disaster phase, i.e.,

$$\mathcal{L}^d(I(\pi,d)) = \beta + (\rho-1)\mu(d) + \frac{1}{2}\gamma(1-\rho)\sigma(d)^2 - v\frac{1}{\theta}\left[\left(\frac{I(\pi,n)}{I(\pi,d)}\right)^\theta - 1\right] - \frac{1}{I(\pi,d)}. \tag{A.30}$$

Denote by $\left(\hat{I}(\pi,n), \hat{I}(\pi,d)\right)$ the solution to (A.7)–(A.8) and (A.22). In terms of the above operators, we must have

$$\mathcal{L}^n\left(\hat{I}(\pi,n)\right) = \mathcal{L}^n\left(\hat{I}(\pi_0,n)\right) = 0. \tag{A.31}$$

$$\mathcal{L}^d\left(\hat{I}(\pi,d)\right) = 0. \tag{A.32}$$

Step 3: We appeal to the Chebyshev Interpolation Theorem (e.g., Judd (1998)) to find an approximate solution to (A.31)–(A.32), which we denote by $\left(I^C(\pi,n), I^D(\pi,d)\right)$. The approximation is obtained by evaluating $\mathcal{L}\left(\hat{I}\left(.\right)\right)$ at a chosen set of points, and setting it to zero at each of these points. More specifically, we write

$$I\left(\pi,s\right)=\sum_{j=1}^{m}c_{j}\left(s\right)p_{j}\left(\pi\right),$$

for $s\in\{n,d\}$, where $\{c_j(s)\}_{j=1}^{m}$ are basis coefficients to be determined. For both $I^C(\pi,n)$ and $I^C(\pi,d)$, we use basis Chebyshev polynomials up to the 15th order. Adding polynomials with higher orders change little the results. To solve (A.31), we set $\mathcal{L}(I(\pi,n))=0$ at the first 14 Chebyshev nodes given by (A.27) with m set to 14. This restriction, together with $\mathcal{L}(I(\pi_0,n))=0$, yields a total of 15 equations for determining $\{c_j(n)\}_{j=1}^{15}$. To solve (A.32), we set $\mathcal{L}(I(\pi,d))=0$ at the first 15 Chebyshev nodes, which yields 15 equations for determinating $\{c_j(d)\}_{j=1}^{15}$.

Chebyshev basis polynomials in combination with Chebyshev interpolation nodes yield an extremely well-conditioned interpolation equation that can be accurately and efficiently solved. To show it, Panel A of Figure A1 plots the resduals of $\mathcal{L}\left(I^C(\pi,n)\right)$ and $\mathcal{L}\left(I^C(\pi,d)\right)$ at a refined grid of 50 evenly spaced points between zero and one. The implied residuals oscillate fairly evenly through the domain $[0,1]$ with the implied pricing errors generally below 2×10^{-5} and even lower for that of $\mathcal{L}\left(I^C(\pi,d)\right)$. Panel B of the same figure plots residuals in the similar way for the P/D ratio which again shows fairly low pricing errors.

B.3. Computing option prices

We first describe the procedure for simulating the terminal option payoff on a typical path. In particular, we start from $(\pi,s)=(\pi_0,s_0)$ and normalize the initial values of both stock price P_0 and pricing kernel Λ_0 to one.

1. For each $0\le t\le\tau$, check the current economic phase s_t and compute variables related to jumps of the wealth-consumption ratio. For example, if $s_t=n$, the jump intensity is λ_{pt} defined by (2.9) and the jump size of W/C ratio equals

$$\frac{I\left(\pi_t^{+},d\right)}{I\left(\pi_t,n\right)}=\frac{I\left(\pi_t\frac{\lambda^G}{\lambda_{pt}},d\right)}{I\left(\pi_t,n\right)}.$$

2. Simulate the evolution of the state π_t, the dividend D_t, and the pricing kernel Λ_t. In particular, if $s_t=d$, simulate (π_t,D_t,Λ_t) according to (2.7)–(2.8), (3.7), and (3.4). If $s_t=d$, stop updating π_t, and simulate D_t and Λ_t according to (3.7) and (3.5).

3. Compute the terminal P_τ using $I^S(\pi_\tau,s_\tau)\cdot D_\tau$. The terminal payoff for a put option characterized by (τ,K) is thus given by $\frac{\Lambda_\tau}{\Lambda_0}\max(K-P_\tau,0)$, where τ and K denote the time to maturity and the strike price, respectively.

By simulating a large number of paths starting from the same (π_0,s_0) and taking average of the terminal payoffs, we obtain the model-implied put option price which we denote by $I^O(\pi_0, s_0)$, where we focus on $s_0 = n$ in the paper. Next, we approximate $I^O(\pi_0, s_0)$ through

$$I^O(\pi_0, s_0) = \sum_{j=1}^{m} c_j^O(s_0)\, p_j(\pi_0) \tag{A.33}$$

where $p_j(\pi_0)$ is the jth order Chebyshev polynomial introduced Appendix B.2; $\left\{c_j^O(s_0)\right\}_{j=1}^{m}$ are basis coefficients to be determined. Again, we use Chebyshev polynomials up to the 15th order, and set

$$\sum_{j=1}^{m} c_j^O(s_0)\, p_j(\pi_0) - I^{O,s}(\pi_0, s_0) = 0 \tag{A.34}$$

at the first 15 Chebyshev nodes given by (A.27) with $a = 0$ and $b = 1$. The implied restrictions from (A.34) exactly identify $\left\{c_j^O(s)\right\}_{j=1}^{m=15}$. Finally, we simulate a large number of realizations of the probability state in its stationary distribution region denoted by $\{\pi_i\}_{i=1}^{N}$, and the approximate form (A.33) allows us to conveniently compute the implied option price conditional on the given π_i. Taking the average of $I^O(\pi_i, s)$ for all i, we obtain the model-implied unconditional option price for the given economic phase s_0.

C. Solving the general model

C.1. Differential equations

Using a similar procedure to that described in Appendix A.2., we obtain the following differential equations for the W/C ratio in the general model:

$$\begin{aligned}\frac{1}{I(\pi, \pi^v, n)} &= \beta + (\rho - 1)\,\mu(n) + \frac{1}{2}\gamma(1-\rho)\,\sigma(n)^2 \\ &\quad - \frac{1}{I(\pi, \pi^v, n)}\frac{\partial I(\pi, \pi^v, n)}{\partial \pi}\mu_\pi - \lambda_p \frac{1}{\theta}\left[\left(\frac{I(\pi^+, \pi^v, d)}{I(\pi, \pi^v, n)}\right)^\theta - 1\right], \quad \text{(A.35)}\end{aligned}$$

$$\begin{aligned}\frac{1}{I(\pi, \pi^v, d)} &= \beta + (\rho - 1)\,\mu(d) + \frac{1}{2}\gamma(1-\rho)\,\sigma(d)^2 \\ &\quad - \frac{1}{I(\pi, \pi^v, d)}\frac{\partial I(\pi, \pi^v, d)}{\partial \pi^v}\mu_\pi^\nu - v_p \frac{1}{\theta}\left[\left(\frac{I(\pi, , \pi^{v,+}, n)}{I(\pi, , \pi^v, d)}\right)^\theta - 1\right], \quad \text{(A.36)}\end{aligned}$$

where μ_π and μ_π^ν are given by (2.8) and (6.4); we've used that π stops evolving during the disaster phase, and that π^v stops evolving during the normal phase. Similar to the base model, the two boundary conditions for the above differential equations are:

$$\frac{1}{I\left(\pi_0,\pi^v,n\right)}=\beta+(\rho-1)\,\mu\,(n)+\frac{1}{2}\gamma\,(1-\rho)\,\sigma\,(n)^2-\lambda_p^0\frac{1}{\theta}\left[\left(\frac{I\left(\pi_0^+,\pi^v,d\right)}{I\left(\pi_0,\pi^v,d\right)}\right)^{\theta}-1\right], \quad \text{(A.37)}$$

$$\frac{1}{I\left(\pi,\pi_0^v,d\right)}=\beta+(\rho-1)\,\mu\,(d)+\frac{1}{2}\gamma\,(1-\rho)\,\sigma\,(d)^2-v_p^0\frac{1}{\theta}\left[\left(\frac{I\left(\pi,,\pi_0^{v,+},n\right)}{I\left(\pi,,\pi_0^v,d\right)}\right)^{\theta}-1\right], \quad \text{(A.38)}$$

where

$$\pi_0^v=\frac{\left(\nu^G-\nu^B+\psi_B+\psi_G\right)-\sqrt{\left(\nu^G-\nu^B+\psi_B+\psi_G\right)^2-4\left(\nu^G-\nu^B\right)\psi_B}}{2\left(\nu^G-\nu^B\right)}; \quad \text{(A.39)}$$

$\pi_{0,}$ λ_p^0, and π_0^+ are given by (A.23)–(A.25);

$$\pi_0^{v,+}\equiv\pi_0^v\frac{\nu^G}{\nu_p^0}; \quad \text{(A.40)}$$

$$\nu_p^0\equiv\pi_0^v\nu^G+\left(1-\pi_0^v\right)\nu^B \quad \text{(A.41)}$$

$$\mu_\pi=-\pi\left(\lambda^G-\bar{\lambda}\right)+(1-\pi)\,\phi_B-\pi\phi_G=0.$$

Derivations of differential equations and boundary conditions for the P/D ratio $I^S(.)$ are similar, and we omit the details.

C.2. Approximate solution in the general model

We still use $I(.)$ as the example to illustrate the Chebyshev approximation of solutions to the general model in which the recovery rate is also inaccurately observed. Given that we have two states now: π and π^v, we need to construct a two-dimensional function basis defined on

$$S\equiv\left\{(\pi,\pi^v):0\leq\pi\leq1,0\leq\pi^v\leq1\right\}.$$

This is done through the tensor product of univariate Chebyshev polynomials adjusted to the domain $[0,1]^2$, i.e.,

$$p_{j_1,j_2}\left(\pi,\pi^v\right)=p_{j_1}(\pi)\,p_{j_2}\left(\pi^v\right),$$

where $p_{j_1}(\pi)$ and $p_{j_2}(\pi^v)$ are both one-dimensional Chebyshev polynomials restricted to $[0,1]$. The W/C ratio is thus approximated by

$$I\left(\pi,\pi^v,s\right)=\sum_{j_2=1}^{m_2}\sum_{j_1=1}^{m_1}c_{j_1,j_2}(s)\,p_{j_1,j_2}\left(\pi,\pi^v\right),$$

where $\{c_{j_1,j_2}(s)\}$ represent $m = m_1 m_2$ basis coefficients to be determined. On the other hand, we form a grid of $m = m_1 m_2$ interpolation nodes by forming the tensor product of univariate interpolation nodes which are denoted by

$$\left\{\left(\pi_{j_1}, \pi^v_{j_2}\right) : j_1 = 1, ..., m_1; j_2 = 1, ..., m_2\right\}.$$

Similarly to that for the base model, we use $\mathcal{L}^n$ as the residual operator associated with differential equation and boundary condition for $I(., n)$ given by (A.35) and (A.37), and use $\mathcal{L}^d$ as the residual operator associated with differential equation and boundary condition for $I(., d)$ given by (A.36) and (A.38). Denote by $\left(\hat{I}(\pi, \pi^v, n), \hat{I}(\pi, \pi^v, d)\right)$ the solution to (A.35)–(A.38), which must satisfy:

$$\mathcal{L}\left(\hat{I}(\pi, \pi^v, n)\right) = \mathcal{L}\left(\hat{I}(\pi_0, \pi^v, n)\right) = 0. \tag{A.42}$$

$$\mathcal{L}\left(\hat{I}(\pi, \pi^v, d)\right) = \mathcal{L}\left(\hat{I}(\pi, \pi^v_0, d)\right) = 0. \tag{A.43}$$

To solve (A.42), we set $\mathcal{L}(I(\pi, \pi^v, n)) = 0$ at the $(m_1 - 1) m_2$ nodes formed as the tensor product of the first $m_1 - 1$ Chebyshev nodes for π and the first m_2 Chebyshev nodes for π^v. We then set $\mathcal{L}(I(\pi_0, \pi^v, n)) = 0$ at the m_2 Chebyshev nodes for π^v. Altogether, we have $m_1 m_2$ restrictions which exactly identify the coefficients for the normal phase, $\{c_{j_1,j_2}(n)\}$. To solve (A.43), we set $\mathcal{L}(I(\pi, \pi^v, d)) = 0$ at the $m_1(m_2 - 1)$ nodes formed as the tensor product of the first m_1 Chebyshev nodes for π and the first m_2- Chebyshev nodes for π^v. We then set $\mathcal{L}(I(\pi, \pi^v_0, d)) = 0$ at the m_1 Chebyshev nodes for π. Put together, we again have $m_1 m_2$ restrictions which exactly identify the coefficients for the disaster phase $\{c_{j_1,j_2}(d)\}$. In the actual implemention, we choose $m_1 = m_2 = 15$, and the program for solving the implied 225×2 equations converge quickly with fairly high accuracy.

References

[1] Abel, A. B., 1990, Risk premia and term premia in general equilibrium. Journal of Monetary Economics 43, 3–33.

[2] Ai, H., 2010, Information quality and long-run risk: asset pricing implications. Journal of Finance 65, 1333–67.

[3] Bansal, R., and A. Yaron, 2004, Risks for The long-run: a potential resolution of asset pricing puzzles, Journal of Finance 59, 1481-1509

[4] Barro, R. J., 2006, Rare disasters and asset markets in the twentieth century, Quarterly Journal of Economics 121, 823-866.

[5] Barro, R. J., 2009, Rare Disasters, Asset Prices, and Welfare Costs," American Economic Review 99, 243–264.

[6] Barro, R., and J. Ursua, 2008, Macroeconomic crises since 1987, Brookings Papers on Economic Activity, 255–335.

[7] Beeler, J., and J. Y. Campbell, 2009, The long-run risks model and aggregate asset prices: an empirical assessment, working paper, Havard University.

[8] Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, Journal of Financial Economics 101, 552-573.

[9] Chen, H., S. Joslin, and N. Tran, 2012, Rare disasters and risk sharing with heterogeneous beliefs, Review of Financial Studies, forthcoming.

[10] Cochrane, J., 2001. Asset Pricing. Princeton University Press, Priceton, NJ.

[11] Collin-Dufresne, P., R. Goldstein, and J. Helwege, 2008, Is credit event risk priced? Modeling contagion via the updating of beliefs. Working paper, Columbia University.

[12] David, A., 1997, Fluctuating confidences in stock markets: implications for returns and volatility, Journal of Financial and Quantitative Analysis 32, 427–462.

[13] David, A., 2008, Heterogeneous beliefs, speculation, and the equity premium, Journal of Finance 63, 41–83.

[14] David, A., and P. Veronesi, 2002, Option prices with uncertain fundamental, working paper, University of Calgary and University of Chicago

[15] Du, D., 2011, General equilibrium pricing of options with habit formation and event risks, Journal of Financial Economics 99, 400–426.

[16] Duffie, D., and L. G. Epstein, 1992, Stochastic differential utility, Econometrica 60, 353–394.

[17] Gabaix, X., 2011, Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance, Quarterly Journal of Economics, forthcoming.

[18] Gourio, F., 2008, Disasters and Recoveries, American Economic Review - p&p 98, 68-73.

[19] Gourio, F., 2011, Credit risk and disaster risk, working paper, Boston University.

[20] Judd, K. L., 1998, Numerical Methods in Economics. MIT Press, Cambridge, MA.

[21] Kreps, D. M., and E. L. Porteus, 1978, Temporal resolution of uncertainty and dynamic choice theory, Econometrica 46, 185–200.

[22] Lettau, M., Ludigson, S., and Watcher, J., 2007, The declining equity premium: What role does macroeconomic risk play? Review of Financial Studies 21, 1653-1687.

[23] Liptser, R. S., and A. N. Shiryayev, 2001, Statistics of Random Processes: I, II, Springer-Verlag, New York.

[24] Liu, J., J. Pan, and T. Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, Review of Financial Studies 18, 131–164.

[25] Mao, X., C. Yuan, 2006, Stochastic differential equations with markovian switch-ing, Imperial College Press.

[26] Mehra, R., E. Prescott, 1985, The equity premium puzzle, Journal of Monetary Economics 15, 145-161.

[27] Miranda, M. J., and P. L. Fackler, 2002, Applied Computational Economics and Finance. The MIT Press, Cambridge, MA.

[28] Nakamura, E., J. Steinsson, R. J. Barro, and J. Ursua, 2011, Crises and recoveries in an empirical model of consumption disasters, working paper, Columbia University and Havard University.

[29] Pan, J., 2002, The jump-risk premia implicit in options: evidence from an integrated time series study. Journal of Financial Economics 63, 3–50.

[30] Rietz, T., 1988, The equity premium puzzle: a solution? Journal of Monetary economics 21, 117–132

[31] Veronesi, P. 1999, Stock market overreaction to bad news in good times: a rational expectations equilibrium model. Review of Financial Studies 12, 975-1007.

[32] Veronesi, P. 2000, How does information quality affect stock returns? Journal of Finance 55, 807–837.

[33] Wachter, J., 2011, Can time-varying risk of rare disasters explain aggregate stock market volatility? working paper, Wharton.

[34] Weitzman, M. L., 2007, Subjective expectations and asset-return puzzles, American Economic Review 97, 1102-1130.

[35] Working, H., 1960, Note on the correlation of first differences of averages in a random chain, Econometrica 28, 916–918.

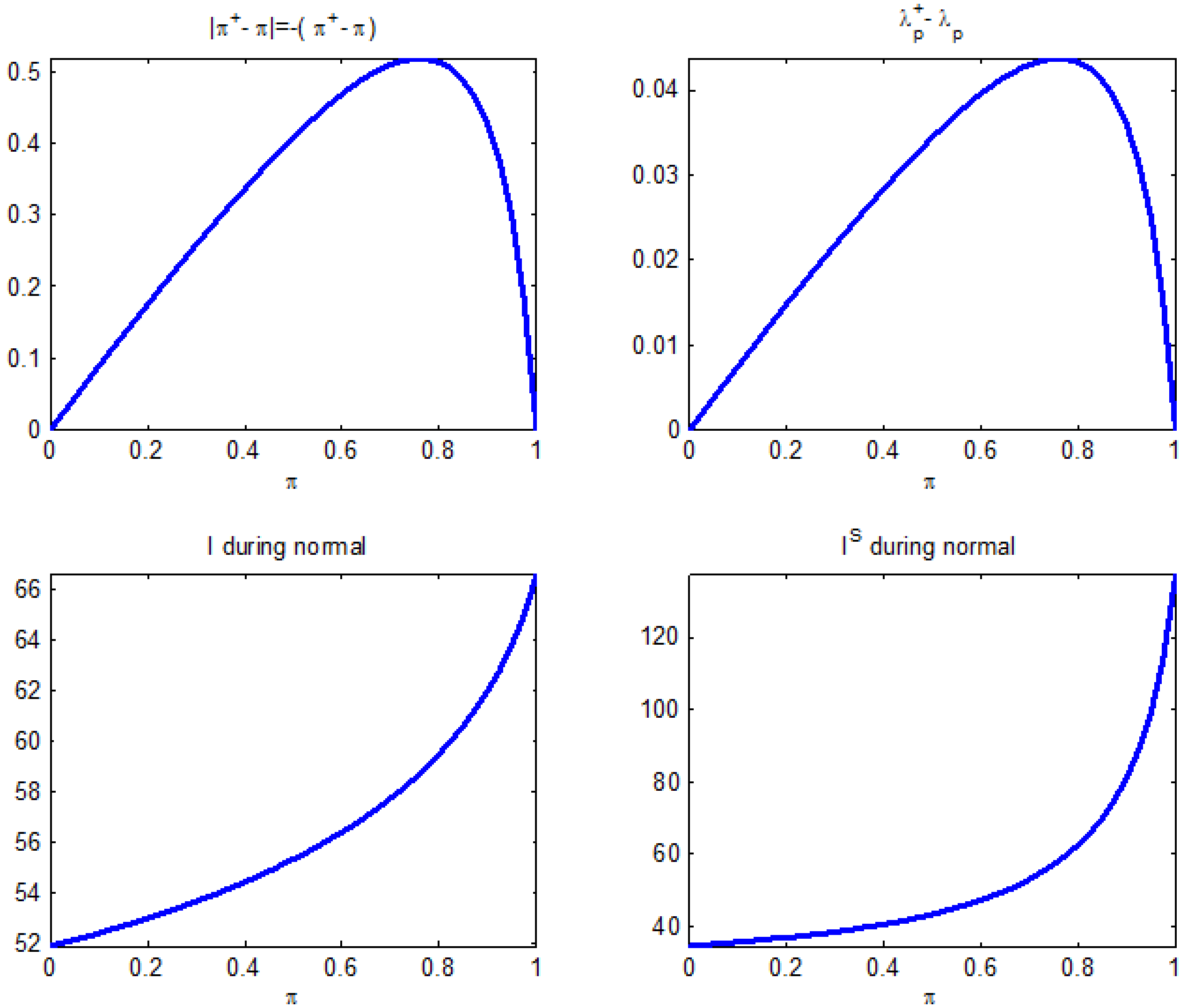

Figure 1: Jump sizes of the state and the value-fundamental ratios in the base model. In the base model, π denotes the posterior probability that the disaster rate λ is in the good regime. The top two panels plot the state dependences of the absolute jump sizes of π and λ_p, where λ_p denotes the posterior estimation of λ. The bottom two panels plot the state dependences of the wealth-consumption ratio $I(.)$ and the price dividend ratio $I^S(.)$ conditional on the normal phase.

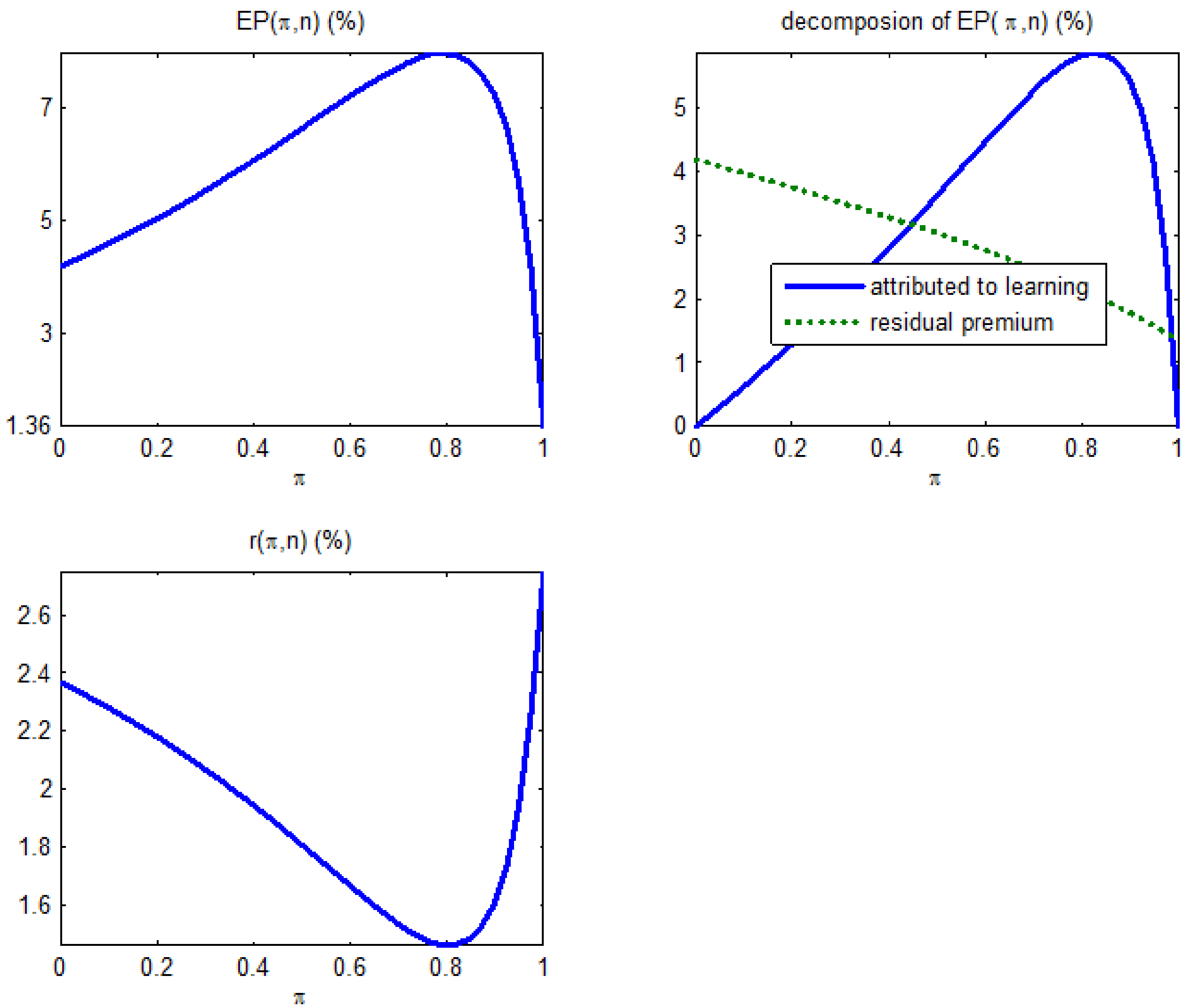

Figure 2: State dependences of equity premium and short-term rate. The top two panels plot the state dependences of the equity premium conditional on the normal phase, $EP(\pi, n)$, and its decompositions. In particular, the part of $EP(\pi, n)$ attributed to learning is given by (4.7). The bottom left panel plots the state dependences of the short-term rate conditional on the normal phase.

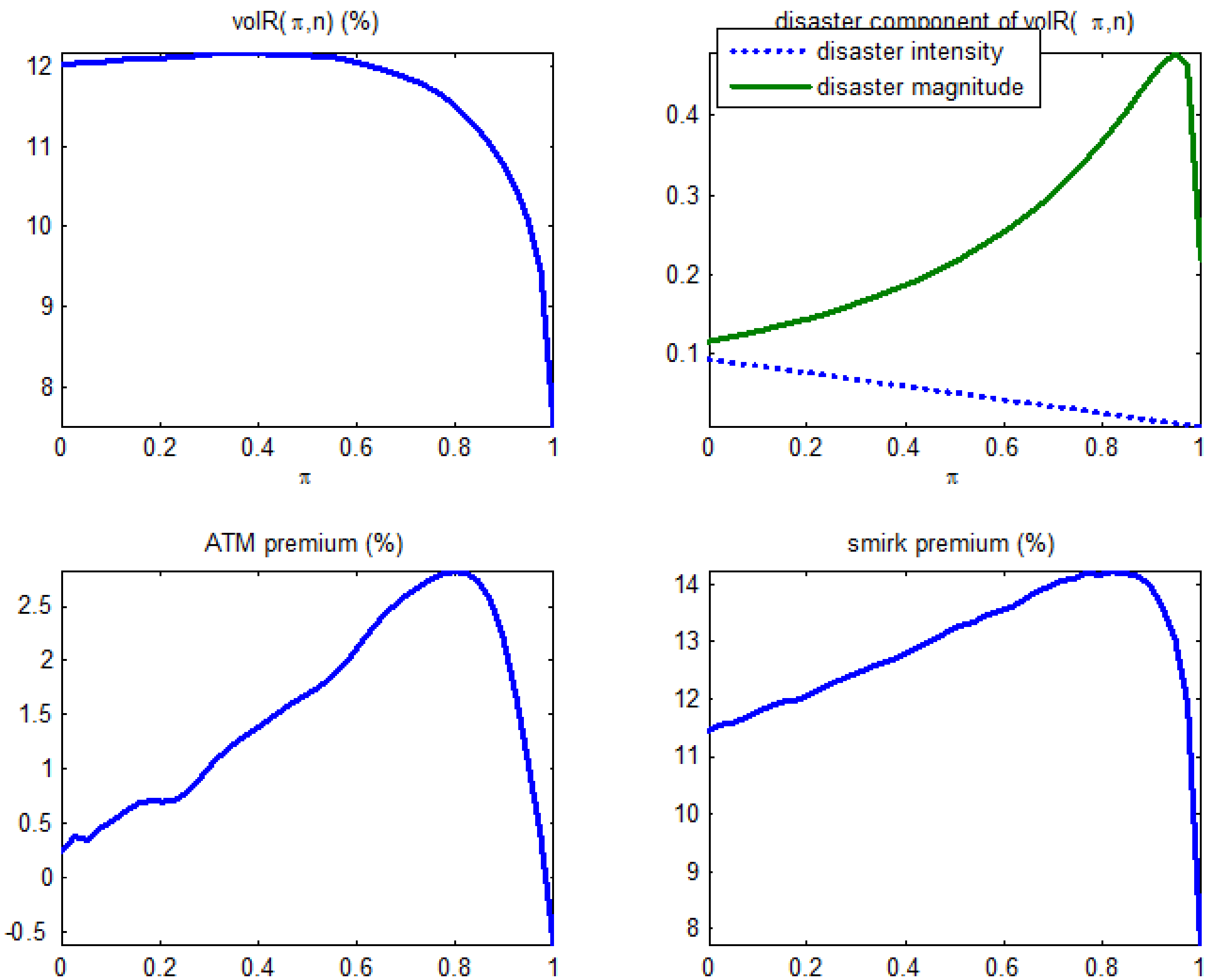

Figure 3: State dependences of return volatility and option implied premiums in the base model. The top left panel plots the state dependence of the stock return volatility and the top right panel plots the decomposition of its jump component in terms of i) disaster intensity λ_p; and ii) the squared jump size $\left(\frac{I^S(\pi_t^+,d)}{I^S(\pi_t,n)} - 1\right)^2$. The bottom left panel plots the state-dependences of ATM premium defined as the difference between ATM implied volatility and the total stock return volatility. The bottom right panel plots the state-dependences of smirk premium defined as the difference between 10% OTM implied volatilities and the ATM implied volatilities. All moments are conditional on the normal phase.

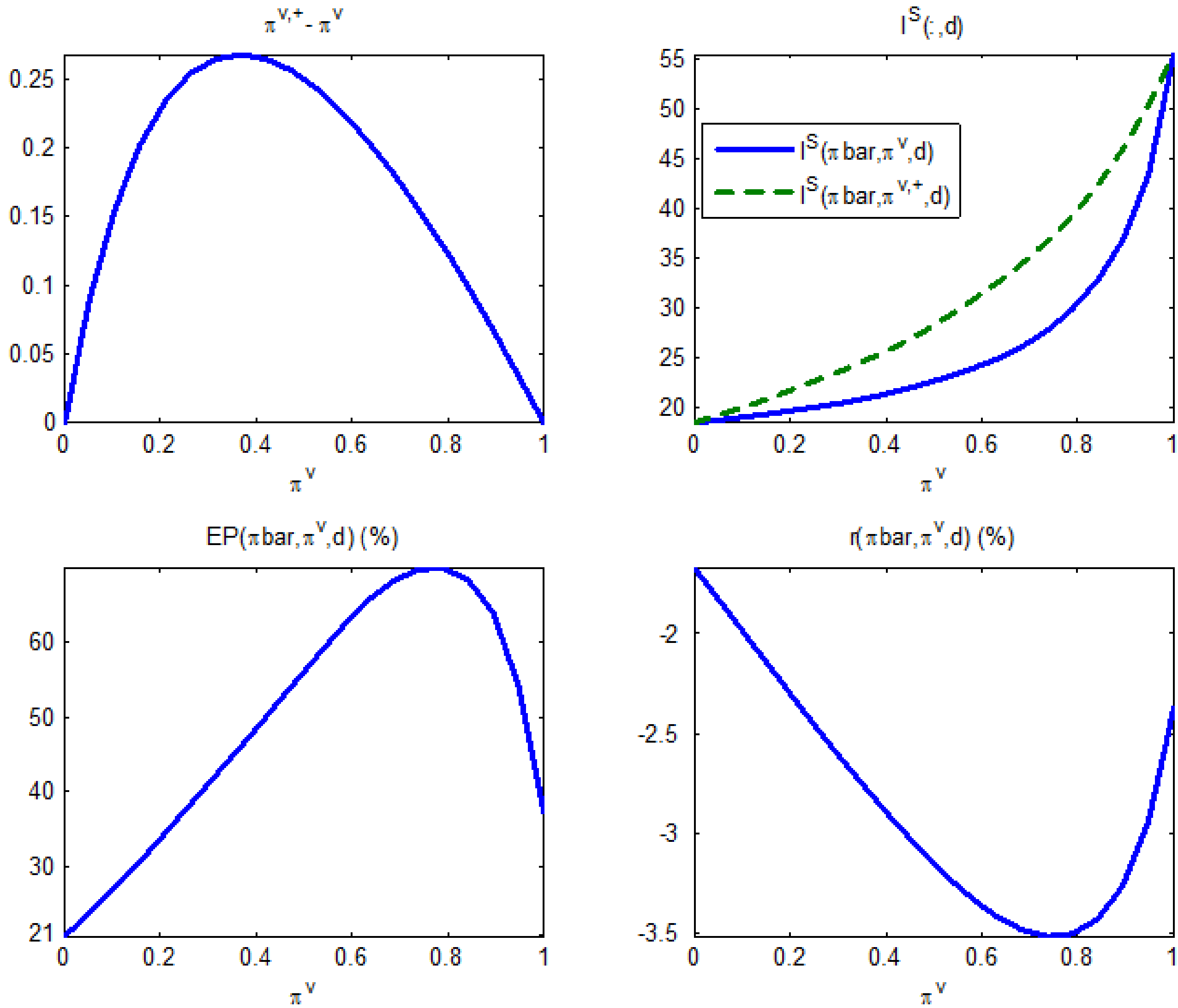

Figure 4: State dependences in the general model. π and π^{ν} denote the posterior probability that the disaster rate λ is in the good regime and the posterior probability that the recovery rate ν is in its good regime, respectively, which serve as the states for pricing. In all panels, we plot state dependences as the function of π^{ν} while π is fixed at its long-run average $\bar{\pi}$. The top left panels plot the jump sizes of ν_p, the posterior estimation about the recovery rate. The top right panel plots the P/D ratio conditional on the disaster phase. The bottom two panels plot the equity premium and the short term rate, both of which are conditional on the disaster phase.

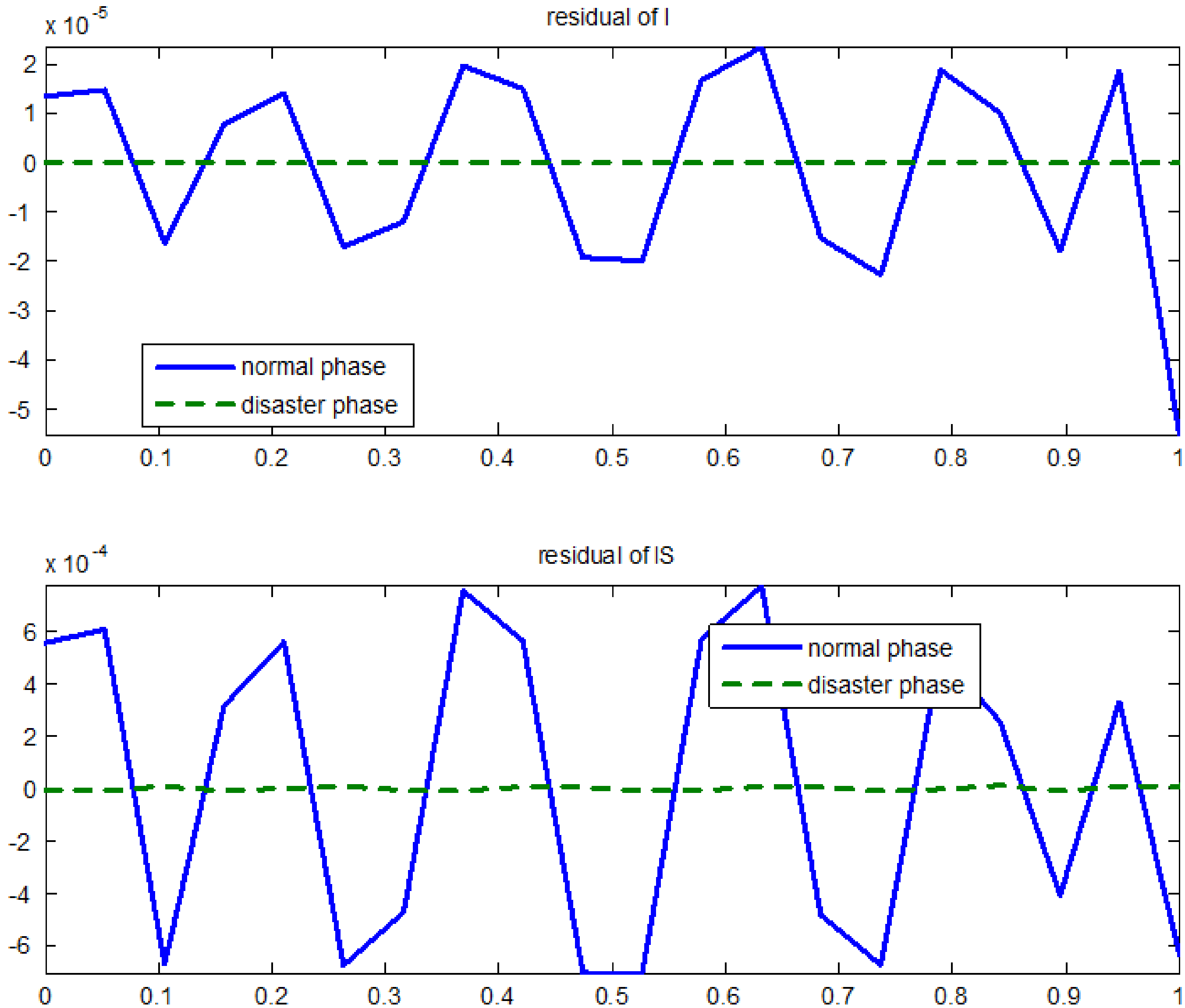

Figure A1: Residuals from computing the W/C ratio I and P/D ratio I^S using Chebyshev interpolation. Within our setup, the implied W/C ratio $I(.)$ and P/D ratio $I^S(.)$ cannot be solved analytically. We thus resort to its numerical solution using Chebyshev interpolation with details in Appendix B.2 for the base model and in Appendix C.2 for the general model. The top and the bottom panel of Figure A1 plot the solution residuals for $I(.)$ and $I^S(.)$, respectively, in the base model conditional on both the normal phase and the disaster phase.

Table I. Implications about the P/D ratio, the short-term rate, the equity premium, and the return volatility

Panel A reports implications about the P/D ratio, the short-term rate, the equity premium, and the return volatility in a degenerated model without learning. We report moment values conditional on both the normal and the disaster phases, as well as their weighted averages across the two economic phases with weights set proportional to the durations of the two phases.Panel B reports implications of the same moments from the base model, in which the posterior probability that the disaster rate λ is in the good regime, denoted by π, serves as the state. We report both the conditional moment values evaluated at $\bar{\pi}$, the long-run average of π, and the unconditional moment values, computed as averages of the moment realizations over the stationary region of π.

Panel A: without learning				
	P/D	r (%)	EP (%)	$volR$ (%)
normal	894	2.63	2.29	8.97
disaster	437	-1.85	26.6	62.1
average	872	2.41	3.48	11.6

Panel B: the base model								
	evaluated at $\pi = \bar{\pi}$				averages over $\pi-$distribution			
	P/D	r (%)	EP (%)	$volR$ (%)	P/D	r (%)	EP (%)	$volR$ (%)
normal	83.4	1.66	6.99	10.6	102	2.11	4.88	9.93
disaster	47.2	-1.65	19.3	47.0	55.7	-1.68	20.2	48.7
Average	82.1	1.49	7.59	12.4	99.3	1.93	5.63	11.8

Table II. Implications about option pricing

Table III reports implications about option pricing in terms of the stock return volatility $volR$, the implied volatilities from ATM options, and the implied volatilities from 10% OTM options. Panel A reports implications from a degenerated model without learning. Panels B&C reports implications of the same moments from the base model. In this case, the posterior probability that the disaster rate λ is in the good regime, denoted by π, serves as the state. More specifically, Panel B reports values computed at $\bar{\pi}$, the long-run average of π; Panel C reports the unconditional values computed over the stationary region of π. All reported moment values are for the normal phase.

Panel A: without learning		
volR (%)	ATM vol (%)	OTM vol (%)
8.97	7.48	17.9

Panel B: base model: evaluated at $\pi = \bar{\pi}$		
volR (%)	ATM vol (%)	OTM vol (%)
10.6	12.8	26.6

Panel C: base model: averages over $\pi-$distribution		
volR (%)	ATM vol (%)	OTM vol (%)
9.93	11.8	24.3

Table III. Implications from the general model

Panel A reports implications about the P/D ratio, the short-term rate, the equity premium, and the stock return volatility implied from the general model. In this case, the posterior probability that the disaster rate λ is in the good regime and the posterior probability that the recovery rate ν is in its good regime, denoted π and π^{ν}, respectively, serve as the states. We report moment values conditional on the normal and disaster phases, as well as their weighted averages across the two economic phases with weights set proportional to the durations of the two phases. In addition, we report both the moment values evaluated at $(\pi, \pi^v) = (\bar{\pi}, \bar{\pi}^v)$, where $\bar{\pi}$ and $\bar{\pi}^v$ are the long-run average of π and π^v, and the unconditional moment values, computed as averages of the moment realizations over the stationary region of π and π^v. Panel B&C reports option pricing implications during the normal phase in terms of the stock return volatilities, $volR$, the implied volatilities from ATM options, and the implied volatilities from 10% OTM options. Again, we report both the conditional moments evaluated at $(\pi, \pi^v) = (\bar{\pi}, \bar{\pi}^v)$ and the unconditional moments computed as averages of the moment realizations over the stationary region of π and π^v. All reported moments are conditional on the normal phase.

Panel A: bond and stock pricing								
	Evaluated at $(\pi, \pi^{\nu}) = (\bar{\pi}, \bar{\pi}^{\nu})$				Averages over (π, π^{ν}) distribution			
	P/D	r (%)	EP (%)	$volR$ (%)	P/D	r (%)	EP (%)	$volR$ (%)
normal	41.3	1.09	8.22	10.3	52.6	1.69	6.16	9.71
disaster	20.1	-2.46	37.4	76.5	24.5	-2.30	38.3	78.4
Average	40.3	0.91	9.64	13.3	51.3	1.49	7.73	13.1

Panel B: option pricing: evaluated at $(\pi, \pi^{\nu}) = (\bar{\pi}, \bar{\pi}^{\nu})$		
volR (%)	ATM vol (%)	OTM vol (%)
10.3	14.3	28.1

Panel C: option pricing: averages over (π, π^{ν}) distribution		
volR (%)	ATM vol (%)	OTM vol (%)
9.71	12.8	25.2