ESSAYS IN ASSET PRICING AND TAIL RISK

Sang Byung Seo

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Signature _____

Jessica A. Wachter

Richard B. Worley Professor of Financial Management, Professor of Finance

Graduate Group Chairperson

Signature _____

Eric T. Bradlow

K.P. Chao Professor of Marketing, Statistics, and Education

Dissertation Committee

Jessica A. Wachter, Richard B. Worley Professor of Financial Management, Professor of Finance

Nikolai Roussanov, Associate Professor of Finance

Amir Yaron, Robert Morris Professor of Banking & Finance

ACKNOWLEDGEMENT

First and foremost, I am deeply grateful to my advisors Nikolai Roussanov, Jessica Wachter, and Amir Yaron for their invaluable guidance.

I would also like to thank my friends and classmates who made my time at Wharton more enjoyable: Ian Appel, Christine Dobridge, Michael Lee, Gil Segal, and Colin Ward. I would like to especially thank Mete Kilic for his amazing help and feedback.

Finally, I am grateful to my family and Rosie for always supporting me no matter what.

ABSTRACT

ESSAYS IN ASSET PRICING AND TAIL RISK

Sang Byung Seo

Jessica A. Wachter

The first chapter "Option Prices in a Model with Stochastic Disaster Risk," co-authored with Jessica Wachter, studies the consistency between the rare disaster mechanism and options data. In contrast to past work based on an iid setup, we find that a model with stochastic disaster risk can explain average implied volatilities well, despite being calibrated to consumption and aggregate market data alone. Furthermore, we extend the stochastic disaster risk model to a two-factor model and show that it can match variation in the level and slope of implied volatilities, as well as the average implied volatility curves.

The second chapter "Do Rare Events Explain CDX Tranche Spreads?," also co-authored with Jessica Wachter, investigates the rare disaster mechanism based on the data on the CDX index and its tranches. Senior tranches are essentially deep out-of-the-money options because they do not incur any losses until a large number of investment-grade firms default. Using the two-factor stochastic disaster model, we jointly explain the spreads on each CDX tranche, as well as prices on put options and the aggregate market. This paper demonstrates the importance of beliefs about rare disasters and shows a basic consistency in these beliefs across different asset markets.

In the third chapter, "Correlated Defaults and Economic Catastrophes: Linking the CDS Market and Asset Returns," I consider economic catastrophes as massive correlated defaults and construct a catastrophic tail risk measure from joint default probabilities based on my model as well as information contained in the CDS data. Using the rich information contained in this measure, I find that investors put more weight on future extreme events even after the stock market showed signs of recovery from the recent financial crisis.

Furthermore, I show that high catastrophic tail risk robustly predicts high future excess returns for various assets, including stocks, government bonds, and corporate bonds. This risk is negatively priced, generating substantial dispersion in the cross section of stock returns. These results consistently indicate that seemingly impossible economic catastrophes are considered as an important risk source when trading assets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Option Prices in a Model with Stochastic Disaster Risk

(with Jessica A. Wachter)

1.1. Introduction

Century-long evidence indicates that the equity premium, namely the, expected return from holding equities over short-term debt, is economically significant. Ever since Mehra and Prescott (1985) noted the puzzling size of this premium relative to what a standard model would predict, the source of this premium has been a subject of debate. One place to look for such a source is in options data. By holding equity and a put option, an investor can, at least in theory, eliminate the downside risk in equities. For this reason, it is appealing to explain both options data and standard equity returns together with a single model.

Such an approach is arguably of particular importance for a class of macro-finance models that explain the equity premium through the mechanism of consumption disasters (e.g. Rietz (1988), Barro (2006) and Weitzman (2007)). In these models, consumption growth rates and thus equity returns are subject to shocks that are rare and large. Options (assuming away, for the moment, the potentially important question of counterparty risk), offer a way to insure against the risk of these events. Thus it is of interest to know whether these models have the potential to explain option prices as well as equity prices.

Thus, option prices convey information about the price investors require to insure against losses, and therefore indirectly about the source of the equity premium. Indeed, prior literature uses option prices to infer information about the premium attached to crash risk (e.g. Pan (2002)); recent work connects this risk to the equity premium as a whole (Bollerslev and Todorov (2011b) and Santa-Clara and Yan (2010)). Despite the clear parallel to the literature on consumption disasters and the equity premium, the literatures have advanced separately. The macro-finance literature takes as its source not options data, but rather international data on large consumption declines (see Barro and Ursua (2008)). There is

also a difference in terminology (which reflects a difference in the philosophy). In the options literature, the focus is on sudden negative changes in prices, called crashes; connections to economic fundamentals are not modeled. In the macro-finance literature, the focus is on disasters as reflected in economic fundamentals like consumption or GDP, rather than the behavior of stock prices. However, in both literatures, risk premia are specifically attributed to events whose size and probability of occurrence renders a normal distribution essentially impossible.

Motivated by the parallels between these very different literatures, Backus, Chernov, and Martin (2011) study option prices in a consumption disaster model similar to that of Rietz (1988) and Barro (2006). They find, strikingly, that options prices implied by the consumption disaster model are far from their data counterparts. In particular, the implied volatilities resulting from a calibration to macroeconomic consumption data are lower than in the data, and are far more downward sloping as a function of the strike price. Thus options data appear to be inconsistent with consumption disasters as an explanation of the equity premium puzzle.

Like Barro (2006) and Rietz (1988), Backus, Chernov, and Martin (2011) assume that the probability of a disaster occurring is constant. Such a model can explain the equity premium, but cannot account for other features of equity markets, such as the volatility. Recent rare events models therefore introduce dynamics that can account for equity volatility (Gabaix (2012), Gourio (2012), Wachter (2013)). We derive option prices in a model based on that of Wachter (2013), and in a significant generalization of this model that allows for variation in the risk of disaster at different time scales. We show that allowing for a stochastic probability of disaster has dramatic effects on implied volatilities. Namely, rather than being much lower than in the data, the implied volatilities are at about the same level. The slope of the implied volatility curve, rather than being far too great, also matches that of the data. We then apply the model to understanding other features of option prices in the data, such as time-variation in the level and slope of the implied volatility curve. We show

that the model can account for these features of the data as well. In other words, a model calibrated to international consumption data on disasters can explain option prices after all.

We use the fact that our model generalizes the previous literature to better understand the difference in results. Because of time-variation in the disaster probability, the model endogenously produces the stock price changes that occur during normal times and that are reflected in option prices. These changes are absent in iid rare event models, because, during normal times, the volatility of stock returns is equal to the (low) volatility of dividend growth. Moreover, by assuming recursive utility, the model implies a premium for assets that covary negatively with volatility. This makes implied volatilities higher than what they would be otherwise.

Our findings relate to those of Gabaix (2012), who also reports average implied volatilities in a model with rare events.[1] Our model is conceptually different in that we assume recursive utility and time-variation in the probability of a disaster. Gabaix assumes a linearity-generating process (Gabaix (2008)) with a power utility investor. In his calibration, the sensitivity of dividends to changes in consumption is varying; however, the probability of a disaster is not. There are several important implications of our choice of approach. First, the tractability of our framework implies that we can use the same model for pricing options as we do for equities. Moreover, our model nests the simpler one with constant risk of disaster, allowing us to uncover the reason for the large difference in option prices between the models. We also show that the assumption of recursive utility is important for the accounting for the average implied volatility curve. Finally, we go beyond the average implied volatility curve for 3-month options, considering the time series in the level and slope of the implied volatility curve across different maturities. These considerations lead us naturally to a two-factor model for the disaster probability.

[1] Recent work by Nowotny (2011) reports average implied volatilities as well. Nowotny focuses on the implications of self-exciting processes for equity markets rather than on option prices.

Other recent work explores implications for option pricing in dynamic endowment economies. Benzoni, Collin-Dufresne, and Goldstein (2011) derive options prices in a Bansal and Yaron (2004) economy with jumps to the mean and volatility of dividends and consumption. The jump probability can take on two states which are not observable to the agent. Their focus is on the learning dynamics of the states, and the change in option prices before and after the 1987 crash, rather than on matching the shape of the implied volatility curve. Du (2011) examines options prices in a model with external habit formation preferences as in Campbell and Cochrane (1999) in which the endowment is subject to rare disasters that occur with a constant probability. His results illustrate the difficulty of matching implied volatilities assuming either only external habit formation, or a constant probability of a disaster. Also related is the work of Drechsler and Yaron (2011), who focus on the volatility premium and its predictive properties. Our paper differs from these in that we succeed in fitting both the time series and cross-section of implied volatilities, as well as reconciling the macro-finance evidence on rare disasters with option prices.

A related strand of research on endowment economies focuses on uncertainty aversion or exogenous changes in confidence. Drechsler (2012) builds on the work of Bansal and Yaron (2004), but incorporates dynamic uncertainty aversion. He argues that uncertainty aversion is important for matching implied volatilities. Shaliastovich (2009) shows that jumps in confidence can explain option prices when investors are biased toward recency. These papers build on earlier work by Bates (2008) and Liu, Pan, and Wang (2005), who conclude that it is necessary to introduce a separate aversion to crashes to simultaneously account for data on options and on equities. Buraschi and Jiltsov (2006) explain the pattern in implied volatilities using heterogenous beliefs. Unlike these papers, we assume a rational expectations investor with standard (recursive) preferences. The ability of the model to explain implied volatilities arises from time-variation in the probability of a disaster rather than a premium associated with uncertainty.

Finally, a growing line of work focuses on the links between macroeconomic events, index

option prices, and risk premia. Brunnermeier, Nagel, and Pedersen (2008) links option prices to the risk of currency crashes, Kelly, Lustig, and Van Nieuwerburgh (2012) use options to infer a premium for too-big-to-fail financial institutions, Gao and Song (2013) price crash risk in the cross-section using options, and Kelly, Pastor, and Veronesi (2014) demonstrate a link between options and political risk. The focus of these papers is empirical. Our model provides a model that can help in interpreting these results.

The remainder of this paper is organized as follows. Section 1.2 discusses a single-factor model for stochastic disaster risk (SDR). This model turns out to be sufficient to explain why stochastic disaster risk can explain the level and slope of implied volatilities, as explained in Section 1.3, which also compares the implications of the single-factor SDR model to the implications of a constant disaster risk (CDR) model. As we explain in Section 1.4, however, the single-factor model cannot explain many interesting features of options data. In this section, we explore a multi-factor extension, which we fit to the data in Section 1.5. We show that the multi-factor model can explain the time-series variation in the implied volatility slope. Section 1.6 concludes.

1.2. Option prices in a single-factor stochastic disaster risk model

1.2.1. Assumptions

In this section we describe a model with stochastic disaster risk (SDR). We assume an endowment economy with complete markets and an infinitely-lived representative agent. Aggregate consumption (the endowment) solves the following stochastic differential equation

$$dC_t = \mu C_{t^-}\, dt + \sigma C_{t^-}\, dB_t + (e^{Z_t} - 1) C_{t^-}\, dN_t, \tag{1.1}$$

where B_t is a standard Brownian motion and N_t is a Poisson process with time-varying intensity λ_t. This intensity follows the process

$$d\lambda_t = \kappa(\bar{\lambda} - \lambda_t)\, dt + \sigma_\lambda \sqrt{\lambda_t}\, dB_{\lambda,t}, \tag{1.2}$$

where $B_{\lambda,t}$ is also a standard Brownian motion, and B_t, $B_{\lambda,t}$ and N_t are assumed to be independent. For the range of parameter values we consider, λ_t is small and can therefore be interpreted to be (approximately) the probability of a jump. We thus will use the terminology probability and intensity interchangeably, while keeping in mind the that the relation is an approximate one.

The size of a jump, provided that a jump occurs, is determined by Z_t. We assume Z_t is a random variable whose time-invariant distribution ν is independent of N_t, B_t and $B_{\lambda,t}$. We will use the notation E_ν to denote expectations of functions of Z_t taken with respect to the ν-distribution. The t subscript on Z_t will be omitted when not essential for clarity.

We will assume a recursive generalization of power utility that allows for preferences over the timing of the resolution of uncertainty. Our formulation comes from Duffie and Epstein (1992), and we consider a special case in which the parameter that is often interpreted as the elasticity of intertemporal substitution (EIS) is equal to 1. That is, we define continuation utility V_t for the representative agent using the following recursion:

$$V_t = E_t \int_t^\infty f(C_s, V_s)\, ds, \tag{1.3}$$

where

$$f(C, V) = \beta(1-\gamma)V\left(\log C - \frac{1}{1-\gamma}\log((1-\gamma)V)\right). \tag{1.4}$$

The parameter β is the rate of time preference. We follow common practice in interpreting γ as relative risk aversion. This utility function is equivalent to the continuous-time limit (and the limit as the EIS approaches one) of the utility function defined by Epstein and Zin (1989) and Weil (1990).

1.2.2. Solving for asset prices

We will solve for asset prices using the state-price density, π_t.[2] Duffie and Skiadas (1994) characterize the state-price density as

$$\pi_t = \exp\left\{\int_0^t \frac{\partial}{\partial V} f\left(C_s, V_s\right) ds\right\} \frac{\partial}{\partial C} f\left(C_t, V_t\right). \tag{1.5}$$

There is an equilibrium relation between utility V_t, consumption C_t and the disaster probability λ_t. Namely,

$$V_t = \frac{C_t^{1-\gamma}}{1-\gamma} e^{a+b\lambda_t},$$

where a and b are constants given by

$$a = \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + b\frac{\kappa\bar{\lambda}}{\beta} \tag{1.6}$$

$$b = \frac{\kappa+\beta}{\sigma_\lambda^2} - \sqrt{\left(\frac{\kappa+\beta}{\sigma_\lambda^2}\right)^2 - 2\frac{E_\nu\left[e^{(1-\gamma)Z}-1\right]}{\sigma_\lambda^2}}. \tag{1.7}$$

It follows that

$$\pi_t = \exp\left(\eta t - \beta b\int_0^t \lambda_s ds\right)\beta C_t^{-\gamma} e^{a+b\lambda_t}, \tag{1.8}$$

where $\eta = -\beta(a+1)$. Details are provided in Appendix A1.2.1.

Following Backus, Chernov, and Martin (2011), we assume a simple relation between dividends and consumption: $D_t = C_t^\phi$, for leverage parameter ϕ.[3] Let $F(D_t, \lambda_t)$ be the value of the aggregate market (it will be apparent in what follows that F is a function of D_t and λ_t). It follows from no-arbitrage that

$$F(D_t, \lambda_t) = E_t\left[\int_t^\infty \frac{\pi_s}{\pi_t} D_s\, ds\right].$$

[2]Other work on solving for equilibria in continuous-time models with recursive utility includes Benzoni, Collin-Dufresne, and Goldstein (2011), Eraker and Shaliastovich (2008), Fisher and Gilles (1999) and Schroder and Skiadas (1999).

[3]This implies that dividends respond more than consumption to disasters, an assumption that is plausible given the U.S. data (Longstaff and Piazzesi (2004)).

The stock price can be written explicitly as

$$F(D_t, \lambda_t) = D_t G(\lambda_t), \tag{1.9}$$

where the price-dividend ratio G is given by

$$G(\lambda_t) = \int_0^\infty \exp\left\{a_\phi(\tau) + b_\phi(\tau)\lambda_t\right\}$$

for functions $a_\phi(\tau)$ and $b_\phi(\tau)$ given by:

$$\begin{aligned} a_\phi(\tau) &= \left(\mu_D - \mu - \beta + \gamma\sigma^2(1-\phi) - \frac{\kappa\bar{\lambda}}{\sigma_\lambda^2}(\zeta_\phi + b\sigma_\lambda^2 - \kappa)\right)\tau \\ &\qquad - \frac{2\kappa\bar{\lambda}}{\sigma_\lambda^2}\log\left(\frac{(\zeta_\phi + b\sigma_\lambda^2 - \kappa)\left(e^{-\zeta_\phi\tau} - 1\right) + 2\zeta_\phi}{2\zeta_\phi}\right) \\ b_\phi(\tau) &= \frac{2E_\nu\left[e^{(1-\gamma)Z} - e^{(\phi-\gamma)Z}\right]\left(1 - e^{-\zeta_\phi\tau}\right)}{(\zeta_\phi + b\sigma_\lambda^2 - \kappa)\left(1 - e^{-\zeta_\phi\tau}\right) - 2\zeta_\phi}, \end{aligned}$$

where

$$\zeta_\phi = \sqrt{\left(b\sigma_\lambda^2 - \kappa\right)^2 + 2E_\nu\left[e^{(1-\gamma)Z} - e^{(\phi-\gamma)Z}\right]\sigma_\lambda^2}$$

(see Wachter (2013)). We will often use the abbreviation $F_t = F(D_t, \lambda_t)$ to denote the value of the stock market index at time t.

1.2.3. Solving for implied volatilities

Let $P(F_t, \lambda_t, \tau; K)$ denote the time-t price of a European put option on the stock market index with strike price K and expiration $t+\tau$. For simplicity, we will abbreviate the formula for the price of the dividend claim as $F_t = F(D_t, \lambda_t)$. Because the payoff on this option at expiration is $(K - F_{t+\tau})^+$, it follows from the absence of arbitrage that

$$P(F_t, \lambda_t, T-t; K) = E_t\left[\frac{\pi_T}{\pi_t}\left(K - F_T\right)^+\right].$$

Let $K^n = K/F_t$, the normalized strike price (or "moneyness"), and define

$$P^n(\lambda_t, T-t; K^n) = E_t\left[\frac{\pi_T}{\pi_t}\left(K^n - \frac{F_T}{F_t}\right)^+\right]. \tag{1.10}$$

We will establish below that P^n is indeed a function of λ_t, time to expiration and moneyness alone Clearly $P_t^n = P_t/F_t$. Because our ultimate interest is in implied volatilities, and because, in the formula of Black and Scholes (1973), normalized option prices are functions of the normalized strike price (and the volatility, interest rate and time to maturity), it suffices to calculate P_t^n.[4]

Returning to the formula for P_t^n, we note that, from (1.9), it follows that

$$\frac{F_T}{F_t} = \frac{D_T}{D_t}\frac{G(\lambda_T)}{G(\lambda_t)}. \tag{1.11}$$

Moreover, it follows from (1.8) that

$$\frac{\pi_T}{\pi_t} = \left(\frac{C_T}{C_t}\right)^{-\gamma} \exp\left\{\int_t^T (\eta - \beta b\lambda_s)\, ds + b(\lambda_T - \lambda_t)\right\}. \tag{1.12}$$

At time t, λ_t is sufficient to determine the distributions of consumption and dividend growth between t and T, as well as the distribution of λ_s for $s = t, \ldots, T$. It follows that normalized put prices (and therefore implied volatilities) are a function of λ_t, the time to expiration, and moneyness.

[4]Given stock price F, strike price K, time to maturity $T-t$, interest rate r, and dividend yield y, the Black-Scholes put price is defined as

$$\mathrm{BSP}(F, K, T-t, r, y, \sigma) = e^{-r(T-t)}KN(-d_2) - e^{-y(T-t)}FN(-d_1)$$

where

$$d_1 = \frac{\log(F/K) + \left(r - y + \sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T-t}$$

Given the put prices calculated from the transform analysis, inversion of this Black-Scholes formula gives us implied volatilities. Specifically, the implied volatility $\sigma_t^{\mathrm{imp}} = \sigma^{\mathrm{imp}}(\lambda_t, T-t; K^n)$ solves

$$P_t^n(\lambda_t, T-t; K^n) = \mathrm{BSP}\left(1, K^n, T-t, r_t^b, 1/G(\lambda_t), \sigma_t^{\mathrm{imp}}\right)$$

where r_t^b is the model's analogue of the Treasury Bill rate, which allows for a probability of a default in case of a disaster (see Barro (2006); as in that paper we assume a default rate of 0.4).

Appendix A1.2.5 describes the calculation of (1.10). We first approximate the price-dividend ratio $G(\lambda_t)$ by a log-linear function of λ_t, As the Appendix describes, this approximation is highly accurate. We can then apply the transform analysis of Duffie, Pan, and Singleton (2000) to calculate put prices.

The implied volatility curve in the data represents an average of implied volatilities at different points in time. We follow the same procedure in the model, calculating an unconditional average implied volatility curve. To do so, we first solve for the implied volatility as a function of λ_t. We numerically integrate this function over the stationary distribution of λ_t. This stationary distribution is Gamma with shape parameter $2\kappa\bar{\lambda}/\sigma_\lambda^2$ and scale parameter $\sigma_\lambda^2/(2\kappa)$ (Cox, Ingersoll, and Ross (1985)).

1.2.4. The constant disaster risk model

Taking limits in the above model as σ_λ approaches zero implies a model with a constant probability of disaster (Appendix A1.2.2 shows that this limit is indeed well-defined and is what would be computed if one were to solve the constant disaster risk model from first principles). We use this model to evaluate the role that stochastic disaster risk plays in the model's ability to match the implied volatility data. We refer to this model in what follows as the CDR (constant disaster risk) model, to distinguish it from the more general SDR model.

The CDR model is particularly useful in reconciling our results with those of Backus, Chernov, and Martin (2011). Backus et al. solve a model with a constant probability of a jump in consumption, calibrated in a manner similar to Barro (2006). They call this the consumption-based model, to distinguish it from the reduced-form options-based model which is calibrated to fit options data. While Backus et al. assume power utility, their model can be rewritten as one with recursive utility with an EIS of one. The reason is that the endowment process is iid. In this special case, the EIS and the discount rate are not separately identified. Specifically, a power utility model has an identical stochastic discount

factor, and therefore identical asset prices to a recursive utility model with arbitrary EIS as long as one can adjust the discount rate (Appendix A1.2.3).[5]

1.3. Average implied volatilities in the model and in the data

In this section we compare implied volatilities in the two versions of the model we have discussed with implied volatilities in the data. For now, we focus on three-month options.

Table 1.1 shows the parameter values for the SDR and the benchmark CDR model. The parameters in the SDR model are identical to those of Wachter (2013), and thus do not make use of options data. The parameters for the benchmark CDR model are as in the consumption-based model of Backus, Chernov, and Martin (2011). That is, we consider a calibration that is isomorphic to that of Backus, Chernov, and Martin (2011) in which the EIS is equal to 1. Assuming a riskfree rate of 2% (as in Backus et al.), we calculate a discount rate of 0.0189 . We could also use the same parameters as in the SDR model except with $\sigma_\lambda = 0$; the results are very similar.

The two calibrations differ in their relative risk aversion, in the volatility of normal-times consumption growth, in leverage, in the probability of a disaster, and of course in whether the probability is time-varying. The net effect of some of these differences turns out to be less important than what one may think: for example, higher risk aversion and lower disaster probability roughly offset each other. We explore the implications of leverage and volatility in what follows.

The two models also assume different disaster distributions. For the SDR model, the disaster distribution is multinomial, and taken from Barro and Ursua (2008) based on actual consumption declines. The benchmark CDR model assumes that consumption declines are

[5]Backus et al. define option payoffs not in terms of the price, but in terms of the total return. In principal one could correct for this; however the correction requires a well-defined price-dividend ratio. For their parameters, and a riskfree rate of 2%, the price-dividend ratio is not well-defined (the value of the infinitely-lived dividend-claim does not converge). One can assess the difference between options on returns and the usual option definition with higher levels of the riskfree rate, which does lead to well-defined prices. It is small for ATM and OTM options which are the focus of our (and their) study. In what follows, for consistency with their study, we use their definition and pricing method when reporting results for the benchmark CDR calibration.

log-normal. For comparison, we plot the smoothed density for the SDR model along with the density of the consumption-based model in Figure 1.1. Compared with the lognormal model, the SDR model has more mass over small declines in the 10–20% region, and more mass over large declines in the 50-70% region.[6]

Figure 1.2 shows the resulting implied volatilities as a function of moneyness, as well as implied volatilities in the data. Confirming previous results, we find that the CDR model leads to implied volatilities that are dramatically different from those in the data. First, the implied volatilities are too low, even though the model was calibrated to match the volatility of equity returns. Second, they exhibit a strong downward slope as a function of the strike price. While there is a downward slope in the data, it is not nearly as large. As a result, implied volatilities for at-the-money (ATM) options in the CDR model are less than 10%, far below the option-based implied volatilities, which are over 20%.

In contrast, the SDR model can explain both ATM and OTM (out-of-the-money) implied volatilities. For OTM options (with moneyness equal to 0.94), the SDR model gives an implied volatility of 23%, close to the data value of 24%. There is a downward slope, just as in the data, but it is much smaller than that of the CDR model. ATM options have implied volatilities of about 21% in both the model and the data. There are a number of differences between this model and the CDR model. We now discuss which of these differences is primarily responsible for the change in implied volatilities.

1.3.1. The role of leverage

In their discussion, Backus, Chernov, and Martin (2011) emphasize the role of very bad consumption realizations as a reason for the poor performance of the disaster model. Therefore, this seems like an appropriate place to start. The disaster distribution in the SDR bench-

[6]One concern is the sensitivity of our results to behavior in the tails of the distribution. By assuming a multinomial distribution, we essentially assume that this distribution is bounded, which is probably not realistic. However, it makes very little difference if we consider a unbounded distribution that matches the observations in the data. Barro and Jin (2011) suggest this can be done with a power law distribution with tail parameter of about 6.5. We have tried this version of the model and the results are virtually indistinguishable. The reason is that, even though low realizations are possible in theory, their probability is so small as to not affect the model's results.

mark actually implies a slightly higher probability of extreme events than the benchmark CDR model (Figure 1.1). However, the benchmark CDR model has much higher leverage: the leverage parameter is 5.1 for the CDR calibration versus 2.6 for the SDR calibration. Leverage does not affect consumption but it affects dividends, and therefore stock and option prices. A higher leverage parameter implies that dividends will fall further in the event of a consumption disaster. It is reasonable, therefore, to attribute the difference in the implied volatilities to the difference in the leverage parameter.

Figure 1.3 tests this directly by showing option prices in the CDR model for leverage of 5.1 and for leverage of 2.6 (denoted "lower leverage") in the figure. Surprisingly, the slope for the calibration with leverage of 2.6 is slightly higher than the slope for leverage of 5.1. Lowering leverage results in a downward shift in the level of the implied volatility curve, not the slope. Thus the difference in leverage cannot be the explanation for why the slope in our model is lower than the slope for CDR.

Why does the change in leverage result in a shift in the level of the curve? It turns out that in the CDR model, changing normal-times volatility has a large effect. Leverage affects both the disaster distribution and normal-times volatility. Lowering leverage has a large effect on normal-times volatility and thus at-the-money options. This is why the level of the curve is lower, and the slope is slightly steeper.[7]

1.3.2. The role of normal shocks

To further consider the role of normal-times volatility, we explore the impact of changing the consumption volatility parameter σ. In the benchmark CDR comparison, consumption volatility is equal to the value of consumption volatility over the 1889–2009 sample, namely 3.5%. Most of this volatility is accounted for by the disaster distribution, because, while the disasters are rare, they are severe. Therefore normal-times volatility is 1%, lower than the U.S. consumption volatility over the post-war period. The SDR model is calibrated differ-

[7] Yan (2011) shows analytically that, as the time to expiration approaches zero, the implied volatility is equal to the normal-times volatility in the stock price, while the slope is inversely related to the normal-times volatility of the stock price.

ently; following Barro (2006), the disaster distribution is determined based on international macroeconomic data, and the normal-times distribution is set to match postwar volatility in developed countries. The resulting normal-times volatility is 2%. To evaluate the effect of this difference, we solve for implied volatilities in the CDR model with leverage of 5.1 and normal-times volatility of 2%. In Figure 1.3, the result is shown in the line denoted "higher normal-times volatility."

As Figure 1.3 shows, increasing the normal-times volatility of consumption growth in the CDR model has a noticeable effect on implied volatilities: The implied volatility curve is higher and flatter. The change in the level reflects the greater overall volatility. The change in the slope reflects the greater probability of small, negative outcomes. However, the effect, while substantial, is not nearly large enough to explain the full difference. The level of the "higher normal-times volatility" smile is still too low and the slope is too high compared with the data.[8]

While raising the volatility of consumption makes the CDR model look somewhat more like the SDR model (though it does not account for the full difference), it is not the case that lowering the volatility of consumption makes the SDR model more like the CDR model. Namely, reducing σ to 1% (which would imply a normal-times consumption volatility that is lower than in the post-war data) has almost no effect on the implied volatility curve of the SDR model. There are two reasons why this parameter affects implied volatilities differently in the two cases. First, the leverage parameter is much lower in the SDR model than the CDR model. Second, volatility in the SDR model comes from time-variation in discount rates (driven by λ_t) as well as in payouts ($\phi\sigma$). The first of these terms is much larger than the second.[9]

[8]Note further that leverage of 5.1, combined with a normal-times consumption volatility of 2% means that normal-times dividend volatility of dividends is 10.2%. However, annual volatility in postwar data is only 6.5%.

[9]To be precise, total return volatility in the SDR model equals the square root of the variance due to λ_t, plus the variance in dividends. Dividend variance is small, and it is added to something much larger to determine total variance. Thus the effect of dividend volatility on return volatility is very small, and changes in dividend volatility also have relatively little effect.

1.3.3. The price of volatility risk

One obvious difference between the CDR model and the SDR model is that the SDR model is dynamic. Because of recursive utility, this affects risk premia on options and therefore option prices and implied volatilities. As shown in Section 1.2.2, the state price density depends on the probability of disaster. Thus risk premia depend on covariances with this probability: assets that increase in price when the probability rises will be a hedge. Options are such an asset. Indeed, an increase in the probability of a rare disaster raises option prices, while at the same time increasing marginal utility.

To directly assess the magnitude of this effect, we solve for option prices using the same process for the stock price and the dividend yield, but with a pricing kernel adjusted to set the above effect equal to zero. Risk premia in the model arise from covariances with the pricing kernel. We replace the pricing kernel in (1.8) with one in which $b = 0$.[10] Because b determines the risk premium due to covariance with λ_t, setting $b = 0$ will shut off this effect. Figure 1.4 shows, setting $b = 0$ noticeably reduces the level of implied volatilities, though the difference between the $b = 0$ model and the SDR is small compared to the difference between the CDR and the SDR model.

It is worth emphasizing that the assumption of $b = 0$ does not imply an iid model. This model still assumes that stock prices are driven by stochastic disaster risk; otherwise the volatility of stock returns would be equal to that of dividends.

1.3.4. The distribution of consumption growth implied by options

Our results show that a model with stochastic disaster risk can fit implied volatilities, thereby addressing one issue raised by Backus, Chernov, and Martin (2011). Backus et

[10]Note that a and η also depend on b: these expressions are also changed in the experiment. While it may first appear that b should also affect the riskfree rate, this does not occur in the model with EIS= 1. The riskfree rate satisfies a simple expression

$$r_t = \beta + \mu - \gamma\sigma^2 + \lambda_t E_\nu \left[e^{-\gamma Z} \left(e^Z - 1 \right) \right].$$

al. raise a second issue: assuming power utility and iid consumption growth, they back out a distribution for the left tail of consumption growth from option prices (we will call this the "option-implied consumption distribution").[11] Based on this distribution, they conclude that the probabilities of negative jumps to consumption are much larger, and the magnitudes much smaller, than implied by the international macroeconomic data used by Barro (2006) and Barro and Ursua (2008).

The resolution of this second issue is clearly related to the first. For if a model (like the one we describe) can explain average implied volatilities while assuming a disaster distribution from macroeconomic data, then it follows that this macroeconomic disaster distribution is one possible consumption distribution that is consistent with the implied volatility curve. Namely, the inconsistency between the extreme consumption events in the macroeconomic data and option prices is resolved by relaxing the iid assumption.

Of course, this reasoning does not imply that the stochastic disaster model is any better than the iid model with the option-implied consumption distribution. This distribution is, after all, consistent with option prices, the equity premium, and the mean and volatility of consumption growth observed in the U.S. in the 1889-2009 period (provided a coefficient of relative risk aversion equal to 8.7). However, it turns out that this consumption distribution can be ruled out based on other data: because it assumes that negative consumption jumps are relatively frequent (as they must be to explain the equity premium), some would have occurred in the 60-year postwar period in the U.S. The unconditional volatility of consumption growth in the U.S. during this period was less than 2%. Under the option-implied consumption growth distribution, there is less than a 1 in one million chance of observing a 60-year period with volatility this low.[12]

[11] A methodological problem with this analysis is that it assumes that options are on returns rather than on prices. See footnote 5.

[12] There are multiple additional objections to an iid model for returns. Another that arises in the context of options and rare disaster is that of Neuberger (2012), who shows that an iid model is unlikely based on the lack of decay in return skewness as the measurement horizon grows. We discuss this result further in Section 1.5.4.

1.3.5. Summary

A consequence of stochastic disaster risk is high stock market volatility, not just during occurrences of disasters, but during normal periods as well. This is reflected in the level and shallowness of the volatility smile: while the existence of disasters leads to an upward slope for out-of-the money put options, high normal-period volatility implies that the level is high for put options that are in the money or only slightly out of the money. The same mechanism, and indeed the same parameters that allow the model to match the level of realized stock returns enable the model to match implied volatilities.

Previous work suggests that allowing for stochastic volatility (and time-varying moments more generally) does not appear to affect the shape of the implied volatility curve.[13] How is it, then, that this paper comes to such a different conclusion? The reason may arise from the fact that the previous literature mainly focused on reduced-form models, in which the jump dynamics and volatility of stock returns are freely chosen. However, in an equilibrium model like the present one, stock market volatility arises endogenously from the interplay between consumption and dividend dynamics and agents' preferences. While it is possible to match the unconditional volatility of stock returns and consumption in an iid model, this can only be done (given the observed data) by having all of the volatility occur during disasters. In such a model it is not possible to generate sufficient stock market volatility in normal times to match either implied or realized volatilities. While in the reduced-form literature, the difference between iid and dynamic models principally affects the conditional second moments, in the equilibrium literature, the difference affects the level of volatility itself.

[13] Backus, Chernov, and Martin (2011) write: "The question is whether the kinds of time dependence we see in asset prices are quantitatively important in assessing the role of extreme events. It is hard to make a definitive statement without knowing the precise form of time dependence, but there is good reason to think its impact could be small. The leading example in this context is stochastic volatility, a central feature of the option-pricing model estimated by Broadie et al. (2007). However, average implied volatility smiles from this model are very close to those from an iid model in which the variance is set equal to its mean. Furthermore, stochastic volatility has little impact on the probabilities of tail events, which is our interest here."

1.4. Option prices in a multi-factor stochastic disaster risk model

1.4.1. Why multiple factors?

The previous sections show how introducing time variation into conditional moments can substantially alter the implications of rare disasters for implied volatilities. The model presented there was parsimonious, with a single state variable following a square root process.

Closer examination of our results suggests an aspect of options data that may be difficult to fit to this model. Figure 1.5 shows implied volatilities for λ_t equal to the median and for the 1st and 99th percentile value for put options with moneyness as low as 0.85.[14] Implied volatilities increase almost in parallel as λ_t increases. That is, ATM options are affected by an increase in the rare disaster probability almost as much as out-of-the-money options. The model therefore implies that there should be little variation in the slope of the implied volatility curve.

Figure 1.6 shows the historical time series of implied volatilities computed on one month ATM and OTM options with moneyness of 0.85. Panel C shows the difference in the implied probabilities, a measure of the slope of the implied volatility curve. Defined in this way, the average slope is 12%, with a volatility of 2%. Moreover, the slope can rise as high as 18% and fall as low as 6%. While the SDR model can explain the average slope, it seems unlikely that it would be able to account for the time-variation in the slope, at least under the current calibration. Moreover, comparing Panel C with Panels A and B of Figure 1.6 indicates that the slope varies independently of the level of implied volatilities. Thus it is unlikely that any model with a single state variable could account for these data.

The mechanism in the SDR model that causes time-variation in rare disaster probabilities is identical to the mechanism that leads to volatility in normal times. Namely, when λ_t

[14]The figure also shows 99th and 1st percentile implied volatilities in the data at each moneyness level. While the 99th percentile values are high in the model, they are similarly high in the data; thus the single-factor model can accurately account for the range in implied volatilities. Because this graph is constructed looking at implied volatilities for each moneyness level, it does not answer the important question of whether the model can explain time-variation in the slope.

is high, rare disasters are more likely *and* returns are more volatile. In order to account for the data, a model must somehow decouple the volatility of stock returns from the probability of rare events. This is challenging, because volatility endogenously depends on the probability of rare events. Indeed, the main motivation for assuming time-variation in the probability of rare events is to generate volatility in stock returns that seems otherwise puzzling. Developing such a model is our goal in this section of the paper.

1.4.2. Model assumptions

In this section, we introduce a mechanism that decouples the volatility of stock returns from the probability of rare events. We assume the same stochastic process for consumption (1.1) and for dividends. We now assume, however, that the probability of a rare event follows the process

$$d\lambda_t = \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t}, \tag{1.13}$$

where ξ_t solves the following stochastic differential equation

$$d\xi_t = \kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t}. \tag{1.14}$$

We continue to assume that all Brownian motions are mutually independent. The process for λ_t takes the same form as before, but instead of reverting to a constant value $\bar{\lambda}$, λ_t reverts to a value that is itself stochastic. We will assume that this value, ξ_t itself follows a square root process. Though the relative values of κ_λ and κ_ξ do not matter for the form of the solution, to enable an easier interpretation of these state variables, we will choose parameters so that shocks to ξ_t die out more slowly than direct shocks to λ_t. Duffie, Pan, and Singleton (2000) use the process in (1.13) and (1.14) to model return volatility. This model is also related to multi-factor return volatility processes proposed by Bates (2000), Gallant, Hsu, and Tauchen (1999), and Andersen, Fusari, and Todorov (2013).[15] In

[15]The purpose of this model is not to provide a better fit than existing multi-factor reduced form models; all else equal it will be easier for a reduced-form model which has fewer restrictions to fit the data than an equilibrium model.

our study, the process is for jump intensity rather than volatility; however, volatility, which arises endogenously, will inherit the two-factor structure. While we implement a model with two-factors in this paper, our solution method is general enough to encompass an arbitrary number of factors with linear dependencies through the drift term, as in the literature on the term structure of interest rates (Dai and Singleton (2002)).[16]

Like the one-factor SDR model, our two-factor extension is highly tractable. In Appendix A1.3.1, we show that utility is given by

$$V_t = \frac{C_t^{1-\gamma}}{1-\gamma} e^{a+b_\lambda \lambda_t + b_\xi \xi_t} \tag{1.15}$$

where

$$a = \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + \frac{b_\xi \kappa_\xi \bar{\xi}}{\beta} \tag{1.16}$$

$$b_\lambda = \frac{\kappa_\lambda + \beta}{\sigma_\lambda^2} - \sqrt{\left(\frac{\kappa_\lambda + \beta}{\sigma_\lambda^2}\right)^2 - 2\frac{E_\nu\left[e^{(1-\gamma)Z_t} - 1\right]}{\sigma_\lambda^2}} \tag{1.17}$$

$$b_\xi = \frac{\kappa_\xi + \beta}{\sigma_\xi^2} - \sqrt{\left(\frac{\kappa_\xi + \beta}{\sigma_\xi^2}\right)^2 - 2\frac{b_\lambda \kappa_\lambda}{\sigma_\xi^2}} \tag{1.18}$$

Equation 1.5 still holds for the state-price density, though the process V_t will differ. The riskfree rate and government bill rate are the same functions of λ_t as in the one-factor model. In Appendix A1.3.3, we show that the price-dividend ratio is given by

$$G(\lambda_t, \xi_t) = \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda_t + b_{\phi\xi}(\tau)\xi_t\right), \tag{1.19}$$

[16] Equilibrium models that highlight the importance of variation across different time scales include Calvet and Fisher (2007) and Zhou and Zhu (2014). These papers do not discuss the fit to implied volatilities. Rather than rare disasters, these papers induce fluctuations in asset prices based on time-varying first and second moments of consumption growth; evidence from consumption growth itself suggests that this variation is insufficient to produce the observed volatility in stock prices (Beeler and Campbell (2012)). These papers require simultaneously high risk aversion and a high elasticity of intertemporal substitution; this model requires neither.

where a_ϕ, $b_{\phi\lambda}$ and $b_{\phi\xi}$ solve the differential equations

$$\begin{aligned} a'_\phi(\tau) &= -\beta - \mu - \gamma(\phi-1)\sigma^2 + \mu_D + b_{\phi\xi}(\tau)\kappa_\xi\bar{\xi} \\ b'_{\phi\lambda}(\tau) &= -b_{\phi\lambda}(\tau)\kappa_\lambda + \frac{1}{2}b_{\phi\lambda}(\tau)^2\sigma_\lambda^2 + b_\lambda b_{\phi\lambda}(\tau)\sigma_\lambda^2 + E_\nu\left[e^{(\phi-\gamma)Z_t} - e^{(1-\gamma)Z_t}\right] \\ b'_{\phi\xi}(\tau) &= -b_{\phi\lambda}(\tau)\kappa_\lambda - b_{\phi\xi}(\tau)\kappa_\xi + \frac{1}{2}b_{\phi\xi}(\tau)^2\sigma_\xi^2 + b_\xi b_{\phi\xi}(\tau)\sigma_\xi^2, \end{aligned}$$

with boundary condition

$$a_\phi(0) = b_{\phi\lambda}(0) = b_{\phi\xi}(0) = 0.$$

Similar reasoning to that of Section 1.2.3 shows that for a given moneyness and time to expiration, normalized option prices and implied volatilities are a function of λ_t and ξ_t alone. To solve for option prices, we approximate $G(\lambda_t, \xi_t)$ by a log-linear function of λ_t and ξ_t, as shown in Appendix A1.3.3.

1.5. Fitting the multifactor model to the data

1.5.1. Parameter choices

For simplicity, we keep risk aversion γ, the discount rate β and the leverage parameter ϕ the same as in the one-factor model. We also keep the distribution of consumption in the event of a disaster the same. Note that κ_λ and σ_λ will not have the same interpretation in the two-factor model as κ and σ_λ do in the one-factor model.

Our first goal in calibrating this new model is to generate reasonable predictions for the aggregate market and for the consumption distribution. That is, we do not want to allow the probability of a disaster to become too high. One challenge in calibrating representative agent models is to match the high volatility of the price-dividend ratio. In the two-factor model, as in the one-factor model, there is an upper limit to the amount of volatility that can be assumed in the state variable before a solution for utility fails to exist. The more persistent the processes, namely the lower the values of κ_λ and κ_ξ, the lower the

respective volatilities must be so as to ensure that the discriminants in (1.17) and (1.18) stay nonnegative. We choose parameters so that the discrimant is equal to zero; thus there is only one more free parameter relative to to the one-factor model.

The resulting parameter choices are shown in Table 1.2. The mean reversion parameter κ_λ and volatility parameter σ_λ are relatively high, indicating a fast-moving component to the λ_t process, while the mean reversion parameter κ_ξ and σ_ξ are relatively low, indicating a slower-moving component. The parameter $\bar{\xi}$ (which represents both the average value of ξ_t and the average value of λ_t) is 2% per annum. This is lower than $\bar{\lambda}$ in our calibration of the one-factor model. In this sense, the two-factor calibration is more conservative. However, the extra persistence created by the ξ_t process implies that λ_t could deviate from its average for long periods of time. To clarify the implications of these parameter choices, we report population statistics on λ_t in Panel C of Table 1.2. The median disaster probability is only 0.37%, indicating a highly skewed distribution. The standard deviation is 3.9% and the monthly first-order autocorrelation is 0.9858.

Implications for the riskfree rate and the market are shown in Table 1.3. We simulate 100,000 samples of length 60 years to capture features of the small-sample distribution. We also simulate a long sample of 600,000 years to capture the population distribution. Statistics are reported for the full set of 100,000 samples, and the subset for which there are no disasters (38% of the sample paths). The table reveals a good fit to the equity premium and to return volatility. The average Treasury Bill rate is slightly too high, though this could be lowered by lowering β or by lowering the probability of government default.[17] The model successfully captures the low volatility of the riskfree rate in the postwar period. The median value of the price-dividend ratio volatility is lower than in the data (0.27 versus 0.43), but the data value is still lower than the 95th percentile in the simulated sample.[18] For the market moments, only the very high AR(1) coefficient in postwar data falls outside

[17] As in the one-factor model, we assume a 40% probability of government default.

[18] The one-factor model, which was calibrated to match the population persistence of the price-dividend ratio, has a median price-dividend ratio volatility of 0.21 for sample paths without disasters and a population price-dividend ratio volatility of 0.38.

the 90% confidence bounds: it is 0.92 (annual), while in the data, the median is 0.79 and the 95th percentile value is 0.91. As we will show below, there is a tension between matching the autocorrelation in the price-dividend ratio and in option prices. Moreover, so that utility converges, there is a tradeoff between persistence and volatility. One view is that the autocorrelation of the price-dividend ratio observed in the postwar period may in fact have been very exceptional and perhaps is not a moment that should be targeted too stringently.

1.5.2. Simulation results for the two-factor model

We first examine the fit of the model to the mean of implied volatilities in the data. We ask more of the two-factor model than its one-factor counterpart. Rather than looking only at 3-month options across a narrow moneyness range, we extend the range to options of moneyness of 0.85. We also look at 1- and 6-month options, and at moments of implied volatilities beyond the means. Moreover, rather than looking only at the population average, we consider the range of values we would see in repeated samples that resemble the data, namely, samples of length 17 years with no disaster. This is a similar exercise to what was performed in Table 1.3, though calculating option prices is technically more difficult than calculating equity prices.[19]

Figure 1.7 shows means and volatilities of implied volatilities for the three option maturities. We report the averages across each sample path, as well as 90% confidence intervals from the simulation. We see that this new model is successful at matching the average level of the implied volatility curve for all three maturities, even with this extended moneyness range, and even though we are looking at sample paths in which the disaster probability will be lower than average. In fact the slope in the model is slightly below that in the data. Similarly, the model's predictions for volatility of volatility are well within the standard error bars for all moneyness levels and for all three option maturities.

[19]Because of the extra persistence in this multifrequency model, it is important to accurately capture the dynamics when λ_t is near zero. We therefore simulate the model at a daily interval for 17 years. This simulation is repeated 1000 times for the options calculation, and more for the (easier) equity calculations. Along each simulation path, we pick monthly observations of the state variables and calculate option prices for these monthly observations.

One issue that arises in fitting both options and equities with a single model is the very different levels of persistence in the option and equity markets. This tension is apparent in our simulated data as well. As Table 1.3 reports, the annual AR(1) coefficient for the price-dividend ratio in the data is extremely high: 0.92; just outside of our 10% confidence intervals. The median value from the simulations is still a very high 0.79; in monthly simulations, this value is 0.98. Implied volatilities in simulated data have much lower autocorrelations. Median autocorrelations are roughly the same across moneyness levels, and are in the 0.92 to 0.94 range; substantially below the level for the price-dividend ratio. The AR(1) coefficients in the data are lower still, though generally within the 10% confidence intervals.[20] While the same two factors drive equity and option prices, they do so to different extents. The model endogenously captures the greater persistence in equity prices, which represent value in the longer run than do option prices.

1.5.3. Implications for the time series

We now return to the question of whether the two-factor model can explain time-variation in the slope and the level of option prices. Before embarking on this exercise we note that the matching the time series is not usually a target for general equilibrium models because these models operate under tight constraints. We expect that there will be some aspects of the time series that our model will not be able to match.

We consider the time series of one-month ATM and OTM implied volatilities (Figure 1.6). For each of these data points, we compute the implied value of λ_t and ξ_t. Note that this exercise would not be possible if the model were not capable of simultaneously matching the level and slope of the implied volatility curve for one-month options. We show the resulting values in Figure 1.8. For most of the sample period, the disaster probability λ_t varies between 0 and 6%, with spikes corresponding to the Asian financial crisis in the late 1990s and the large market declines in the early 2000s. However, the sample is clearly dominated

[20] It is well known that the implied volatility series exhibits long-memory-like properties. Thus the first-order autocorrelation of volatility may understate the true persistence of the data.

by the events of 2008-2009, in which the disaster probability rises to about 20%. [21] While we choose the state variables to match the behavior of one-month options exactly, Figure 1.9 shows that the model also delivers a good fit to the time series of implied volatilities from three-month and six-month options.

The exercise above raises the question of whether the time series of λ_t and ξ_t are reasonable given our assumptions on the processes for these variables. Table 1.4 calculates the distribution of moments for λ_t and ξ_t. With the exception of the first-order autoregressive coefficients, the data fall well within the 90% confidence intervals. That is, the average values of the state variables and their volatilities could easily have been observed in 17-year samples with no disasters. The persistence in the data is somewhat lower than the persistence implied by the model. This illustrates the same tension between matching the time series of option prices and the time series of the price-dividend ratio discussed in the previous section.

As discussed in Section 1.4.1, it is not clear why adding a second state variable would allow the model to capture the time series variation in both the level and the slope, since both of these observables would be endogenously determined by both state variables. To better understand the mechanisms that allow the model to match these data, we consider the relative contributions of two state variables to the implied volatility curve in Figure 1.10. Panel A of Figure 1.10 fixes ξ_t at its median value and shows implied volatilities for λ_t at its 80th percentile value, at its median, and at its 20th percentile value. Panel B of Figure 1.10 performs the analogous exercise, this time fixing λ_t but varying ξ_t. In contrast to its behavior in the one-factor model, increasing λ_t both increases implied volatilities and increases the slope. The state variable ξ_t has a smaller effect on the level of implied volatilities, but a greater effect on the slope. Moreover, an increase in ξ_t lowers the slope rather than raising it. To summarize, increases in λ_t raise both the level and the slope of

[21] While this model can formally capture the financial crisis, there is a tension in that the large spike in λ_t would be an extremely unlikely event in a model in which state variables are driven by Brownian shocks. Our methods can easily accomodate Poisson jumps in λ_t itself, or a third transitory factor that would drive λ_t.

the implied volatility curve. Increases in ξ_t slightly raise the level and decrease the slope.

Why does ξ_t affect the slope of the implied volatility curve? The reason is that ξ_t has a relatively small effect on the probability of disasters at the time horizon important for option pricing, but a large effect on the volatility of stock prices (because of the square root term on its own volatility). Thus increases in ξ_t flatten the slope of the implied volatility curve because they raise the implied volatility for ATM options much more than for OTM options. The process for λ_t, on the other hand, has a large effect on OTM implied volatilities because it directly controls the probability of a disaster. It has a smaller effect on ATM implied volatilities than in the one-factor model because of its lower persistence.

Finally, we ask what these implied volatilities say about equity valuations. Given the option-implied values of ξ and λ, we can impute a price-dividend ratio using (1.19). This price-dividend ratio uses no data on equities, only data on options and the model. Figure Figure 1.12 shows the results, along with the price-dividend ratio from data available from Robert Shiller's webpage. The model can match the sustained level of the price-dividend ratio, and, most importantly, the time series variation after 2004.[22] Indeed, between 2004 and 2013, the correlation between the option-implied price-dividend ratio and the actual price-dividend ratio is 0.84, strongly suggesting these two markets share a common source of risk.

1.5.4. The implied volatility surface

We now take a closer look at what the model says about the implied volatility surface, namely implied volatilities across moneyness and time to expiration. Figures 1.11 show the implied volatilities for one, three and six-month options as a function of moneyness. We show the average implied volatilities computed using the time series of λ_t and ξ_t, and compare these to the data. Given that we have chosen λ_t and ξ_t to match the time series of implied volatilities on 1-month options, it is not surprising that the model matches these

[22] Not surprisingly, the disaster-risk model is not able to match the run-up in stock prices from the late 90s until around 2004. It may be that time-varying fears of a disaster will not be able to capture the extreme optimism that characterized that period.

data points exactly. More interestingly, the model is able to match the three-month and the six-month implied volatility curves almost exactly, even though it was not calibrated to these curves.

The downward slope in implied volatilities from options expiring in as long as six months indicates that the risk-neutral distribution of returns exhibits considerable skewness at long horizons. This is known as the skewness puzzle (Bates (2008)) because the law of large numbers would suggest convergence toward normality as the time to expiration increases. Recently, Neuberger (2012) makes use of options data to conclude that the skewness in the physical distribution of returns is also more pronounced than has been estimated previously. Neuberger emphasizes the observed negative correlation between stock prices and volatility (French, Schwert, and Stambaugh (1987)) as a reason why skewness in long-horizon returns does not decay as the law of large numbers in an iid model suggests that it would (see also Bates (2000)).

Figures 1.7 and 1.11 show that our model can capture the downward slope in 6-month implied volatilities as well as the slope for shorter-term options. Thus stock returns in the model exhibit skewness at both long and short-horizons. The short-horizon skewness arises from the existence of rare disasters. Long-horizon skewness, however, comes about endogenously because of the time-variation in the disaster probability. An increase in the rare disaster probability leads to lower stock prices, and, at the same time, higher volatilities, thereby accounting for this co-movement in the data. As a result, returns maintain their skewness at long horizons, and the model can explain six-month as well as one-month implied volatility curves.

1.6. Conclusion

Since the early work of Rubinstein (1994), the implied volatility curve has constituted an important piece of evidence against the Black-Scholes Model, and a lens through which to view the success of a model in matching option prices.

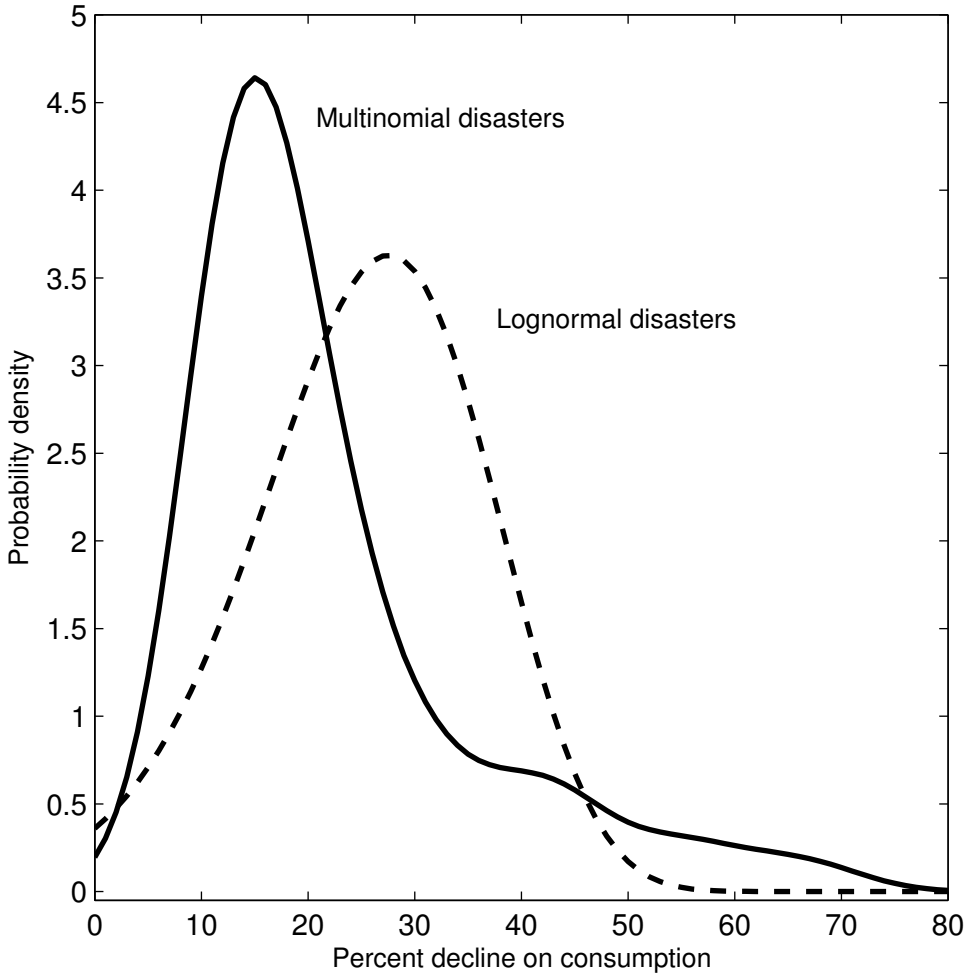
The implied volatility curve, almost by definition, has been associated with excess kurtosis in stock prices. Separately, a literature has developed linking kurtosis in consumption (which would then be inherited by returns in equilibrium) with the equity premium. However, much of the work up to now, as exemplified by a recent paper by Backus, Chernov, and Martin (2011) suggests that, at least for standard preferences, the non-normalities required to match the equity premium are qualitatively different from those required to match implied volatility.

We have proposed an alternative and more general approach to modeling the risk of downward jumps that can reconcile the implied volatility curve and the equity premium. Rather than assuming that the probability of a large negative event is constant, we allow it to vary over time. The existence of very bad consumption events leads to both the downward slope in the implied volatility curve and the equity premium. Moreover, the time-variation in these events moderates the slope, raises the level and generates the excess volatility observed in stock prices. Thus the model can simultaneously match the equity premium, equity volatility, and implied volatilities on index options. Option prices, far from ruling out rare consumption disasters, provide additional information for the existence of what has been referred to as the "dark matter" of asset pricing (Campbell (2008), Chen, Dou, and Kogan (2013)).

The initial model that we develop in the paper is deliberately simple and parsimonious. However, there are some interesting features of option and stock prices that cannot be matched by a model with a single state variable; for example, the imperfect correlation between the slope and the level of the implied volatility curve. For this reason, we investigate a more general model that allows for variation in disaster risk to occur at multiple time scales. This modification naturally produces time-variation in the slope of implied volatilities because it introduces variation in stock price volatility that can be distinguished from the risk of rare disasters. Taken together, these results indicate that options data support the existence of rare disasters in beliefs about the equity premium. Moreover, options

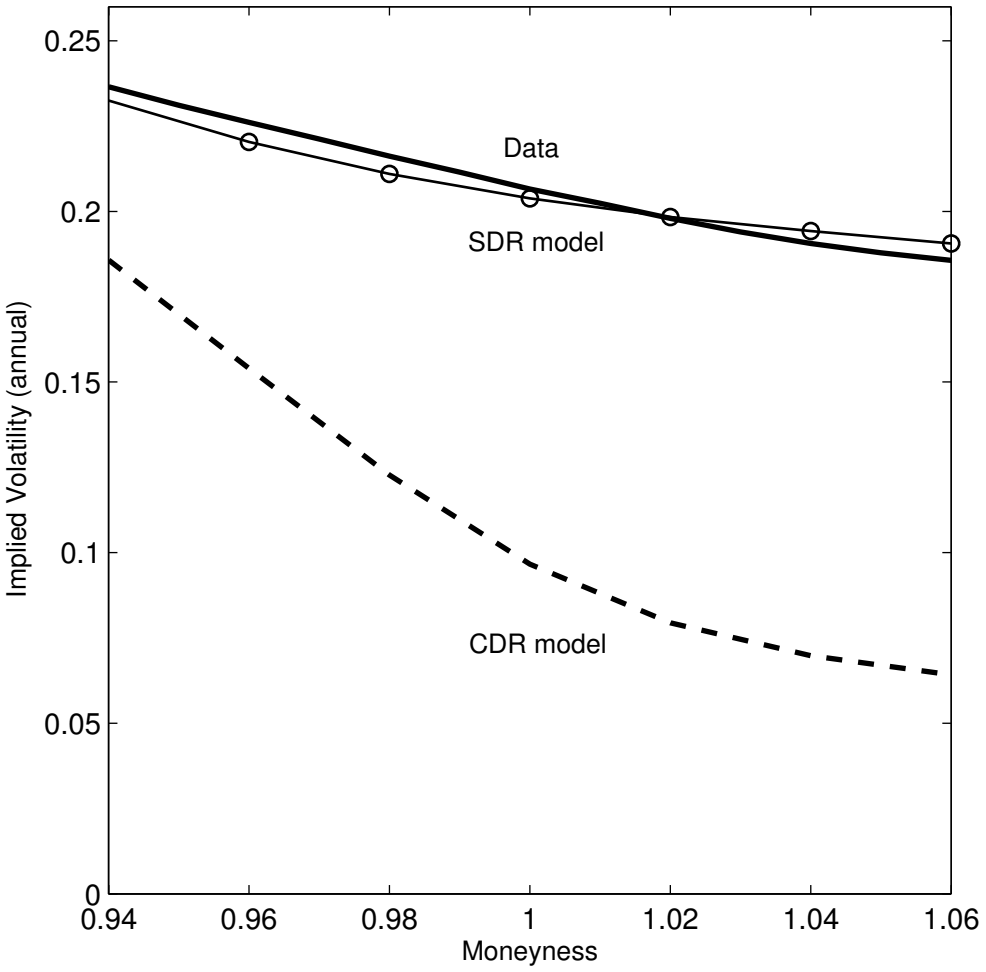
data can provide information about the disaster distribution beyond that offered by stock prices. In particular, data from options suggest that modeling time-variation in disaster risk occurring at multiple time scales may be a fruitful avenue for future work.

Figure 1.1: Probability density functions for consumption declines



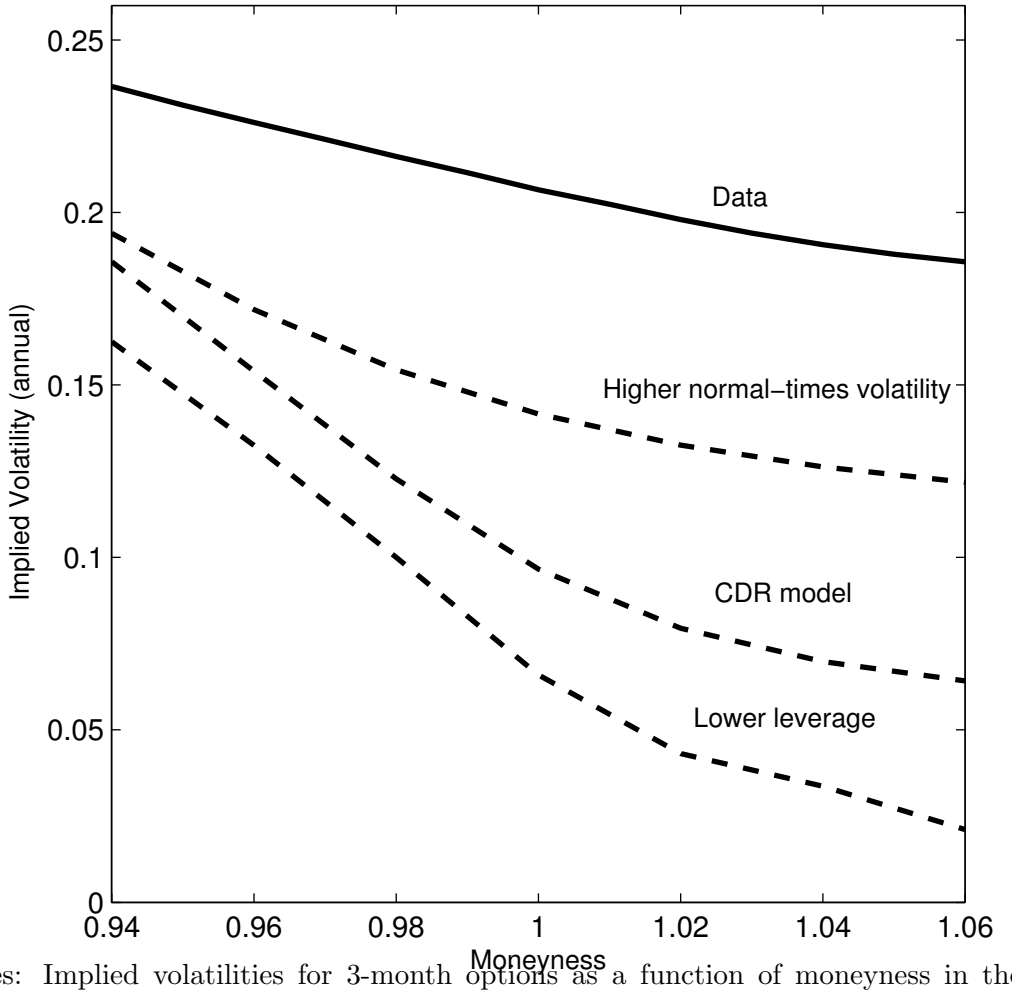
Notes: The probability density functions (pdfs) for consumption declines for log-normally distributed disasters and for the multinomial distribution assumed in the stochastic disaster risk (SDR) model. In the case of the SDR model, the pdf approximates the multinomial distribution from Barro and Ursua (2008). The exact multinomial distribution is used to calculate the results in the paper. The pdfs are for the quantities $1 - e^Z$ in each model.

Figure 1.2: Average implied volatilities in the SDR and CDR models



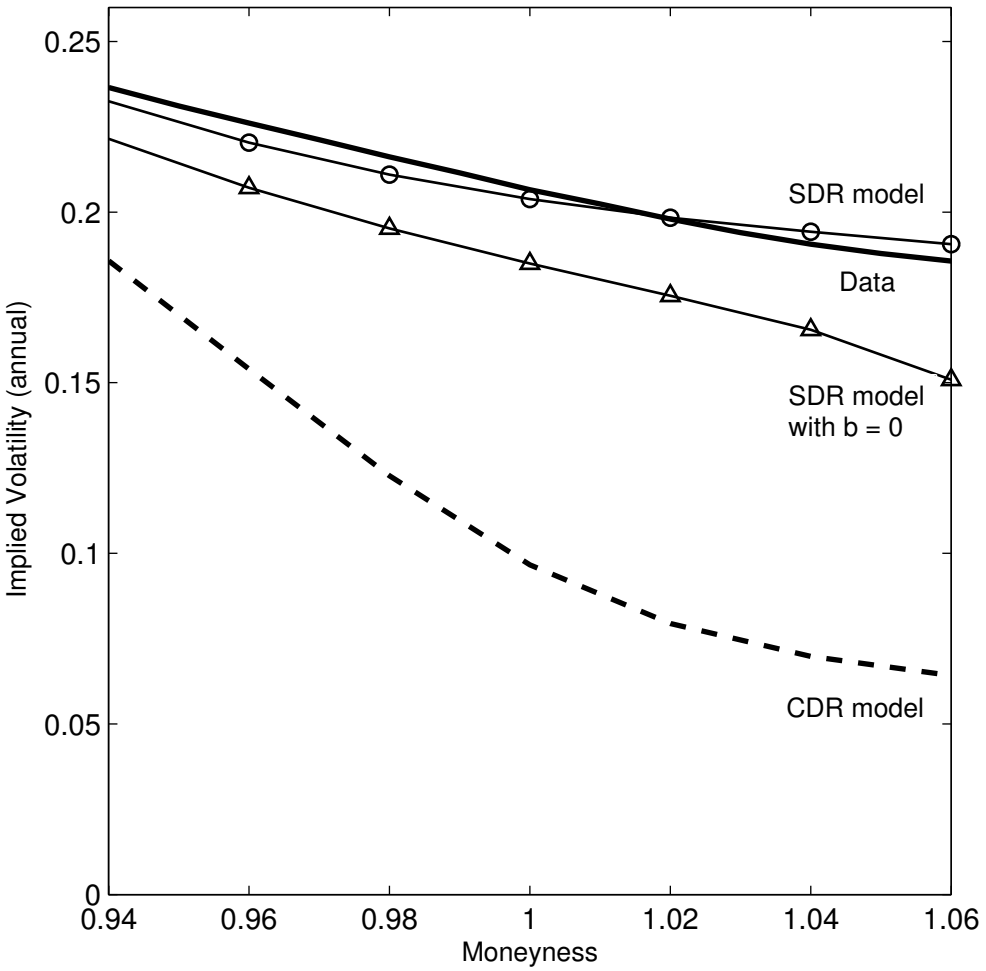
Notes: Average implied volatilities for 3-month options as a function of moneyness for the stochastic disaster risk (SDR) model, for the constant disaster risk (CDR) model (under the benchmark calibration given in Table 1.1) and in the data. Average implied volatilities are shown as functions of moneyness, defined as the exercise price divided by the asset price.

Figure 1.3: Comparative statics for the CDR model



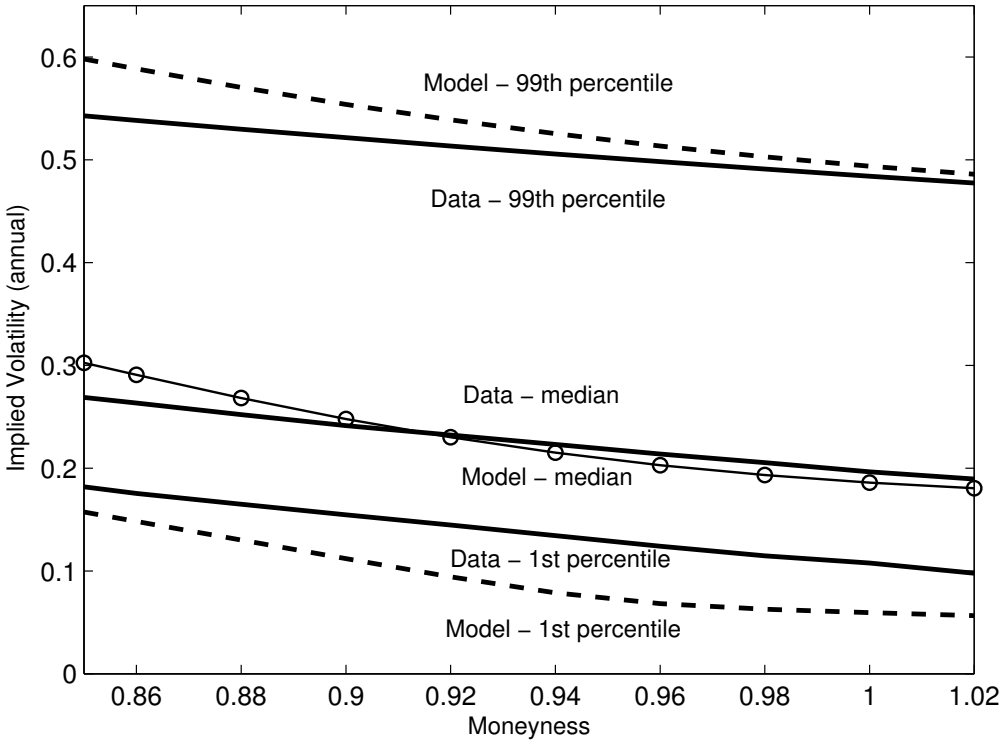
Notes: Implied volatilities for 3-month options as a function of moneyness in the data and for three parameterization of the CDR model. The line labeled "CDR" shows the benchmark calibration. The line labeled "higher normal-times volatility" raises the volatility of consumption shocks that are not associated with disasters from 1% to 2% per annum but keeps all other parameters, including the consumption disaster distribution, the same. The line labeled "lower leverage" lowers the term multiplying dividends from 5.1 to 2.6, while keeping all other parameters the same.

Figure 1.4: Evaluating the role of recursive utility



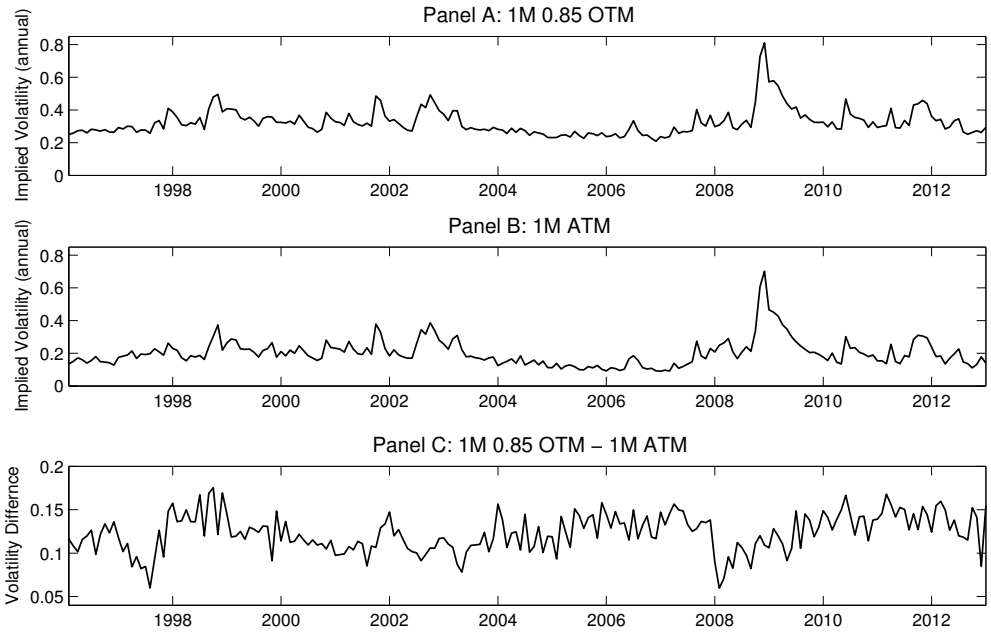
Notes: Implied volatilities for 3-month options as a function of moneyness in the data, in the CDR model (calibrated as in Table 1.1) and in the SDR model. Also shown are implied volatilities in the SDR model computed under the assumption that the premium associated with time-variation in the disaster probability is equal to zero (SDR model with $b = 0$). Note both the benchmark and the $b = 0$ version of the SDR model are dynamic models with stochastic moments. The CDR model is iid.

Figure 1.5: Implied volatilities for given values of the disaster probability



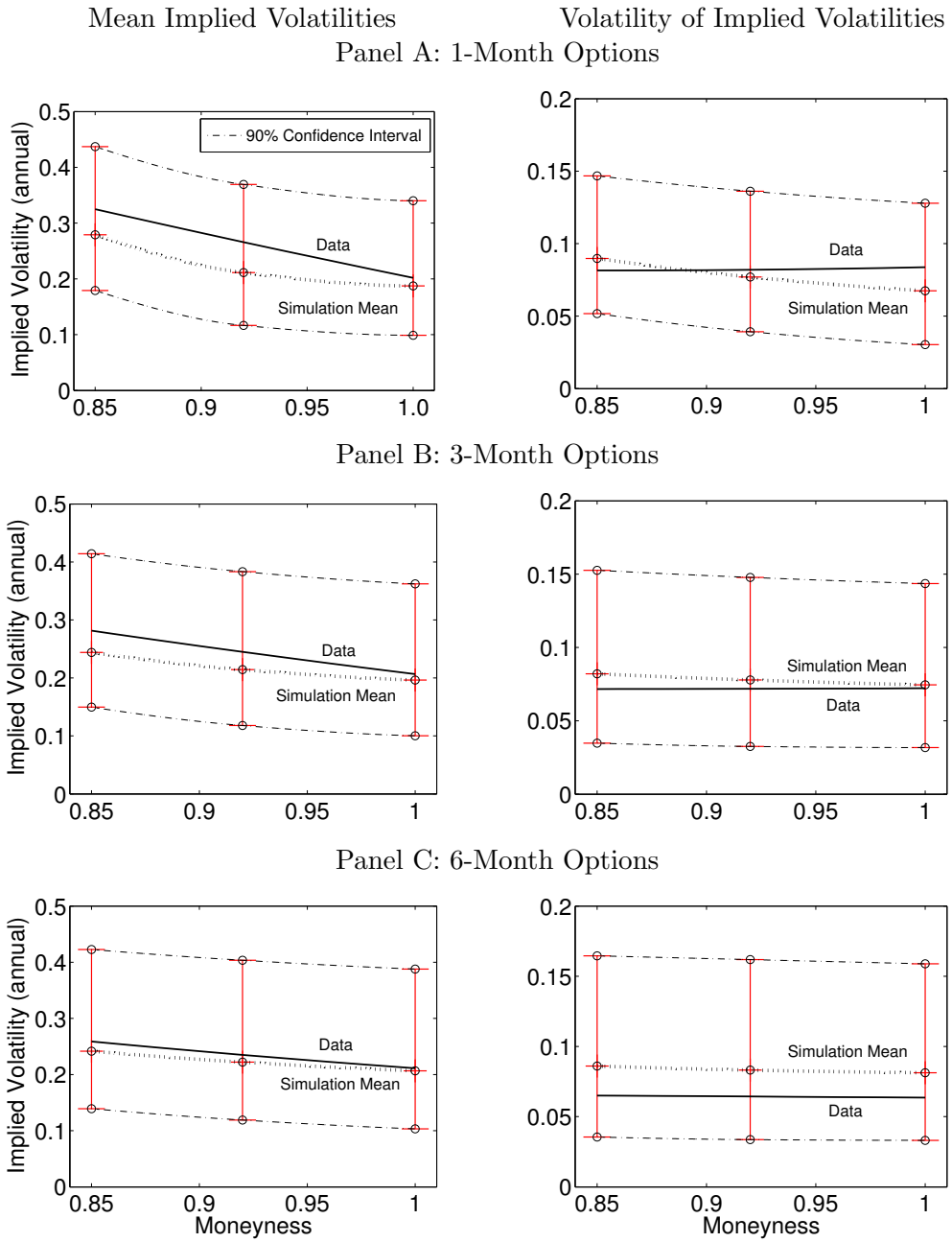
Notes: Implied volatilities on 3-month options as a function of moneyness for various percentile values of the probability of a disaster in the single-factor SDR model. Also shown are average implied volatilities in the data, and, for each moneyness value, 99th and 1st percentile implied volatilities in the data.

Figure 1.6: 1-month implied volatility time series



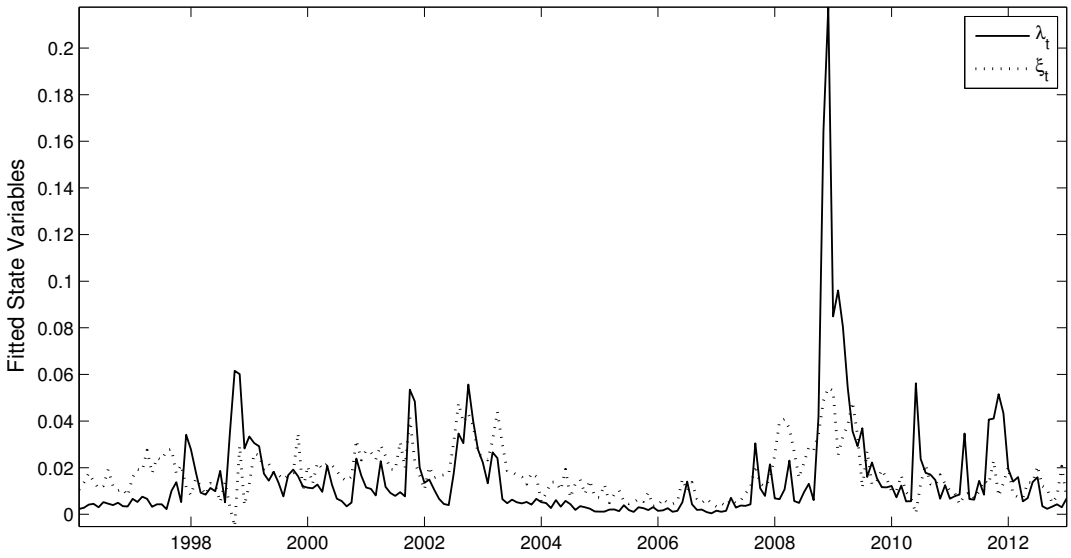
Notes: Monthly time series of implied volatilities on 1-month options in the data and of the difference in implied volatilities. Implied volatilities are calculated for at-the-money (ATM) options and out-of-the-money (OTM) options with moneyness equal to 0.85.

Figure 1.7: Mean and volatility of implied volatilities in simulated data



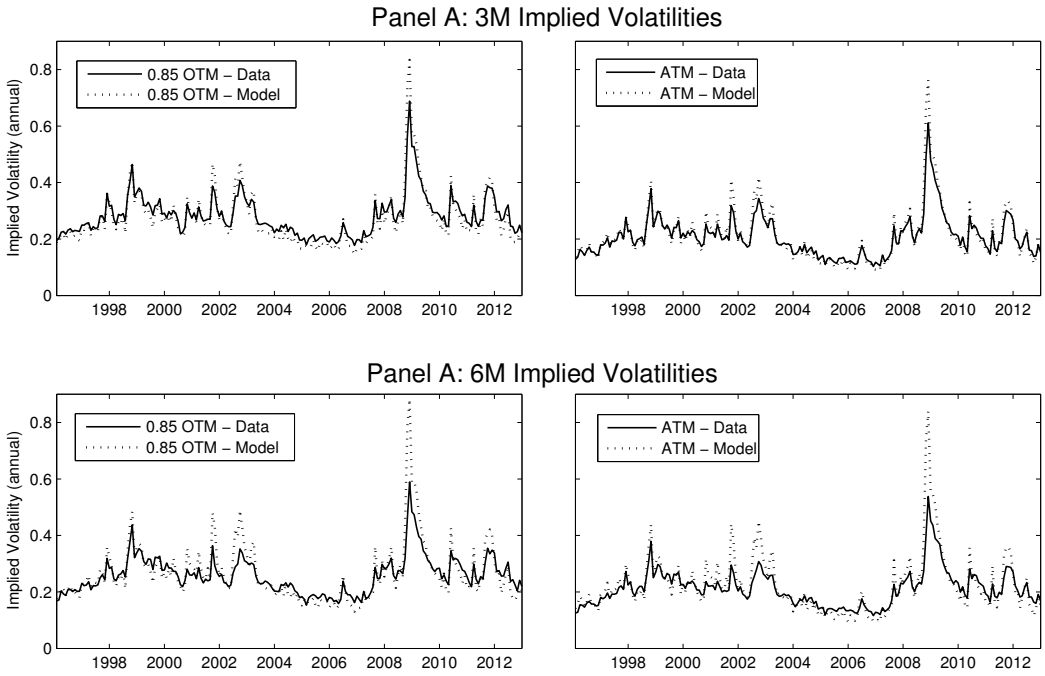
Notes: We simulate 1000 samples of length 17 years from the two-factor model. For each sample path, we compute the mean and volatility of implied volatilities at three different moneyness levels and for three maturities. The dotted line shows the means of these statistics across sample paths, while the dashed-dotted lines show 95th and 5th percentiles. The solid line shows the data.

Figure 1.8: Fitted values of state variables



Notes: Fitted state variables in the two-factor SDR model. State variables are fitted to the time series of 1-month implied volatilities of ATM options and OTM options with moneyness equal to 0.85. The variable λ_t represents the (annual) probability of disaster. λ_t reverts to a time-varying value ξ_t.

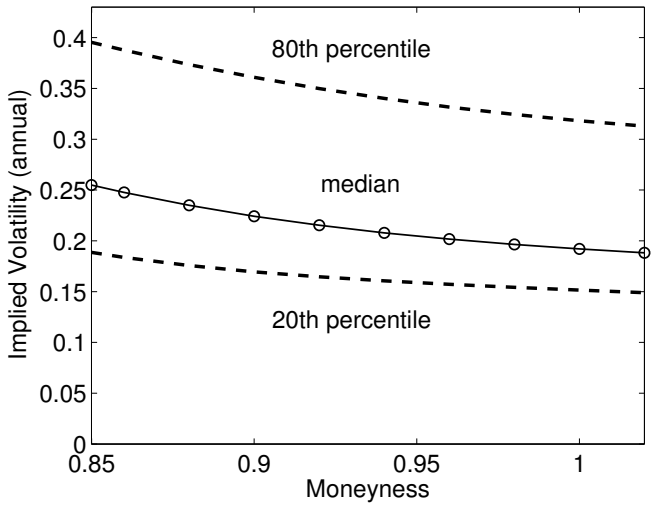
Figure 1.9: 3- and 6-month implied volatility time series



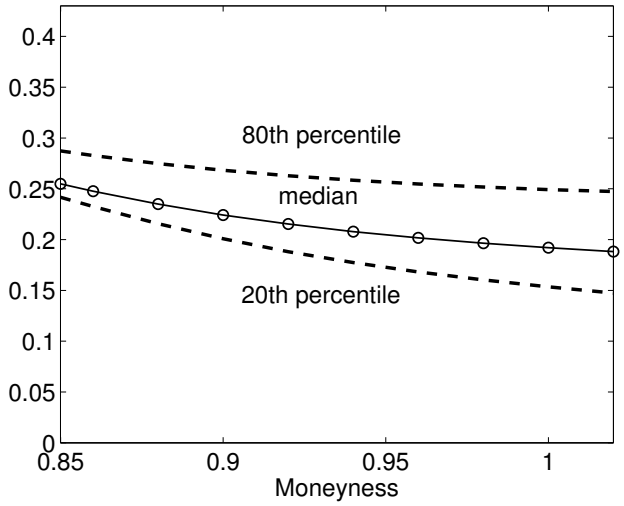
Notes: Monthly time series of implied volatilities on 3- and 6-month options in the data and in the two-factor SDR model. Implied volatilities are calculated for ATM options and OTM options with moneyness equal to 0.85. In the model, we compute option prices using state variables fit to the time series of 1-month ATM and OTM options with moneyness of 0.85.

Figure 1.10: Implied volatilities as functions of the state in the two-factor SDR model

Panel A: Varying λ_t

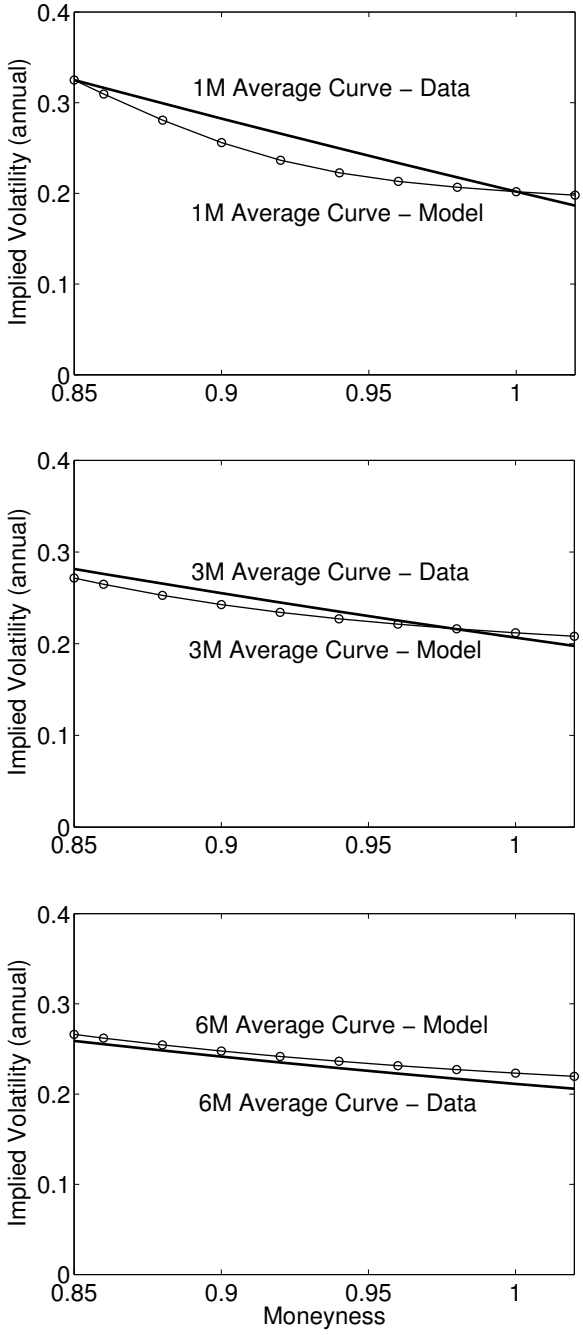


Panel B: Varying ξ_t



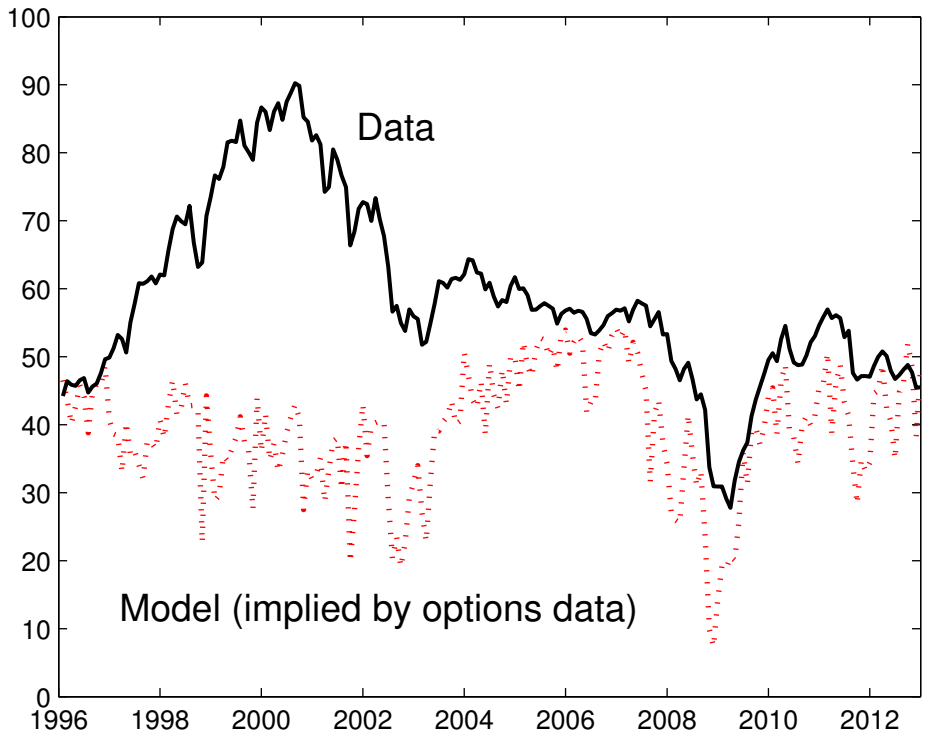
Notes: Implied volatilities on 3-month options as functions of moneyness. The figures show the effects of varying the state variables λ_t (the disaster probability) and ξ_t (the value to which λ_t reverts). Panel A sets ξ_t equal to its median value and varies λ_t, while Panel B sets λ_t equal to its median value and varies ξ_t.

Figure 1.11: Average implied volatilities from the time series of state variables



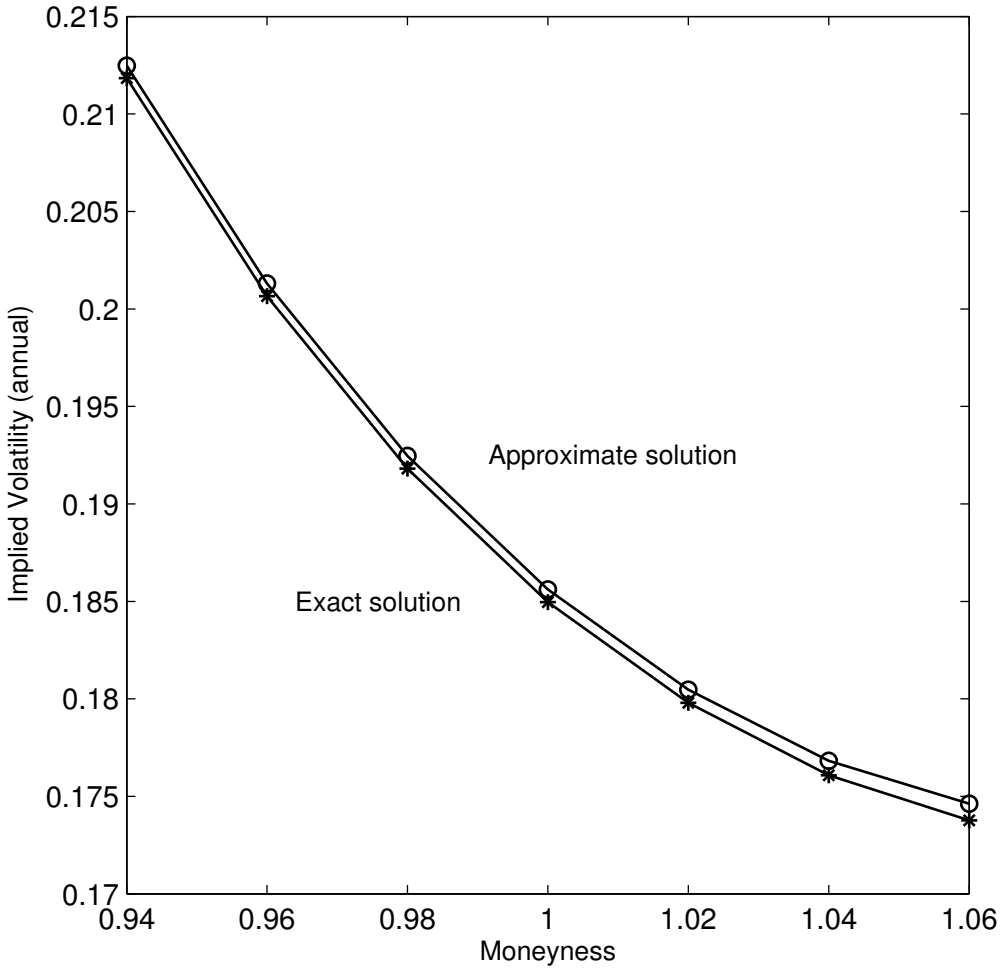
Notes: Average implied volatilities on put options in the data and in the two-factor SDR model. In the model, we compute option prices using state variables fit to the time series of 1-month ATM and OTM options with moneyness of 0.85. We compute an implied volatility for each option price and then take the average.

Figure 1.12: The price-dividend ratio in the data and in the model



Notes: The solid line shows the time series of the price-dividend ratio on the data. The red line shows the price-dividend ratio implied by the model for state prices chosen to fit the one-month ATM and OTM (0.85 moneyness) put options.

Figure A.1: Exact versus approximate solution



Notes: This figure shows implied volatilities when option prices are computed exactly (line with stars) versus when they are computed using an approximation (line with circles) in the SDR model. Implied volatilities assume that the disaster probability is fixed at its mean and that, in the event of disaster, consumption falls by a fixed amount, namely 30%.

Table 1.1: Parameter values

	SDR	CDR
Relative risk aversion γ	3.0	5.19
EIS ψ	1	1
Rate of time preference β	0.0120	0.0189
Average growth in consumption (normal times) μ	0.0252	0.0231
Volatility of consumption growth (normal times) σ	0.020	0.010
Leverage ϕ	2.6	5.1429
Average probability of a rare disaster $\bar{\lambda}$	0.0355	0.010
Mean reversion κ	0.080	NA
Volatility parameter σ_λ	0.067	0

Notes: Parameters for stochastic disaster risk (SDR) model and for the benchmark constant disaster risk (CDR) model, in annual terms.

Table 1.2: Parameter values for the two-factor SDR model

Panel A: λ process	
Mean reversion κ_λ	0.2
Volatility parameter σ_λ	0.1576

Panel B: ξ process	
Mean $\bar{\xi}$	0.02
Mean reversion κ_ξ	0.1
Volatility parameter σ_ξ	0.0606

Panel C: Population statistics of λ	
Median	0.0037
Standard deviation	0.0386
AR(1) coefficient	0.9858

Notes: Parameter values for the two-factor SDR model. The processes are as follows:

$$\begin{aligned} d\lambda_t &= \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} \\ d\xi_t &= \kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t}. \end{aligned}$$

Parameter values are expressed in annual terms. Panel C shows population statistics for the disaster probability λ_t calculated by simulation at a monthly frequency (so that the AR(1) coefficient should be interpreted in monthly terms).

Table 1.3: Moments for the government bill rate and the market return for the two-factor SDR model

		No-Disaster Simulations			All Simulations			
	Data	0.05	0.50	0.95	0.05	0.50	0.95	Population
$E[R^b]$	1.25	1.68	2.96	3.46	-0.47	2.41	3.37	2.02
$\sigma(R^b)$	2.75	0.34	1.07	2.71	0.48	2.06	7.14	3.69
$E[R^m - R^b]$	7.25	5.40	8.01	12.36	5.30	8.49	14.25	9.00
$\sigma(R^m)$	17.8	13.24	19.26	27.91	14.59	22.59	34.38	24.13
Sharpe Ratio	0.41	0.32	0.42	0.55	0.26	0.39	0.53	0.37
$\exp(E[p-d])$	32.5	28.96	40.63	48.88	22.93	36.95	47.41	35.36
$\sigma(p-d)$	0.43	0.15	0.27	0.47	0.17	0.33	0.59	0.43
AR1$(p-d)$	0.92	0.59	0.79	0.91	0.62	0.82	0.92	0.90

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated from simulating data from the two-factor stochastic disaster risk (SDR) model at a monthly frequency for 600,000 years and then aggregating monthly growth rates to an annual frequency. We also simulate 100,000 60-year samples and report the 5th, 50th and 95th percentile for each statistic, both from the full set of simulations and for the subset of samples for which no disasters occur. R^b denotes the government bill return, R^m denotes the return on the aggregate market and $p - d$ denotes the log price-dividend ratio.

Table 1.4: Moments of state variables in the two-factor SDR model

		No-Disaster Simulations			All Simulations		
	Data	0.05	0.50	0.95	0.05	0.50	0.95
$E[\lambda]$	1.57	0.11	0.83	4.64	0.13	1.17	7.54
$\sigma(\lambda)$	2.37	0.22	1.18	4.37	0.26	1.53	5.92
$AR1(\lambda)$	0.74	0.76	0.93	0.98	0.78	0.94	0.98
$E[\xi]$	1.70	0.43	1.42	4.29	0.46	1.58	4.95
$\sigma(\xi)$	1.06	0.32	0.83	2.02	0.34	0.90	2.22
$AR1(\xi)$	0.70	0.90	0.96	0.98	0.91	0.96	0.98

Notes: Data moments of state variables are calculated using state variables extracted from monthly data on option prices from 1996 to 2012. The parameter λ_t is the probability of a disaster. The probability of a disaster reverts to the state variable ξ_t. Means and standard deviations are in percentage terms. We simulate 100,000 17-year samples at a monthly frequency and report the 5th, 50th and 95th percentile for each statistic, both from the full set of simulations and for the subset of samples for which no disasters occur.

CHAPTER 2 : Do Rare Events Explain CDX Tranche Spreads?

(with Jessica A. Wachter)

2.1. Introduction

In both academic work and the popular press, financial engineering has been implicated as a cause of the 2008 financial crisis. One hypothesis is that complex securities were falsely given high ratings. These sold at low interest rates, allowing financial firms to take on more leverage. By making financial firms extremely sensitive to small losses, this artificially high leverage contributed to the crisis.

In this paper, we focus on securities based on an index of credit default swaps. This index is known as the CDX, and tranches on the CDX were actively traded both before and during the crisis. Accordingly, Coval, Jurek, and Stafford (2009) identify mis-pricing in these tranches, based on a model that computes risk using option prices, and fits the spreads assuming that the spread on the 5-year CDX is priced correctly. They find that spreads on the senior tranches are too low and that spreads on the junior tranches are too high, consistent with the hypothesis above. This story is questioned, however, by Collin-Dufresne, Goldstein, and Yang (2012), who argue that the focus on the 5-year spread distorts the results. Indeed, in order to price spreads at all maturities, it is necessary to allow for idiosyncratic jumps. Furthermore, a small probability of a system-wide catastrophic jump is necessary for matching overall credit spreads.

Unlike Coval, Jurek, and Stafford (2009), who use a static model that assumes default at the maturity of a bond, Collin-Dufresne, Goldstein, and Yang (2012) solve a dynamic model of structural default. They specify a stochastic discount factor and allow for both systematic and idiosyncratic jumps in asset value. However, their model is still reduced-form, and therefore disconnected from underlying risks in the economy. For example, their model allows for the specification of jump risk under the risk-neutral measure, but not the

physical measure. Furthermore, they solve their model under the assumption of constant jump risk. In their calibration, however, they assume that jump risk is time-varying.

In this paper, we aim to value tranches on the CDX using a representative-agent model. We require that the model be consistent with facts about the aggregate market, and, importantly, option prices, as the connection between options and credit spreads is well known. Our model allows for a catastrophic jump in prices driven by consumption (a rare disaster, as in Barro (2006) and Rietz (1988)), as well as large idiosyncratic jumps in assets covered in the CDX index. Our model also allows for the probability of both types of jumps to vary over time. We calibrate the model to match prices on index options. In order to match option prices, we use a two-factor model for the risk of rare disasters developed in earlier work (Chapter 1). We find that our model can explain the spreads on both CDX junior and senior spreads, while at the same time generating a reasonable fit to implied volatilities on index options. Rather than arising from mis-pricing, low spreads on senior CDX tranches could plausibly be observed in a period characterized by a low risk of a disaster. Options and equity data also point to 2005 to mid-2007 as being such a period.

2.2. Model

2.2.1. Model primitives and the state-price density

We use the two-factor stochastic disaster risk model in Chapter 1. Namely, we assume an endowment economy with complete markets and an infinitely-lived representative agent. Aggregate consumption (the endowment) solves the following stochastic differential equation:

$$\frac{dC_t}{C_{t^-}} = \mu_C dt + \sigma_C dB_{C,t} + (e^{Z_{C,t}} - 1)dN_{C,t},$$

where $B_{C,t}$ is a standard Brownian motion and $N_{C,t}$ is a Poisson process. The intensity of $N_{C,t}$ is given by λ_t and assumed to be governed by the following system of equations:

$$\begin{aligned} d\lambda_t &= \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} \\ d\xi_t &= \kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t}, \end{aligned}$$

where $B_{\lambda,t}$ and $B_{\xi,t}$ are Brownian motions (independent of each other and of $B_{C,t}$).

We will assume a recursive generalization of power utility that allows for preferences over the timing of the resolution of uncertainty. Our formulation comes from Duffie and Epstein (1992), and we consider a special case in which the parameter that is often interpreted as the elasticity of intertemporal substitution (EIS) is equal to 1. That is, we define continuation utility V_t for the representative agent using the following recursion:

$$V_t = E_t \int_t^\infty f(C_s, V_s)\, ds,$$

where

$$f(C,V) = \beta(1-\gamma)V\left(\log C - \frac{1}{1-\gamma}\log((1-\gamma)V)\right). \tag{2.1}$$

The parameter β is the rate of time preference. We follow common practice in interpreting γ as relative risk aversion. This utility function is equivalent to the continuous-time limit (and the limit as the EIS approaches one) of the utility function defined by Epstein and Zin (1989) and Weil (1990).

As shown in Appendix A2.1, the value function has the solution

$$V_t = J(C_t, \lambda_t, \xi_t) = \frac{C^{1-\gamma}}{1-\gamma}e^{a+b_\lambda\lambda_t+b_\xi\xi_t},$$

with coefficients a, b_λ and b_ξ given in the Appendix. In Appendix A2.2, we show that this

value function implies the following state-price density

$$\pi_t = \exp\left(\eta t - \beta b_\lambda \int_0^t \lambda_s ds - \beta b_\xi \int_0^t \xi_s ds\right) \beta C_t^{-\gamma} e^{a + b_\lambda \lambda_t + b_\xi \xi_t}, \tag{2.2}$$

and risk free rate

$$r_t = \beta + \mu - \gamma\sigma^2 + \lambda_t E\left[e^{(1-\gamma)Z_t} - e^{-\gamma Z_t}\right].$$

Equations for the aggregate market and index options can be found in Chapter 1. Here, we focus on computing CDX prices.

2.2.2. Defaults

Let $D_{i,t}$ be the payout amount of firm i ($i = 1, \cdots, N$). While we use the notation $D_{i,t}$, we intend this to mean the payout not only to the equity holders but the bondholders as well. The payout is driven by common sources of risk B_{ct} and N_{ct} that affect consumption as well, as well as idiosyncratic sources of risk B_t^i and N_t^i. That is,

$$\frac{dD_{i,t}}{D_{i,t^-}} = \mu_i dt + \phi_i \sigma_C dB_{C,t} + (e^{\phi_i Z_{C,t}} - 1) dN_{C,t} + \underbrace{\sigma_i dB_t^i + (e^{Z_t^i} - 1) dN_t^i}_{\text{idiosyncratic risk}}.$$

The processes B_t^i, for $i = 1, \ldots, N$, are independent Brownian motions. The processes Nit have Poisson increments. To avoid a multiplicity of state variables, we assume that the intensity of the Poisson processes are jointly driven by χ_t. However, conditional on χ_t, the increments dN_{it} are independent. We assume that χ follows a CIR process:

$$d\chi_t = \kappa_\chi(\bar{\chi} - \chi_t)dt + \sigma_\chi \sqrt{\chi_t} dB_{\chi,t}.$$

This structure allows for exposure to disasters (through N_{ct}), as well as to large idiosyncratic events. Note that, as is usual in the literature on endowment economies, these firms have greater exposure to disasters than the consumption process itself, as indicated by the

parameter ϕ.[1]

Let $A_{i,t}$ denote the total value of firm i (the equity plus the debt). That is $A_{i,t}$ is the price of the payout stream:

$$A_{i,t} = E_t \left[\int_t^\infty \frac{\pi_s}{\pi_t} D_{i,s} ds \right]. \tag{2.3}$$

We denote $G_i(\lambda_t, \xi_t, \chi_t)$ as the asset-payout ratio $(A_{i,t}/D_{i,t})$ of firm i. We show that this ratio is expressed as

$$G_i(\lambda_t, \xi_t, \chi_t) = \int_0^\infty \exp\left(a_i(\tau) + b_{i\lambda}(\tau)\lambda_t + b_{i\xi}(\tau)\xi_t + b_{i\chi}(\tau)\chi_t\right) d\tau, \tag{2.4}$$

where $a_i(\tau)$, $b_{i\lambda}(\tau)$, $b_{i\xi}(\tau)$, and $b_{i\chi}(\tau)$ solve the system of ordinary differential equations derived in Appendix A2.3.

We define default as the event that a firm's value falls below a threshold following Black and Cox (1976) and many subsequent studies. Let $A_{i,B}$ denote the threshold for firm i. Then the default time is the random variable defined as

$$\tau_i \quad = \quad \inf\{t : A_{i,t} \leq A_{i,B}\}.$$

We let R_i denote the recovery rate of firm i upon default.

2.2.3. CDX pricing

CDX indices are baskets of equally-weighted individual credit default swap (CDS) contracts for a set of large U.S. investment-grade firms.[2] We will let N denote the number of firms in the index (for all CDX indices thus far, the number of firms has been 125). Taking a \$1

[1]Note however that ϕ does not have the interpretation of financial leverage, as D_{it} is total payout. However, $\phi > 1$ is still possible in the presence of labor income. Our calibration of ϕ will reflect the interpretation of D_{it} as including interest as well as dividends.

[2]Each CDS contract involves two parties: protection buyer and protection seller. The protection seller provides the protection buyer an insurance against a credit event of the reference entity (specified in a CDS contract). Thus, the protection buyer pays a series of insurance premiums to the protection seller. In return, if the reference entity experiences a credit event, the protection seller compensates the loss of the protection buyer, either by purchasing the reference obligation from the protection buyer at its par value (physical settlement) or by directly paying the loss amount to the protection buyer (cash settlement).

protection sell position on the CDX index can be viewed as taking protection-sell positions with notional amounts $\$1/N$ on all N individual CDS contracts in the index.

We first develop formulas for the payoffs and pricing of the protection-sell and protection-buy positions on the index. Assume a \$1 notional amount. If firm i defaults, then the protection seller pays the protection buyer $\frac{1}{N}(1-R_i)$. The cumulative loss process on the index, L_t, can therefore be expressed as

$$L_t = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\tau_i \le t\}}(1-R_i). \tag{2.5}$$

The value of the cash flows paid by the protection seller is given by

$$\mathbf{Prot}_{\text{CDX}} = \tilde{E}\left[\int_0^T e^{-\int_0^t r_s ds} dL_t\right]. \tag{2.6}$$

where $\tilde{E}$ denotes the expectation taken under the risk-neutral measure and r_t is the riskfree rate.[3]

In the CDX contract, the protection buyer makes quarterly premium payments that, over the course of a year, add up to a pre-determined spread S. If a firm defaults, the size of premium payments goes down because the outstanding notional amount of the CDS contract is reduced by $\frac{1}{N}$. Let n_t denote the fraction of firms that have defaulted as of time t:

$$n_t = \frac{1}{N}\sum_{i=1}^{N} 1_{\{\tau_i \le t\}}. \tag{2.7}$$

For a given spread S, the expected present value of cash flows paid by the protection buyer is given by

$$\mathbf{Prem}_{\text{CDX}}(S) = S\tilde{E}\left[\sum_{m=1}^{M} e^{-\int_0^{t_m} r_s ds}(1-n_{t_m})\Delta_m + \int_{t_{m-1}}^{t_m} e^{-\int_0^u r_s ds}(u-t_{m-1})dn_s\right], \tag{2.8}$$

[3]The risk-neutral measure and riskfree rate are implied by the model in Section 2.2.1.

where $0 = t_0 < t_1 < \cdots < t_M$ are premium payment dates, S is the premium rate (premium payment per unit notional), and $\Delta_m = t_m - t_{m-1}$ is the m-th premium payment interval.[4] Note that if a default occurs between two premium payment dates, the protection buyer must pay the accrued premium from the last premium payment date to the default date. This is taken into consideration as the second term in the expectation of equation (2.8).

To simplify notation, we have assumed above that the contracts originate at time 0. Because the model is stationary, this assumption is without loss of generality. The CDS spread S_{CDX} is defined as the value of the premium rate that equates the protection and premium legs (i.e. $\mathbf{Prem}_{\text{CDX}}(S_{\text{CDX}}) = \mathbf{Prot}_{\text{CDX}}$). That is, S_{CDX} is determined by

$$S_{\text{CDX}} = \mathbf{Prot}_{\text{CDX}}/\mathbf{Prem}_{\text{CDX}}(1).$$

Appendix A2.4 shows how to compute the spread, which is a function of the state variables λ, ξ and χ.

2.2.4. CDX tranche pricing

Each tranche is defined by its "attachment point" and "detachment point." The attachment point refers to the level of subordination of the tranche, and the detachment point is the level at which the tranche loses its entire notional amount. For example, consider the tranche with a 10% attachment point and 15% detachment point. The protection seller of this tranche does not experience loss until the the entire pool (i.e. the CDX index) accumulates more than a 10% loss. After the 10% level, every loss the pool experiences "attaches" to the tranche. That is, the protection seller of the tranche should compensate the loss after 10%. If the loss of the pool reaches the detachment point of 15%, the protection seller loses the entire notional amount and no longer needs to take the loss of the pool. (i.e. further loss "detaches" from the tranche.)

[4] If T is the maturity of the CDS contract (in years), the total number of premium payments M is equal to $4T$ because payments are quarterly.

The detachment of a tranche is the attachment of the next (higher) tranche. Suppose that there are J tranches based on the CDX index. We denote K_{j-1} as the attachment point and K_j as the detachment point of the j-th tranche tranche. The attachment point of the first tranche (K_0) is equal to 0 and the detachment point of the last tranche (K_J) is equal to 100%.

To properly understand the mechanism of tranche products, it is important to see that a default in the pool not only affects the most junior tranche but also the most senior tranche. To illustrate this, consider a simple case where (1) the entire pool consists of 100 CDS contracts, each of which has \$1 notional amount, and (2) there are only two tranches: 0-50% and 50-100% tranches based on the pool. How does an event of one firm's default in the pool affect the two tranches, assuming that the defaulted firm's recovery rate is 40%? This event affects both the 0-50% and 50-100% tranches. Since the recovery rate is 40% and each CDS contract is with \$1 notional, the loss amount of the pool is \$0.6. Since the 0-50% tranche is the most junior tranche, the protection seller of this tranche takes this loss. The difference between \$1 (notional amount of the defaulted CDS contract) and \$0.6 (loss amount from the defaulted CDS contract) reduces the notional amount of the 50-100% tranche exactly by \$0.4.[5]

The reason why the notional amount of the most senior tranche decreases is straightforward if we consider a case for funded CDO products: if a default occurs in the pool, the loss amount goes to the most junior tranche holder, and the recovery amount from the default goes to the most senior tranche holder as an early redemption. That is, the notional amount of the senior tranche is reduced by the redemption amount. Since CDX tranches are unfunded CDO products, which do not require initial investments, there is no actual early redemption in cash as the case for funded CDO products, but still the notional amount of the most senior tranche is reduced by the recovery amount.

[5]This is not a loss to the protection seller of this tranche. The notional amount simply reduces to \$49.6. This means that the maximum loss the protection seller can experience is \$49.6. Thus, the protection seller receives insurance premium based on the reduced notional amount \$49.6.

In sum, a default in the pool affects both the most junior tranche (by incurring loss) and the most senior tranche (by reducing its notional amount). Keeping this in mind, we derive the expressions for the loss and recovery amount of the j-th tranche as a fraction of the notional amount of the tranche $(K_j - K_{j-1})$:

$$\begin{aligned} T^L_{j,t} &= \frac{\min\{L_t, K_j\} - \min\{L_t, K_{j-1}\}}{K_j - K_{j-1}} \\ T^R_{j,t} &= \frac{\min\{n_t - L_t, 1 - K_{j-1}\} - \min\{n_t - L_t, 1 - K_j\}}{K_j - K_{j-1}}. \end{aligned} \tag{2.9}$$

Let $\mathbf{Prot}_{\text{Tran},j}$ be the protection leg of the j-th CDX tranche. The expression for $\mathbf{Prot}_{\text{Tran},j}$ is the same as the one for $\mathbf{Prot}_{\text{CDX}}$ except that the integral is with respect to the tranche loss $(T^L_{j,t})$ rather than the entire pool loss (L_t):

$$\mathbf{Prot}_{\text{Tran},j} = \tilde{E}\left[\int_0^T e^{-\int_0^t r_s ds} dT^L_{j,t}\right].$$

Let $\mathbf{Prem}_{\text{Tran},j}(U,S)$ be the premium leg of the j-th CDX tranche when the protection buyer pays quantity U up front and quarterly premiums with premium rate S. Since insurance premiums accrue based on the outstanding notional amount (which can be decreased from 100% due to a tranche loss $(T^L_{j,t})$ or a tranche recovery $(T^R_{j,t})$), we can show that

$$\mathbf{Prem}_{\text{Tran},j}(U,S) = U + S\tilde{E}\left[\sum_{m=1}^{Tf}\left\{e^{-\int_0^t r_s ds}\int_{t_{m-1}}^{t_m}\left(1 - T^L_{j,t} - T^R_{j,t}\right) ds\right\}\right].$$

In our sample period, the entire CDX pool has been partitioned into 6 tranches, which are traded as separate products: 0-3%, 3-7%, 7-10%, 10-15%, 15-30% and 30-100% tranches. The first tranche is called the equity tranche and the second tranche is called the mezzanine tranche. The four remaining tranches are called senior tranches, and, in particular, the last (most senior) tranche is called the super senior tranche.

Except for the equity tranche, tranches are traded with zero upfront amount. Thus, for

each of those tranches, the CDX tranche spread ($S_{\text{Tran},j}$) is determined by

$$S_{\text{Tran},j} = \frac{\mathbf{Prot}_{\text{Tran},j}}{\mathbf{Prem}_{\text{Tran},j}(0,1)} \quad \text{if } j = 2, \cdots, 6.$$

The equity tranche is traded with the upfront amount, fixing the premium rate at 500bp. Thus, the upfront amount for the equity tranche ($U_{\text{Tran},j}$) is determined by

$$U_{\text{Tran},j} = \mathbf{Prot}_{\text{Tran},j} - \mathbf{Prem}_{\text{Tran},j}(0, 0.05) \quad \text{if } j = 1.$$

Again, in Appendix A2.4, we show that the spread and upfront amount for CDX tranches can be calculated as a function of the state variables λ, ξ and χ.

2.3. Evaluating the model

2.3.1. Data

Our analysis requires the use of pricing data from the options and CDX markets. First, our sample of options consists of implied volatilities on S&P 500 European put options. We collect our sample from OptionMetrics, which provides the time series and the cross section of implied volatilities on individual stocks and equity indices from January 1996 to December 2012. To construct monthly time-series of implied volatility smiles, we pick observations using the data from the Wednesday of every option expiration week. Following the approach in Chapter 1, we apply standard filters to extract contracts with meaningful trade volumes and prices. The implied volatility smile for each date can be obtained by regressing implied volatilities on a polynomial in strike price and maturity.

CDX is a family of credit default swap (CDS) indices, mainly covering firms and entities in North America. Among many different indices, we focus on the CDX North America Investment Grade (CDX NA IG) index, which is the most actively traded. This index is composed of 125 equally weighted CDX contracts on representative North American investment grade firms.

Unlike equity indices, CDS indices have expiration dates because they consist of multiple contracts with certain maturities. Three-, five-, seven-, and ten-year indices are typically traded. Since each index has an expiration date, the time-to-maturity decays. For this very reason, CDX indices roll every 6 months in March and September. That is, every March and September, a new series of indices is introduced to the market and the previous series becomes off-the run. There have been a total of 20 series to date.

Our data set contains daily market pricing information on CDX NA IG indices from MARKIT. Although the first series of CDX (CDX1) was traded from September 2003, MARKIT provides five- and ten-year data from series 5 and seven-year data from series 6. We extract the monthly time series of on-the-run series using the same dates for our option sample. We only use on-the-run series data because the latest series are the most liquid.

In this paper, our period of interest is from October 2005 to September 2008, which corresponds to CDX series 5 to 10. We divide our sample period into two sub-periods: pre-crisis and post-crisis. Pre-crisis sample is from October 2005 to September 2007 (CDX5 to CDX8). Post-crisis sample is from October 2007 to September 2008 (CDX9 to CDX10). Note that our sample period ends before SNAC was introduced. Therefore we do not need to consider changes in the trading convention in the market. In our entire sample, CDX index and all tranches are quoted in terms of spreads and the equity tranche is quoted in terms of upfront amount with 500bps fixed running spread.

2.3.2. Calibration

The parameter values for aggregate consumption, market dividends, and for utility are shown in Table 2.1. To ensure that our model can match consumption data, aggregate market data and data on index put options, we use similar parameter values as in Chapter 1. That is, we choose risk aversion of 3, and EIS equal to 1 (which allows for closed-form solutions), and standard parameters for temporal discounting, normal-times consumption growth, normal-times consumption volatility, and the leverage for the dividend claim. For

convenience, rather than using the multinomial distribution of Barro and Ursua (2008), we assume a power law distribution with a minimum value of 10% and a tail parameter of 7. This is a reasonable approximation to the multinomial distribution (Barro and Jin (2011)), but is slightly more conservative in that the tail probabilities of very bad events are lower. To keep other quantities in the model the same, we make a slight adjustment to the volatility parameters. Otherwise, parameter values for the λ_t and ξ_t processes are the same as in Chapter 1.

Table 2.2 reports parameter values for the individual firm dynamics. Collin-Dufresne, Goldstein, and Yang (2012) estimate the asset beta of the portfolio of firms in the CDX index as 0.56-0.66, the idiosyncratic volatility as 0.179-0.192, and the leverage ratio as 29-37%. To be consistent with their estimates, we set ϕ_i, σ_i, and the leverage ratio equal to 1.456, 0.19, and 32%, respectively. The recovery rate upon default is assumed to be 40% during normal times but 20% during disasters. As recent literature suggests, we set the default boundary much smaller than the leverage ratio. Specifically, we set the default boundary to 60% of the leverage ratio following Collin-Dufresne, Goldstein, and Yang (2012).

Lastly, we assume that the long run mean of the idiosyncratic jump intensity process ($\bar{\chi}$) is 1.5%, the mean reversion (κ_χ) is 0.1, and the volatility parameter (σ_χ) is 0.05. Since we want to capture the possibility of sudden defaults of individual firms using idiosyncratic jumps, the consequences of jumps are severe. That is, we set Z_i so that the firm value drops 75% when idiosyncratic jumps happen.

2.3.3. Choice of state variables

A standard approach to comparing an endowment economy model with the data is to simulate population moments and compare them with data moments. In a model with rare disasters, this may not be the right approach if one is looking at a historical period that does not contain a disaster. An alternative approach is to simulate many samples from the stationary distribution implied by the model, and see if the data moments fall between the

5th and 95th percentile values simulated from the model. For this study, this approach is not ideal either, for two reasons. First, the short length of the CDX time series will likely mean that the error bars implied by the model will be very wide. Thus this test will have low power to reject the model. Second, unlike stock prices which are available in semi-closed-form, and options, which are available up to a (hard-to-compute, but nonetheless one-dimensional) integral, CDX prices must be simulated for every draw from the state variables. Thus the simulation approach is computationally infeasible.

For these reasons, we adopt a different approach. We use three time series from the data to infer time series for our three state variables. We then generate predictions for the remaining quantities in the data based on this series of state variables. We are thus setting up a more stringent test than endowment economy models are usually subject to. Namely, we are asking that the model match not only moments, but the actual time series of variables of interest.

In the model, CDX and CDX tranche prices are determined by all three state variables λ_t, ξ_t, and χ_t while option prices are determined by the first two. For each date, we back out the values of the state variables so that the model matches (1) the five-year 15-30% tranche spread, (2) the five-year 0-3% tranche upfront amount, and (3) the one-month at-the-money (1M ATM) implied volatility. The spread senior tranche is almost entirely driven by catastrophic risk, and so it is a good source of information for rare-disaster intensity λ_t. The equity tranche, on the other hand, is particularly sensitive to idiosyncratic risk χ_t. Finally, ξ_t is a major determinant of stock market volatility , which can be proxied for by 1 month ATM options. Given these state variables, we compute model-implied CDX spreads and option prices.

2.3.4. Results

Our goal in this paper is not only to match CDX spreads, but to do so in a way that is consistent with asset prices more broadly. In Chapter 1, we show that our model, calibrated

at these parameters is capable of matching the mean and volatility of Treasury bills and of excess stock returns. Thus, at these parameters, the model can resolve the riskfree rate, equity premium, and volatility puzzles, even with a low risk aversion for the representative agent. The model generates very little predictability of consumption growth (in samples without disasters there is no consumption growth predictability). Moreover, excess stock returns, are predictable, consistent with the data. In samples without disasters, consumption growth volatility matches the low volatility of the postwar sample in the U.S. Because the pricing of CDX products is closely tied to aggregate volatility, matching the volatility puzzle is particularly important. These results show that our model explains aggregate stock market volatility without counterfactually introducing volatile or predictable consumption, or volatile interest rates.

The link between put options and the pricing of default risk is well-known, so verifying the ability of the model to match option prices is also important. While options are the subject of Chapter 1, we check that this is still the case, given that we use a different algorithm to back out the values of the state variables. Indeed, Figure A.1 verifies that the model can match the implied volatility smirk at the 1-month, 3-month, and 6-month maturities. As Chapter 1 discusses, the model's ability to match option prices arises in part from the fact that it can match the volatility of stock returns. Moreover, out-of-the-money put options are a particularly good hedge against increases in the risk of a disaster, leading implied volatilities to be higher than realized volatilities. These affects persist for 6-month options (Panel C) because the model endogenously generates long-horizon skewness in asset returns.

We now turn to the pricing of CDX products. Figure A.2 shows average spreads for the index as a whole, and for various tranches at maturities of 5, 7, and 10 years (for the equity tranche, we show the upfront amount because the spread is fixed). The model prices the 5-year spreads almost exactly. One one level this is not surprising, as the state variables are chosen to match the upfront amount of the equity tranche and the spread of 15-30% tranche. However, it is not automatic that the model would match the spreads of all

tranches in between, as well as the CDX index as a whole. Moreover, the debate between Coval, Jurek, and Stafford (2009) and Collin-Dufresne, Goldstein, and Yang (2012) pertains to the 5-year spread. Coval, Jurek, and Stafford (2009) find that the spreads of all tranches reported in Figure A.2 except the equity tranche are too low in the data compared with their model. In our model, the 5-year spreads are very close to the data for all tranches.

We go a step beyond Coval, Jurek, and Stafford (2009) and Collin-Dufresne, Goldstein, and Yang (2012) and report data and model spreads for longer-term CDX contracts in Figure A.2. Our model is also able to explain the average spreads for the 7-year contracts. Though the model was not fit to these contracts, the spreads on the total CDX index are matched almost exactly. For 10-year contracts, the model predicts spreads that are slightly higher than in the data, with the difference being most notable for the 10-15% and 15%-30% tranche. However, this difference is small compared to the mispricings reported by Coval, Jurek, and Stafford (2009), and similar in magnitude to the errors of the preferred model in Collin-Dufresne, Goldstein, and Yang (2012) for the 5-year spread.

As Collin-Dufresne, Goldstein, and Yang (2012) show, the prices of CDX contracts change dramatically during the time period of interest. Moreover, results in Coval, Jurek, and Stafford (2009) are (naturally) in reference to the pre-crisis period. The puzzle noted by Coval et al is that average spreads on all but the most junior tranche were too low. When data for the crisis is included, spreads are higher; perhaps this is the reason why we can match average spreads in Figure A.2. To test this possibility, we split the sample into pre- and post-crisis periods. [6] The solid lines in Figures A.3–A.5 show average tranche spreads in the data when the sample is split in this way. The financial crisis lead to a dramatic increase in spreads for all tranches from the mezzanine to senior level, and for all maturities. Consider the 5-year spreads, shown in Figure A.3. For the 15-30% senior tranche, the average spread went from less than 10 basis points to close to 70 basis points. At

[6]The pre-crisis period is defined as ending in September 2007, as in Collin-Dufresne, Goldstein, and Yang (2012). While this may seem an early date for the financial crisis, the time series of CDX spreads do undergo a dramatic change at this time.

the other end of the seniority spectrum, for the 3-7% mezzanine tranche, the average spread increased from about 100 basis points to 500 basis points. Given that the state variables are chosen to match the time series of the senior 5-year tranche, it is not surprising that this data point is matched exactly for both the pre- and post-crises periods. However, the state variables were not chosen to match the other tranches (the equity tranche will be discussed later), but the model fits these very well. The reason the model is able to capture the shift in average spreads is because it allows for time-varying probabilities of catastrophic economy-wide events, as well as idiosyncratic jumps. These possibilities are built into the dynamics of the model. In contrast, Collin-Dufresne, Goldstein, and Yang (2012) assume constant jump intensities but allow the parameters to shift over time to match the data, interpreting these as structural breaks. Their structure, however, assumes that investors do not build the possibility of such breaks into the model.

Figures A.4 and A.5 show 7 and 10-year tranche spreads also increase dramatically, though less than the 5-year spread. This is sensible, economically, as it indicates the beliefs of investors that the stresses firms faced during the crisis were unlikely to be permanent. Such mean reversion in disaster probabilities is built into our model, and explains why our model can capture the change in average spreads for these maturities, despite the fact that the state variables were not chosen based on these maturities at all.

Finally, Figures A.6–A.8 show the actual time series of tranche spreads in the model and in the data. The solid lines, showing the data, illustrate the substantial change that these prices undergo beginning in late 2007. Consider the time series for 5-year contracts (Figure A.6). After a period of relatively low spreads, spreads rose and became much more volatile. The degree to which the activity in the CDX market anticipated the broader financial crisis is interesting. An initial uptick in the level of spreads occurred in late 2007. Spreads fell slightly, but they did not return to their previous levels. Then, beginning in early 2008, spreads rose steadily to several orders of magnitude beyond anything that had been seen prior to September 2007. Spreads fell substantially, only to rise again in late 2008. This

pattern holds for all tranches except for the equity tranche; it was relatively expensive to insure this tranche even early in the sample, and, while the patterns are the same, the difference in the crisis period is much less notable.

Figure A.7 and A.8 show the time series for the 7- and 10-year contracts. Similar patterns are apparent as in the 5-year contracts, with several important qualifications. First, as noted previously, the increase around the crisis was smaller in magnitude for 7-year as compared with 5-year contracts, and smaller still for 10-year contracts. The differences in maturity are most apparent at the junior tranches. Indeed, for the 10-year equity tranche, the crisis appears to have had little effect.

One pervasive narrative of the financial crisis focuses on a massive liquidity disruption that led to widespread mispricing, especially in structured debt products like those based on CDX tranches. This paper is based on another narrative that emphasizes changes in investor beliefs about rare events. These narratives are not mutually exclusive, as many models have emphasized the link between liquidity events and subsequent economic declines. In this paper, however, it is investor beliefs alone that drive prices. Our intention is not to rule out other models, but simply to point out that there are some features of the data that support the view that prices on CDX products accurately reflected investor beliefs at the time. While all the products moved to a certain extent in tandem, we see much less of a change in the longer-term contracts as compared with the shorter-term contracts. This suggests the view of investors that the crisis would ultimately be short-lived. Moreover, senior tranches were far more affected than junior tranches, reflecting the fact that what had changed was primarily a view of catastrophic risk, not the idiosyncratic risk of firms.

Our model formalizes this intuition. A basic test of the plausibility of the model is that we are able, at some value of the state variables, to match the time series of prices for equity and senior tranches for 5-year contracts. Figure A.6–A.8 show that, in addition to these, the model can match the time series of prices for the other tranches, and for other maturities. Though it is not calibrated to the longer-term contracts, the mean reversion that is built

into the model allows it to price these as well. Thus both the low levels of the spreads on senior tranches prior to the crisis, as well as the much higher spreads during the crisis can be explained within a single rational and frictionless framework.

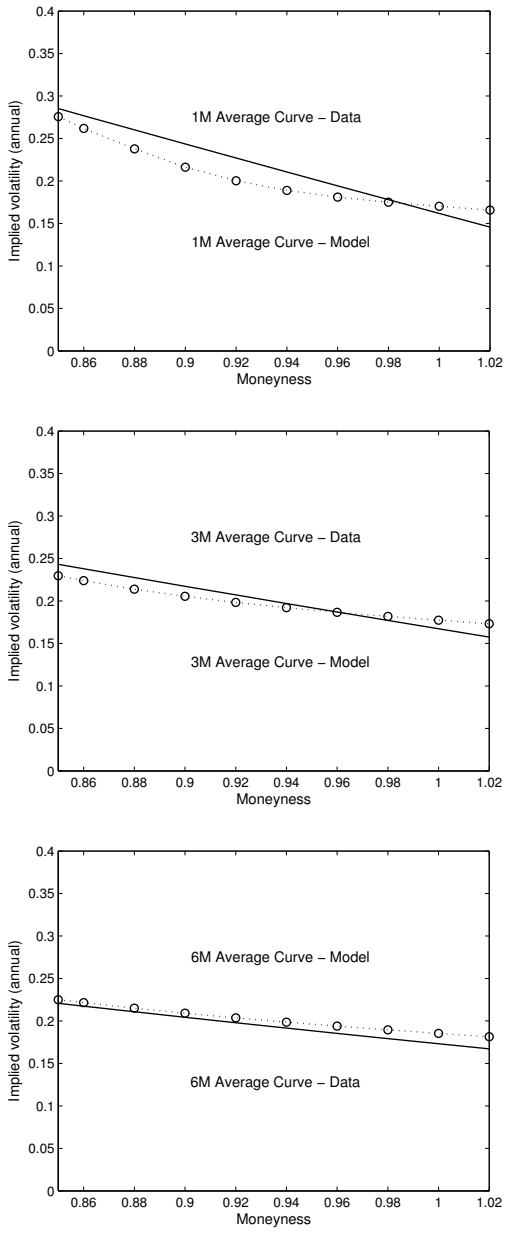
2.4. Conclusion

CDX senior tranches are analogous to extremely deep out-of-the-money options on the US economy because they do not incur any loss until a highly significant number of investment grade firms go into default. In this sense, CDX senior tranche spreads are an important source of information about the probability of catastrophic events. In this paper, we utilize this information to study equilibrium-based asset pricing models. Specifically, we focus on rare disaster models since we believe that the large, infrequent shocks are a natural choice to generate positive prices for extremely bad states.

We extend the two-factor stochastic disaster risk model (Chapter 1) to explain the characteristics of CDX and CDX tranche spreads over different maturities in both the pre-crisis and the post-crisis samples. Our model can explain the equity and option markets along with the CDX market, which emphasizes the importance of beliefs about rare disasters in the asset markets.

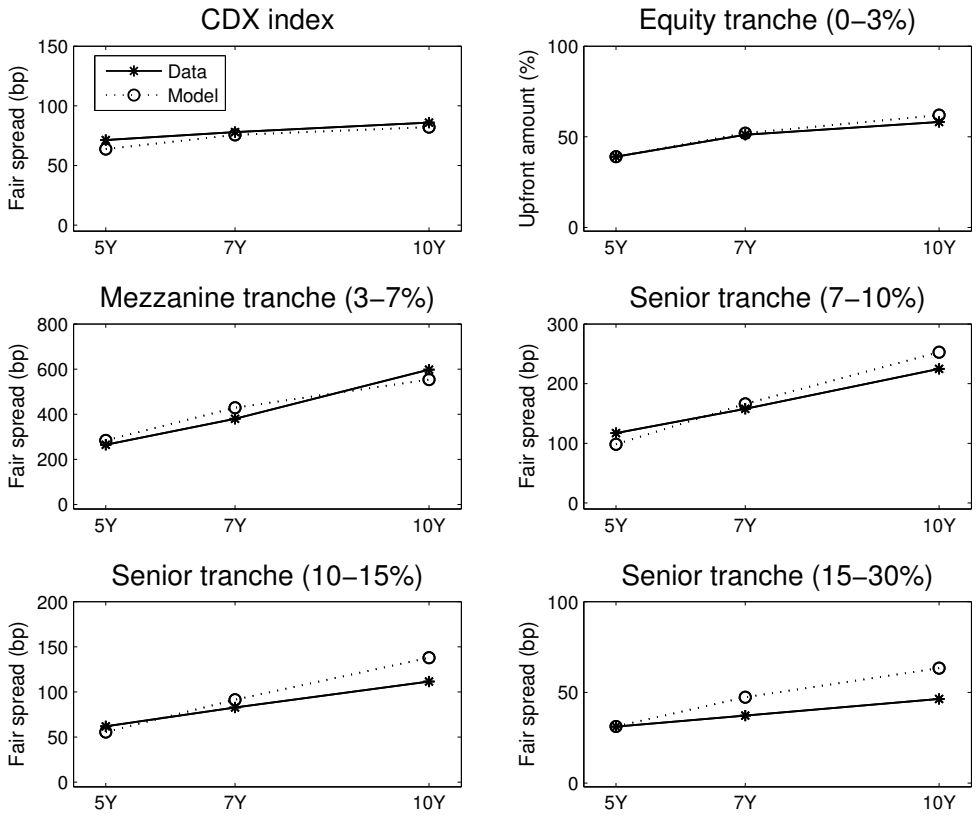
We also point out that senior tranche spreads are very crucial source to see how the market assesses the risk of rare disasters. In this sense, our analysis potentially provides an alternative to the conventional Barro-Ursua calibration of rare disasters.

Figure A.1: Average implied volatilities



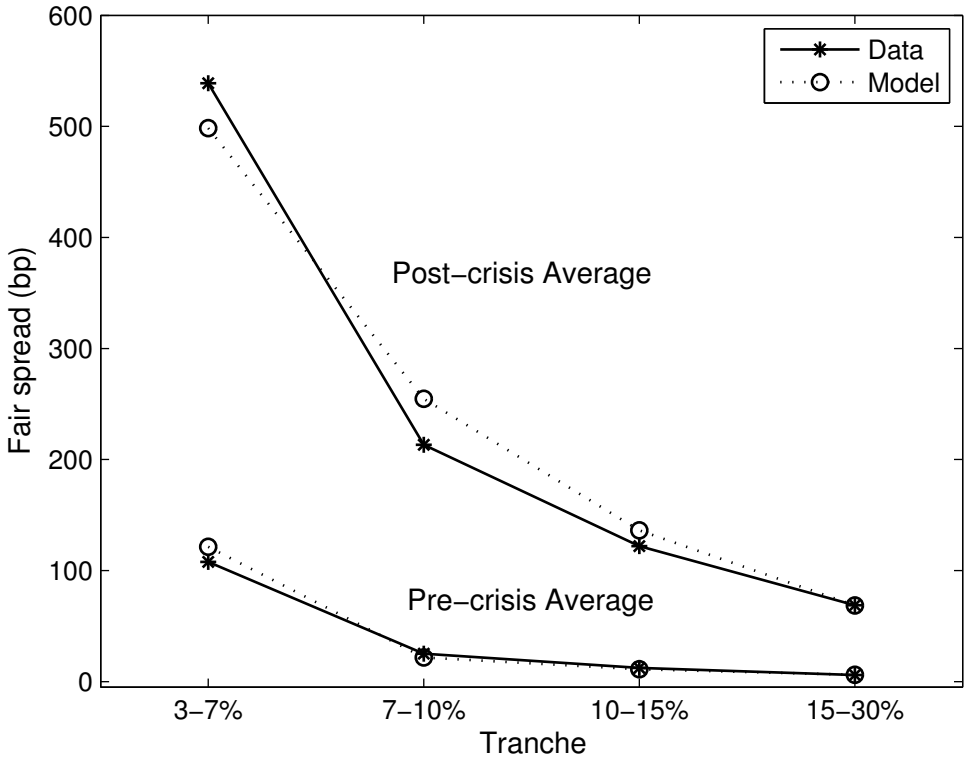
Notes: Average implied volatilities on put options in the data and in the model as a function of moneyness (exercise price divided by index price). Data are monthly from January 1996 to December 2012. In the model, implied volatilities are computed assuming state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the 1 month ATM implied volatility. We compute the implied volatility for each month and take the average.

Figure A.2: Term structure of tranche spreads



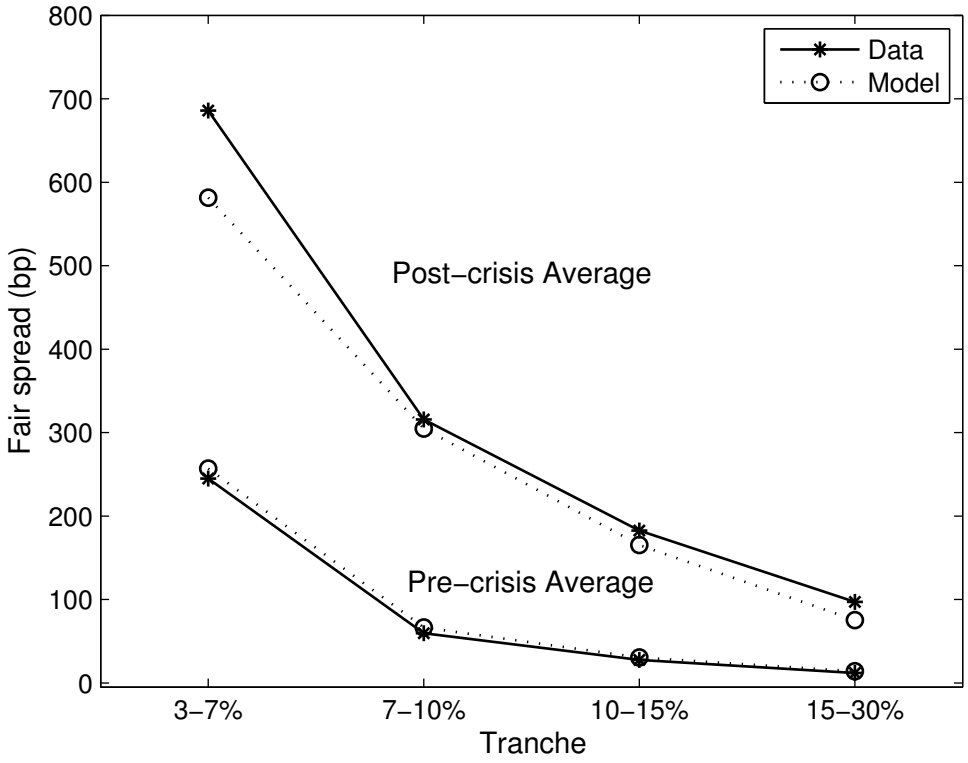
Notes: This figure shows the term structure of historical and model-implied average CDX tranche spreads. We compute the average spreads for five-, seven-, and ten-year maturities. Monthly data start in October 2005 for the five- and 10-year maturities and in April 2006 for the seven-year maturity and end in September 2008. For the equity tranche, we show the upfront amount, assuming a spread of 500 basis points. Model quantities are computed assuming state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the 1 month ATM implied volatility.

Figure A.3: Average CDX tranche spreads (5Y)



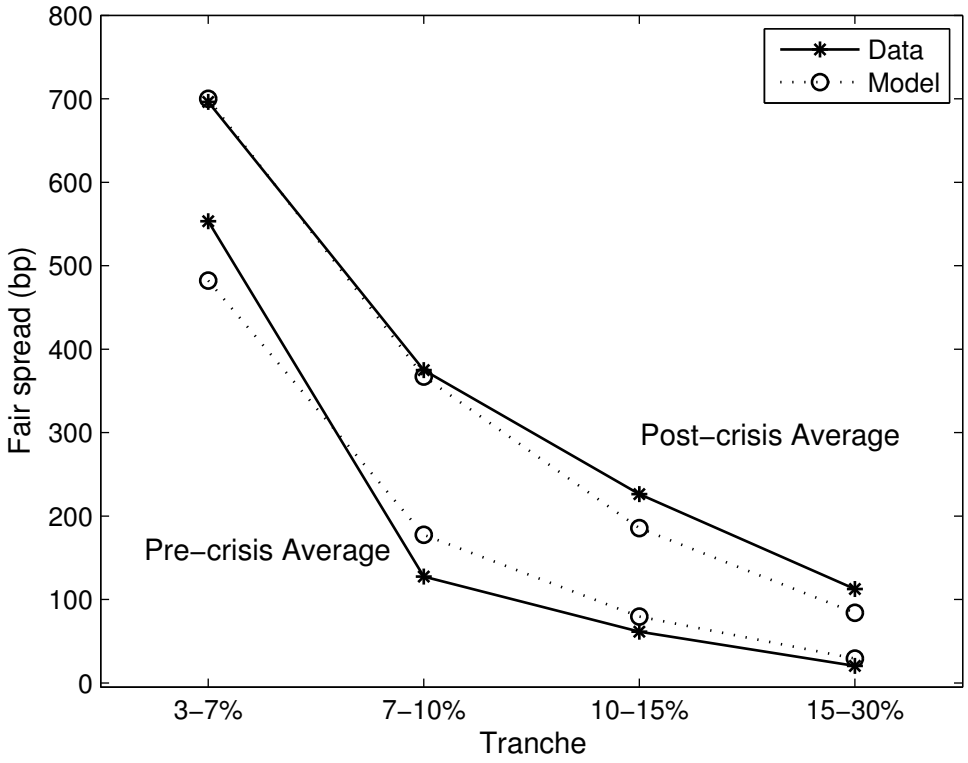
Notes: Historical and model-implied average five-year CDX tranches in the pre-crisis sample and the post-crisis sample. We divide our sample into two sub-periods: pre-crisis and post-crisis. The pre-crisis sample is from October 2005 to September 2007 (CDX5 to CDX8). The post-crisis sample is from October 2007 to September 2008 (CDX9 to CDX10). Model-implied spreads are computed using the state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the ATM implied volatility. We compute the spread for each month and take the average.

Figure A.4: Average CDX tranche spreads (7Y)



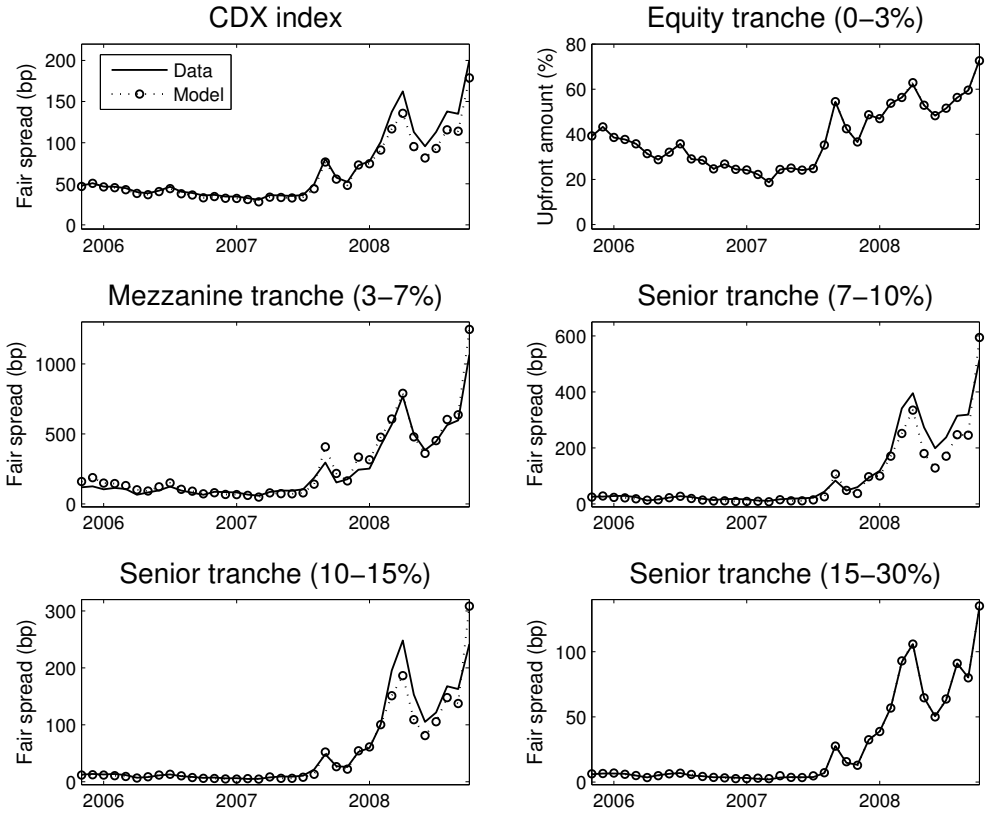
Notes: Historical and model-implied average CDX tranches in the pre-crisis sample and the post-crisis sample. We divide our sample into two sub-periods: pre-crisis and post-crisis. Since the data on seven-year spreads start from April 2006, the pre-crisis sample is from April 2006 to September 2007 (CDX6 to CDX8). The post-crisis sample is from October 2007 to September 2008 (CDX9 to CDX10). Model-implied spreads are computed using the state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the ATM implied volatility. We compute the spread for each month and take the average.

Figure A.5: Average CDX tranche spreads (10Y)



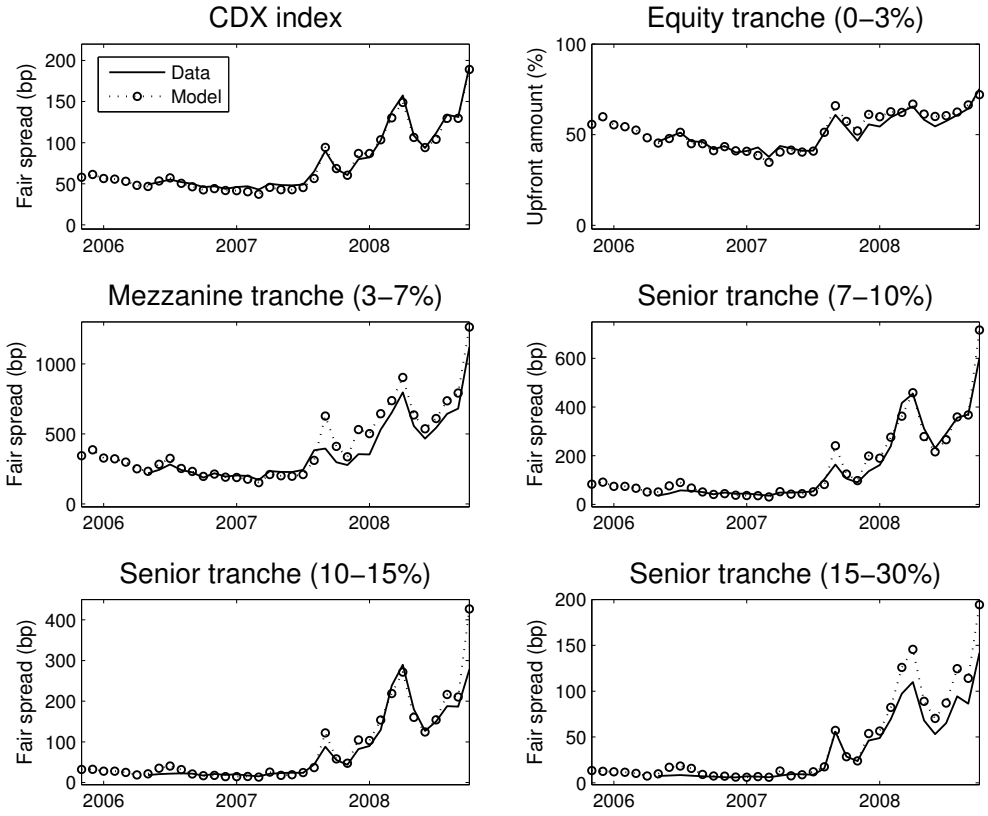
Notes: Historical and model-implied average ten-year CDX tranches in the pre-crisis sample and the post-crisis sample. We divide our sample into two sub-periods: pre-crisis and post-crisis. The pre-crisis sample is from October 2005 to September 2007 (CDX5 to CDX8). The post-crisis sample is from October 2007 to September 2008 (CDX9 to CDX10). Model-implied spreads are computed using the state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the ATM implied volatility. We compute the spread for each month and take the average.

Figure A.6: CDX index and CDX tranches time series (5Y)



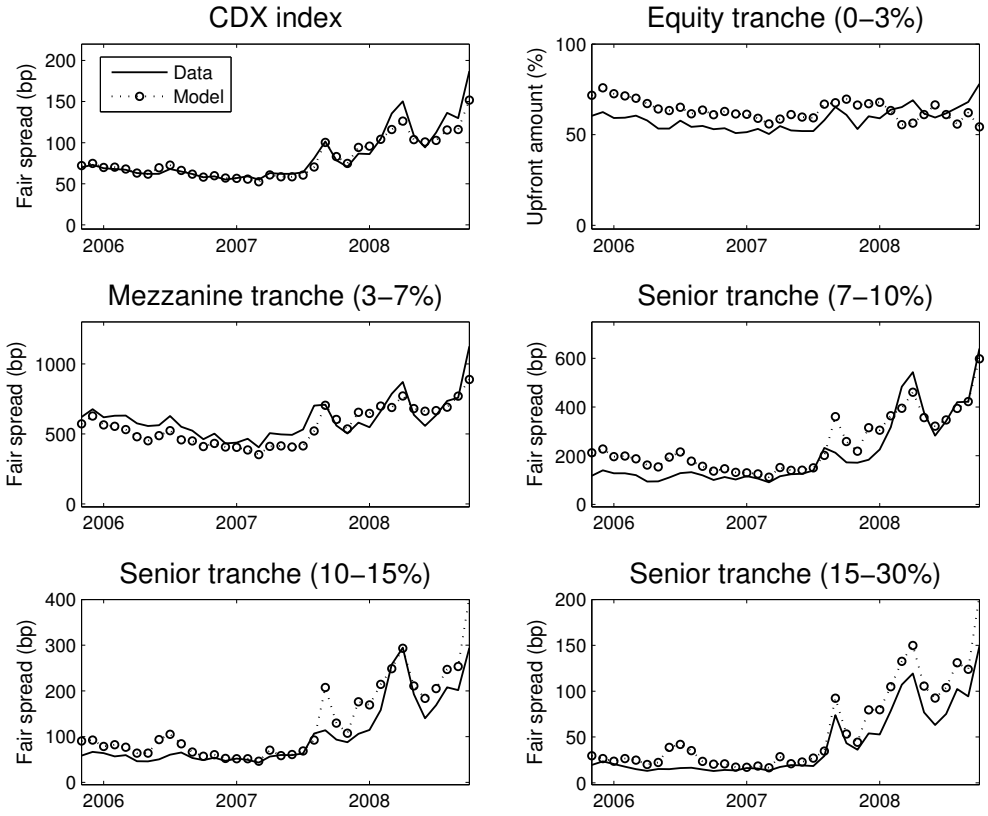
Notes: Monthly time series of five-year CDX and CDX tranche spreads in the data and the model. For the equity tranche we report the upfront amount because the spread is fixed at 500 basis points. Spreads in the data are computed using on-the-run contracts. Data are monthly from October 2005 to September 2008. Spreads in the model are computed using state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the 1-month ATM implied volatility.

Figure A.7: CDX index and CDX tranches time series (7Y)



Notes: Monthly time series of seven-year CDX and CDX tranche spreads in the data and the model. For the equity tranche we report the upfront amount because the spread is fixed at 500 basis points. Spreads in the data are computed using on-the-run contracts. Data are monthly from April 2006 to September 2008. Spreads in the model are computed using state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the 1-month ATM implied volatility.

Figure A.8: CDX index and CDX tranches time series (10Y)



Notes: Monthly time series of seven-year CDX and CDX tranche spreads in the data and the model. For the equity tranche we report the upfront amount because the spread is fixed at 500 basis points. Spreads in the data are computed using on-the-run contracts. Data are monthly from October 2005 to September 2008. Spreads in the model are computed using state variables fit to the time series of the 0-3% tranche, the 15-30% tranche and the 1-month ATM implied volatility.

Table 2.1: Parameter values for the model

Panel A: Basic parameters	
Relative risk aversion γ	3.0
EIS ψ	1.0
Rate of time preference β	0.012
Average growth in consumption (normal times) μ	0.0252
Volatility of consumption growth (normal times) σ	0.020
Market leverage (equity claim) ϕ	2.6
Panel B: Disaster distribution	
Minimum value	0.10
Power law parameter	7
Panel C: λ parameters	
Mean reversion κ_λ	0.20
Volatility parameter σ_λ	0.1643
Panel D: ξ process	
Mean $\bar{\xi}$	0.02
Mean reversion κ_ξ	0.10
Volatility parameter σ_ξ	0.0632

Notes: Panel A shows parameters for normal-times consumption and dividend processes, and for the preferences of the representative agent. Panel B shows parameters for the power law distribution that governs the tail of the consumption distribution. Panels C and D show the parameter values for λ and ξ processes:

$$\begin{aligned} d\lambda_t &= \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} + Z_{\lambda,t}N_{\lambda,t} \\ d\xi_t &= \kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t} \end{aligned}$$

Note that $\bar{\xi}$ is the average level of the probability of a disaster. Parameter values are in annual terms.

Table 2.2: Parameter values for an individual firm

Panel A: Individual firm parameters	
Leverage ϕ_i	1.456
Idiosyncratic volatility σ_i	0.19
Leverage ratio	32%
Default boundary (percent of leverage ratio)	60%
Recovery rate (normal times)	40%
Recovery rate (disaster times)	20%
Panel B: χ process	
Mean $\bar{\chi}$	0.015
Mean reversion κ_χ	0.10
Volatility parameter σ_χ	0.05

Notes: This table reports the parameters for the total payout process on an individual firm. The default boundary in the model is computed as the leverage ratio (0.32) multiplied by 60%. Panel B shows parameter values for the χ_t process, where χ_t is the intensity of idiosyncratic jumps. The process is given by

$$d\chi_t \quad = \quad \kappa_\chi(\bar{\chi} - \chi_t)dt + \sigma_\chi\sqrt{\chi_t}dB_{\chi,t} + Z_{\chi,t}N_{\lambda,t}.$$

All parameter values are expressed in annual terms.

CHAPTER 3 : Correlated Defaults and Economic Catastrophes: Linking the CDS Market and Asset Returns

3.1. Introduction

The bankruptcy of Lehman Brothers illustrates how large of an impact a single significant firm's default can have on the economy. Following this failure, trust and confidence in the financial system plummeted and the real economy experienced a severe contraction. This single event was undoubtedly *bad*. However, consider a scenario in which not only one but twenty, or even thirty, significant firms collectively go into default. This would not just be bad, but *catastrophic*.

Using this simple intuition, I consider economic catastrophes as events in which a number of large firms representing various industries collectively go into default. Over the last 150 years, the U.S. has experienced several events that entailed severe correlated defaults in both the financial and non-financial sectors. One notable example is the Panic of 1893 and the subsequent economic downturn, which resulted in not only a series of bank failures, but also a large number of non-financial corporate defaults.[1] The Great Depression is another well-known example.

Motivated by the observation that massive correlated defaults are not simply a hypothetical situation, I ask the following questions: (1) Is it possible to quantify catastrophic tail risk based on the possibility of massive correlated defaults? (2) Given that catastrophic tail risk is successfully measured, is this risk an important factor for asset pricing? (3) How has the recent financial crisis affected investors' beliefs about future extreme events?

Catastrophic events are, by definition, very rare. Thus, given the relatively short history of economic data, it is difficult to investigate the risk of such events. To address this issue, I use the data on credit default swap (CDS) spreads of 215 firms, each of which directly

[1] According to Giesecke, Longstaff, Schaefer, and Strebulaev (2011) and Giesecke, Longstaff, Schaefer, and Strebulaev (2014), almost 30% of non-financial corporate bonds went into default during this period.

reflects an individual firm's credit risk.[2] Since CDS contracts trade with various maturities, which provides a term structure of CDS spreads, it is possible to obtain comprehensive information about individual firms' default risks.

Based on this information extracted from the CDS spread data, I define the catastrophic tail risk (CAT) measure as the probability that many healthy firms jointly default. However, calculating this joint default probability is a considerable challenge due to difficulties in realistically capturing the co-movements among default risks of multiple firms. To deal with this issue, I develop a model of correlated defaults where the state of the economy is subject to regime-switching: if the economy switches from the normal state to the frailty state, each firm is exposed to an additional risk of default. I derive each firm's default probability as a function of investors' beliefs by assuming that the current economic state (normal vs frailty) and true severity of the frailty state (moderate vs extreme) are both hidden. As a result, the beliefs about the two conditions of the economy serve as the common source of variation in default probabilities, generating default correlations among a large number of firms in a realistic but tractable way. Since the two beliefs disparately impact different horizons, the model is able to account for variations in not only the level but also the slope of CDS term structures.

The resulting CAT measure has three unique properties that other tail risk measures (such as those based on the options market) do not have: (1) a term structure, (2) the ability to capture different extremities of economic tail risk, and (3) the ability to pick up very long-term as well as extremely bad tail risks. These novel characteristics provide further insights into the dynamics of catastrophic tail risk as well as the influence of different tail extremities on asset returns.

The CAT measure offers a useful channel to understand investors' beliefs about future extreme events. For instance, investors' beliefs show marked differences between the pre-crisis (before 2007) and post-crisis (after 2010) periods. In the post-crisis period, I find

[2]The structure of a CDS contract with its pricing is explained in Section 3.5.1.

that investors put much more weight on future extreme events even after the stock market showed signs of recovery. After experiencing the recent financial crisis, investors started to believe that if something bad happens in the future, that event is going to be particularly severe.

Most importantly, by using the CAT measure as a proxy for catastrophic tail risk, I investigate its implications for asset pricing. First, I find that high catastrophic tail risk predicts high future returns for various assets, including stocks, government bonds, and corporate bonds. For example, a 1% point increase in the CAT measure predicts 0.529% higher excess stock returns over the next year. Although the CAT measure is constructed solely based on the CDS market, it generates high predictive power for all three markets, outperforming or driving out classic return predictors.[3]

Moreover, catastrophic tail risk can explain the cross section of stock returns controlling for exposures to the Fama and French (1993) three factors.[4] I sort individual stocks into quintile portfolios based on their sensitivities to catastrophic tail risk at the end of each month, and investigate returns in the next month. I find that the zero-cost portfolio that goes long the first quintile portfolio (most exposed to catastrophic tail risk) and short the fifth quintile portfolio (least exposed to catastrophic tail risk) exhibits significantly positive mean return, CAPM alpha, and Fama-French three-factor alpha. This substantial return dispersion (5.4% per annum) implies that catastrophic tail risk is negatively priced in the cross section of stock returns.

In short, these results consistently indicate that the risk of economic catastrophes should be regarded as an important risk factor to better understand and explain asset markets.

Related literature

[3]Classic return predictors compared with the CAT measure include the log Price-Dividend ratio and default spread for stock returns, the Cochrane and Piazzesi (2005) factor and term spread for government bond returns, and the default spread for corporate bond returns.

[4]The results are robust to adding other control factors such as the momentum factor of Carhart (1997) or the liquidity factor of Pastor and Stambaugh (2003).

There have been several papers that use information contained in the CDS data to make inferences about implied default risk in the economy. Feldhütter and Nielsen (2012) propose a Bayesian Markov Chain Monte Carlo approach to estimate a doubly stochastic model proposed by Mortensen (2006) using the time-series of CDS spreads.[5] As a result, they find that systematic default risk is an explosive process with low volatility. Giglio (2014) exploits information about counterparty risk in CDS contracts to measure pairwise default probabilities between two banks. Based on these pairwise probabilities, he derives the bounds on joint default probabilities of several banks. Benzoni, Collin-Dufresne, Goldstein, and Helwege (2014) develop an equilibrium model where investors have fragile beliefs. Based on the model, they study the level and variation of European sovereign CDS spreads. While these papers focus on understanding the patterns or implied default risk in the CDS data, my goal is to use the CDS data to make inferences about other asset markets.

My work is also related to studies on CDX tranches, prices of which critically depend on default correlations among CDS contracts in the CDX index.[6][7] While the prices of CDX tranches are a potential source of information about systematic risk, these products are subject to debate about mis-pricing: Coval, Jurek, and Stafford (2009) view senior CDX tranches as economic catastrophe bonds and claim that investors traded these products without recognizing substantial systematic risk, which leads to their mis-pricing. In contrast, Collin-Dufresne, Goldstein, and Yang (2012) find no evidence of mis-pricing if the model is constructed and estimated to match the term structure of CDX spreads so that defaults are not backloaded. Separately, the significantly reduced liquidity of tranche products after the recent financial crisis makes the price data of these products less attractive. In my estimation, I circumvent these issues by relying on the CDS data.

My paper also contributes to the literature on rare disasters and tail risk. The rare disaster mechanism has been proposed to account for the large equity premium apparent in the post-

[5] In Section 3.2.4, I discuss the features of my model in comparisons with other typical models, including doubly stochastic models (i.e. conditionally independent models), frailty models, and contagion models.

[6] I provide further information about the CDX index in Section 3.3.1.

[7] Using the prices of CDX tranches, Longstaff and Rajan (2008) study how defaults are clustered.

war data (e.g. Rietz (1988), Barro (2006), and Weitzman (2007)). As extensions, Gabaix (2012), Gourio (2012), and Wachter (2013) also address equity volatility by implementing time variation in disaster risk. The success of these models highlights the importance of tail risk in financial markets.[8] However, the rare nature of economic disasters makes it difficult to empirically test the rare disaster hypothesis solely based on the time series of consumption.

In response to this point, the literature focuses on other data to find evidence of tail risk. Bollerslev and Todorov (2011a) use high-frequency data for the S&P 500 to estimate jump tails in an essentially non-parametric setup. Bollerslev and Todorov (2011b) also propose an Investor Fears index that captures significant time-varying compensations for fears of disasters using high-frequency intraday data and short maturity options.[9] Kelly and Jiang (2014) estimate a tail risk measure from the cross section of stock returns by assuming a power law tail distribution. In contrast, the CAT measure pinpoints catastrophic events by looking at the risk of massive correlated defaults in the economy. The term structure and various tail extremities of the measure provide deeper insights regarding how catastrophic tail risk is perceived and priced by investors, and how it evolves in financial markets.

The remainder of this paper is organized as follows. I propose a model of correlated defaults (Section 2), and based on the model I construct the CAT measure (Section 3). After explaining the CDS spread data (Section 4), I provide the model estimation procedure (Section 5). The estimated results are interpreted with an emphasis on the extracted investors' beliefs (Section 6). Furthermore, I provide the implications of the CAT measure for catastrophic tail risk (Section 7) and asset pricing (Section 8). I then conclude (Section 9).

[8] Bansal and Shaliastovich (2011), Drechsler and Yaron (2011), and Benzoni, Collin-Dufresne, and Goldstein (2011) extend the results of Bansal and Yaron (2004) to an economy with jump risks.

[9] Tail risk implied by options data has been actively studied in the literature. Examples include Bates (2000), Pan (2002), Broadie, Chernov, and Johannes (2007), Bates (2008), and Yan (2011).

3.2. Model of correlated defaults

To capture catastrophic economic tail risk implied by the CDS market, I present a model of correlated defaults. Following the typical intensity approach, each firm's default is modeled as the first jump of a Poisson process with time-varying arrival rate. This arrival rate is usually called *default intensity*, and is the key modeling object that incorporates the correlation structure between multiple firms in my model.

Intuitively, each firm's default intensity can be decomposed into idiosyncratic and systematic components. The idiosyncratic component is entirely firm-specific and follows an exogenous mean-reverting process. On the other hand, the systematic component depends on the state of the economy, which continuously regime-shifts between the *normal* state and the *frailty* state. If the economy switches from the normal state to the frailty state, each firm is exposed to an additional risk of default.[10]

In the model, not only regime shifts but also belief shifts play a crucial role in generating default correlations between multiple firms. This is due to the assumption that investors do not have perfect information about the economy. Specifically, investors do not know exactly (1) which state the economy is currently in (state: normal vs frailty) and (2) how severe the frailty state is (severity: moderate vs extreme). Thus, investors form two beliefs: the probability that the economy is currently in the frailty state and the probability that the true severity of the frailty state is extreme. These two beliefs act as important state variables in the economy, and, coupled with the possibility of future economic regime shifts, determine the systematic co-movement of firms' default probabilities in the economy. Since the two beliefs disparately impact different horizons (short-run vs long-run), belief shifts serve as a key channel through which the term structure of default probabilities changes.

My model is similar to Benzoni, Collin-Dufresne, Goldstein, and Helwege (2014) in that the state of the economy is hidden. However, while the (true) state of the economy is fixed

[10]Each firm is subject to a different magnitude of frailty. A detailed explanation of the setup is provided in Section 3.2.1.

in their model, I assume that it shifts between the normal and frailty states to account for the observation that crises periods exhibit regime-switching patterns (e.g. Gorton (2014)). Furthermore, in my model investors form beliefs not only about the current state of the economy but also the severity of the frailty state, which enables the model to capture the level as well as slope of CDS term structures.

3.2.1. Defaults in the regime-switching economy

Suppose that there are I firms of interest. A default event of a specific firm i is modeled as the first jump of Poisson process N_t^i with arrival rate λ_t^i. That is, N_t^i, or the *default process* for firm i, remains zero unless the firm goes into default, in which case it becomes one.

$$N_t^i = \begin{cases} 0: & \text{no default} \\ 1: & \text{default} \end{cases} \qquad \text{until time } t, \text{ for } i = 1, 2, ..., I.$$

The arrival rate λ_t^i, or the *default intensity* for firm i, represents the firm's risk of default within a very short period of time (between t and $t+dt$). Since the risk of default fluctuates over time, I assume that this intensity is time-varying. Specifically, each firm's default intensity consists of two components: $\lambda_t^i = X_{S,t}^i + X_t^i$.

First, $X_{S,t}^i$ represents the systematic component of firm i's default intensity, which depends on the state of the economy. To reflect the counter-cyclical nature of default risk and to implement the default correlation among multiple firms, I assume that the economy is subject to (continuous-time) regime-switching. The first regime ($\mathcal{S}_t = 0$) represents the *normal* state of the economy, in which each firm is exposed to normal default intensity. The second regime ($\mathcal{S}_t = 1$) represents the *frailty* state. Under this regime, as the name of the state implies, each firm is exposed to an additional source of default as follows:

$$X_{S,t}^i = \begin{cases} \lambda_n^i & \text{if } \mathcal{S}_t = 0 \\ \lambda_n^i + \lambda_f^i & \text{if } \mathcal{S}_t = 1 \end{cases} .$$

In other words, when the economy is hit by a large negative shock (regime-switch from the normal state to the frailty state), the systematic component of firm i's default intensity rises from λ_n^i to $\lambda_n^i + \lambda_f^i$. This captures the notion that crisis periods exhibit regime-switching patterns that lead to substantial jumps in default risks in the economy.[11] The dynamics of regime-switching can be summarized by the infinitesimal generator of the chain G,

$$G = \begin{bmatrix} -\phi_0 & \phi_0 \\ \phi_1 & -\phi_1 \end{bmatrix},$$

which implies that the infinitesimal transition probabilities are given as

$$\begin{aligned} P(\mathcal{S}_{t+dt} = 1 | \mathcal{S}_t = 0) &= \phi_0 dt \\ P(\mathcal{S}_{t+dt} = 0 | \mathcal{S}_t = 1) &= \phi_1 dt. \end{aligned}$$

The second component, X_t^i, represents the idiosyncratic component of firm i's default intensity. Since this component is, by definition, entirely firm-specific, I simply specify this component as an independent Exponential Jump-extended Ornstein-Uhlenbeck (EJ-OU) process, which mean-reverts back to zero, namely,

$$dX_t^i = \kappa^i(\bar{X}^i - X_t^i)dt + \sigma^i dB_t^i + Z_t^i dJ_t^i \qquad \text{for } i = 1, 2, ..., I,$$

where $\bar{X}^i$ is zero, Z_t^i is exponentially distributed with mean ν^i, and J_t^i is a Poisson process with constant intensity ℓ^i. For convenience, I use the following notation:

$$X_t^i \sim \text{EJ-OU}\left(X_0^i, \Theta^i\right) \quad \text{with} \quad \Theta^i = \left[\kappa^i, \bar{X}^i, \sigma^i, \ell^i, \nu^i\right].$$

The idiosyncratic component can have not only positive but also negative values due to the following reason. In my estimated model, the systematic component drives major variations in firms' default probabilities through firms' different sensitivities towards the frailty state:

[11] For example, Gorton (2014) argues in depth that financial crises are regime-switch type events.

a firm's λ_f^i tends to be larger than average if the firm is more sensitive to the business cycle compared to the average firm, while λ_f^i tends to be smaller if the firm is less sensitive. The idiosyncratic component is designed to pick up residual movements of the firm's default probabilities that are not aligned with the average movements of the market. Thus, if a specific firm performs well beyond the level its systematic component implies (in the sense of exceptionally low implied default probabilities), the idiosyncratic component should be negative to pick up this phenomenon, and vice versa.[12] Note that if a firm is not subject to any firm-specific shocks, its idiosyncratic component gradually mean-reverts back to zero ($\bar{X}^i = 0$).

3.2.2. Term structure of default probabilities

Let $P_D^i(t, t+T)$ denote the time-t default probability of firm i up to time $t+T$. Then, by definition, it follows that

$$P_D^i(t, t+T) = 1 - P_S^i(t, t+T), \tag{3.1}$$

where $P_S^i(t, t+T)$ denotes the time-t survival probability of firm i up to time $t+T$. In the following, I derive the expression for this survival probability. Then, the expression for the default probability is simply obtained from equation (3.1).

3.2.2.1. With perfect information

As a starting point, I first calculate the term structure of default probabilities of individual firm i with the assumption that investors are with perfect information about the economy.[13] Let $\mathcal{G}_t$ be the information set which represents perfect information, containing both the current economic state ($\mathcal{S}_t$) and the true severity of the frailty state (λ_f^i). I denote $P_S^i(t, t+$

[12]For example, if a firm is involved in several lawsuits that result in exceptionally high implied default probabilities, this episode has nothing to do with the systematic component of the firm's default intensity; positive realizations of the idiosyncratic component should capture this instead. In Section 3.6.2, I indeed provide a real example of this case.

[13]In contrast, my main model in Section 3.2.2.2 is based on the assumption that investors do not know exactly about the current economic state ($\mathcal{S}_t$) and the true severity of the frailty state (λ_f^i).

$T;\mathcal{G}_t)$ as the time-t default probability of firm i up to time $t+T$ under the perfect information set $\mathcal{G}_t$.

Recall that the default process for each firm is defined as the first jump of a Poisson process. Thus, the event where a firm still survives until time $t+T$ is equivalent to observing zero corresponding Poisson arrival until time $t+T$. If the arrival rate of a Poisson process is constant, say λ, the probability of no arrival between t and $t+T$ is simply equal to $e^{-\lambda T}$. In my setup, however, the arrival rate λ_t^i is stochastic, so this probability is calculated as $E\left[e^{-\int_t^{t+T}\lambda_s^i ds}\,\middle|\,\mathcal{G}_t\right]$ instead. In other words, the survival probability under perfection information can be expressed as

$$P_S^i\left(t,t+T;\mathcal{G}_t\right)=E\left[e^{-\int_t^{t+T}\lambda_s^i ds}\,\middle|\,\mathcal{G}_t\right].$$

During periods when the economy is in the frailty state, each firm is exposed to an additional default intensity λ_f^i. This implies that each firm's default intensity λ_t^i can be decomposed into three parts: $\lambda_t^i=\lambda_n^i+\lambda_f^i 1_{\{\mathcal{S}_t=1\}}+X_t^i$. Thus, in the model, the survival probability is discounted from 1 (i.e. risk-free) due to three risks:

$$\begin{aligned}E\left[e^{-\int_t^{t+T}\lambda_s^i ds}\,\middle|\,\mathcal{G}_t\right] &= E\left[e^{-\int_t^{t+T}\lambda_n^i+\lambda_f^i 1_{\{\mathcal{S}_s=1\}}+X_s^i ds}\,\middle|\,\mathcal{G}_t\right]\\ &= \underbrace{\left[e^{-\lambda_n^i T}\right]}_{\text{normal default risk}}\ \underbrace{E\left[e^{-\int_t^{t+T}X_s^i ds}\,\middle|\,\mathcal{G}_t\right]}_{\text{firm-specific risk}}\underbrace{E\left[e^{-\lambda_f^i\int_t^{t+T}1_{\{\mathcal{S}_s=1\}}ds}\,\middle|\,\mathcal{G}_t\right]}_{\text{frailty risk}}.\end{aligned}\quad(3.2)$$

The last equality holds because the dynamics of idiosyncratic component X_t^i is independent of transitions between regimes. Equation (3.2) suggests that calculating the term structure of survival probabilities (subsequently, default probabilities) boils down to calculating two expectations - (1) the firm-specific risk component and (2) the frailty risk component.

The first expectation (firm-specific risk) is fairly simple to compute. Since X_t^i is in the affine jump-diffusion class, it follows that

$$E\left[e^{-\int_t^{t+T}X_s^i ds}\,\middle|\,\mathcal{G}_t\right] = E\left[e^{-\int_t^{t+T}X_s^i ds}\,\middle|\,X_t^i\right]=e^{\left\{A\left(T;\Theta^i\right)+B\left(T;\Theta^i\right)X_t^i\right\}},$$

where function $A(\cdot)$ and $B(\cdot)$ solve the ODEs of Duffie, Pan, and Singleton (2000). I calculate these two functions in closed form using the formula presented in Chacko and Das (2002). See Appendix A3.1.1 for details.

Though it appears trickier, the second expectation (frailty risk) of equation (3.2) also permits a closed form expression. The key to this derivation is recognizing that the random variable $U_{t,t+T} = \left(\int_t^{t+T} 1_{\{\mathcal{S}_s=1\}} ds\right)$ is the occupation time in the frailty state between time t and $t+T$. Then, it follows that

$$\begin{aligned} E\left[e^{-\lambda_f^i \int_t^{t+T} 1_{\{\mathcal{S}_s=1\}} ds} | \mathcal{G}_t\right] &= E\left[e^{-\lambda_f^i U_{t,t+T}} | \mathcal{G}_t\right] \\ &= M_T\left(-\lambda_f^i | \mathcal{S}_t\right), \end{aligned}$$

where $M_T(\cdot|\cdot)$ denotes the conditional moment generating function of the occupation time $U_{t,t+T}$. According to Darroch and Morris (1968), this function can be derived using a matrix exponential. For details, refer to Appendix A3.2.1. The subscript T signifies that this moment generating function depends only on T (and not t) if the true current state ($\mathcal{S}_t$) is given. Therefore, the survival probability under perfect information (equation (3.2)) is calculated as

$$E\left[e^{-\int_t^{t+T} \lambda_s^i ds} \,\middle|\, \mathcal{G}_t\right] = e^{\left\{-\lambda_n^i T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right) X_t^i\right\}} M_T\left(-\lambda_f^i | \mathcal{S}_t\right).$$

3.2.2.2. With imperfect information: beliefs as state variables

The model used in the empirical analyses incorporates imperfect information into the baseline model.[14] Let $\mathcal{F}_t$ denote the investors' time-t information set, which reflects this point.

First, I assume that investors do not know which regime the economy is currently in. Thus, given $\mathcal{F}_t$, investors form beliefs about the current state. We denote π_t as the probability

[14]The rationale for this extension is provided in Section 3.2.3 and 3.2.4.

investors assign to the frailty state given $\mathcal{F}_t$:

$$\pi_t = P\left(\mathcal{S}_t = 1 | \mathcal{F}_t\right).$$

In addition to the hidden economic regime, investors do not know exactly how severe the frailty state is. For simplicity, the frailty state is either *moderate* ($\mathcal{C} = 0$) or *extreme* ($\mathcal{C} = 1$). The true value of λ_f^i (i.e. extra default intensity under the frailty state) takes λ_{Lf}^i if the frailty state is moderate. On the other hand, it takes λ_{Hf}^i (which is greater than λ_{Lf}^i) if the frailty state is extreme. Since investors do not know the true magnitude (between λ_{Lf}^i and λ_{Hf}^i), they form beliefs about their likelihood. Let ξ_t denote the investors' subjective probability that the true severity is high. That is,

$$\xi_t = P\left(\mathcal{C} = 1 | \mathcal{F}_t\right).$$

With these two beliefs, the survival probability $P_S^i(t, t+T)$, or equivalently $P_S^i(t, t+T; \mathcal{F}_t)$, is calculated by taking the investors' belief-weighted average of survival probabilities under perfect information with four different scenarios.

$$\begin{aligned} P_S^i(t, t+T) &\equiv E\left[e^{-\int_t^{t+T} \lambda_s^i ds} \,\middle|\, \mathcal{F}_t\right] \\ &= e^{\left\{-\lambda_n^i T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right) X_t^i\right\}} F^i(\pi_t, \xi_t; T), \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} F^i(\pi_t, \xi_t; T) \equiv (1-\pi_t) &\left\{(1-\xi_t) M_T(-\lambda_{Lf}^i | \mathcal{S}_t = 0) + \xi_t M_T(-\lambda_{Hf}^i | \mathcal{S}_t = 0)\right\} \\ &+ \pi_t \left\{(1-\xi_t) M_T(-\lambda_{Lf}^i | \mathcal{S}_t = 1) + \xi_t M_T(-\lambda_{Hf}^i | \mathcal{S}_t = 1)\right\}. \end{aligned}$$

Note that investors do not take into account future learning when computing the full-information prices that they average over (Barberis (2000)).

Therefore, from equation (3.1) and (3.3), the default probability is obtained in closed-form as follows:

$$P_D^i(t, t+T) \quad = \quad 1 - e^{\left\{-\lambda_n^i T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right) X_t^i\right\}} F^i(\pi_t, \xi_t; T). \tag{3.4}$$

The term structure of default probabilities is computed by applying equation (3.4) with different values of maturity T.

3.2.3. Role of two beliefs

In the previous section, I show that the two latent beliefs (π_t and ξ_t) play a crucial role in generating the term structure of default probabilities, which subsequently implies the term structure of CDS spreads.[15] To gain a deeper understanding of how these two beliefs affect the term structure of CDS spreads, in Figure A.1, I investigate the comparative statics of the model.[16]

The blue line with circle markers represents the model-implied CDS term structure when the two beliefs (π_t and ξ_t) are all 10%. Taking this blue line as the baseline, I calculate the term structures under two cases: the first one fixes ξ_t but increases π_t to 20% (red line with asterisk markers), while the second one fixes π_t but increases ξ_t to 20% (green line with triangle markers).

In both cases, the level of CDS curves escalates, reflecting higher frailty risk in the economy. However, these two cases exhibit totally different slope behaviors. When π_t increases, short term spreads increase more than long term spreads. This is because the heightened probability of being in the frailty state today strongly affects default probabilities in the near future, but has a relatively small effect on the long term due to regime-switching. Thus, the premium that protection sellers demand for short term CDS contracts increases

[15] In Section 3.5.1, I show that the term structure of CDS spreads is obtained using a model-free approach if the term structure of default probabilities (equivalently, survival probabilities) is given.

[16] In this analysis, I take the estimated model parameters and the dynamics for the average firm in my CDS portfolio, and see the effects of π_t and ξ_t on its CDS term structure. Section 3.5.2 provides a detailed explanation about this average firm.

more significantly compared to that of long term contracts.

On the other hand, when ξ_t increases, short term spreads increase less than long term spreads. In the long run, the economy always has a sizable probability of falling into the frailty state. Hence, a belief shift about the severity of the frailty state matters for long horizons. In contrast, this belief shift might have a smaller effect in the short run because if the economy is in the normal state with high probability, the severity of the frailty state does not matter for default probabilities in the near future. Thus, the premium that protection sellers demand for long term CDS contracts increases more significantly compared to that of short term contracts.

In sum, a shift in π_t creates an inverse relationship between the level and the slope of a CDS term structure: a change in the level is negatively associated with a change in the slope. However, a shift in ξ_t creates a direct relationship: a change in the level is positively associated with a change in the slope. Due to these two opposite responses, the model is able to account for the different behaviors in the level and slope of the CDS market.

3.2.4. Model feature - deviating from doubly stochastic models

Doubly stochastic models refer to a class of models that generate default correlations between different firms through their loadings on observable common factors such as macro or market variables.[17] Since these common factors are all observable, defaults of different firms become independent, conditional on the current information set. Although this assumption provides great tractability to those models, a seminal paper by Das, Duffie, Kapadia, and Saita (2007) finds strong evidence that this widely used assumption is rejected by the U.S. default history data. They highlight that the conditionally independent assumption fails to generate the high level of default clustering found in the data. In response to the shortcomings of doubly stochastic models, the literature has evolved in two directions: *Frailty* vs *Contagion*.

[17] Duffee (1999), Driessen (2005), Duffie and Singleton (1999) are good examples.

The first approach tries to overcome the problem by assuming unobservable common risk factors. Since these factors are not observable, the investors' information set does not contain them. This keeps defaults of different firms dependent even when they are conditioned on the current information set. Duffie, Eckner, Horel, and Saita (2009) find clear evidence for the existence of common latent factors from the U.S. default history data. (They name these unobservable factors "frailty" factors following terminology used in statistics literature.[18])

In contrast, the second approach deviates from the conditionally independent assumption by considering default contagion. By definition, default contagion refers to the phenomenon that at one firms' default, default risks of other firms increase. This approach has intuitive foundations based on general observations in financial markets. Davis and Lo (2001) and Jarrow and Yu (2001) essentially use mutually exciting jumps to model these observations.[19] However, these models are difficult to use for a large pool of firms due to their network nature.[20] Modeling default contagion in a tractable framework remains an ongoing challenge.

These two seemingly distinctive approaches are not entirely separate.[21] In my model, two subjective probabilities (π_t and ξ_t) are latent – as such, they act as frailty factors. Instead of exogenously specifying frailty factors, my model provides a way to economically motivate and generate these factors. In addition, although I do not directly model default contagion using investors' belief updating, abrupt jumps in default probabilities due to default contagion observed in the market can be empirically captured by jumps in the two beliefs in my model. Since investors' beliefs about the aggregate economy serve as a common component of firms' default risks, the model is tractable without having a complicated network

[18] Note that "frailty" here means unobservable factors. When I use this term in other contexts, I am referring to its literal meaning.

[19] Aït-Sahalia, Cacho-Diaz, and Laeven (2013) also model financial contagion using mutually exciting jumps and analyze the model using international equity market data.

[20] For example, we need to consider ${}_{215}P_2 = 46{,}010$ number of channels for 215 firms.

[21] Azizpour, Giesecke, and Schwenkler (2014) examine the potential sources of default clustering and find that both frailty factors and default contagion are important in explaining the recent U.S. default history data, which is consistent with my model.

structure.[22]

3.3. CAT measure - Implied catastrophic tail risk

As I discuss in the introduction, by calculating the joint default probability of many *healthy* firms, one can develop a measure that captures investors' beliefs about the likelihood of extremely catastrophic events. In other words, it is possible to extract a catastrophic tail risk measure in the economy by utilizing the joint default risk implied by the CDS market.

I emphasize that this extraction of catastrophic tail risk is not possible when a small number of defaults are considered. The credit risk of a single firm alone (no matter how large or influential the firm is) does not provide accurate information about systematic risk in the economy. This is because each firm's credit risk is affected not only by systematic risk, but also by its idiosyncratic shocks. High credit risk of a certain firm might reflect high systematic risk, but it also might be attributed to high idiosyncratic risk. On the other hand, the possibility of collective defaults provides clearer information because the shock that causes this disastrous joint collapse of firms must be systematic.[23]

In the following two subsections, I explain how I construct the CDS portfolio that my analyses rely on, and derive the catastrophic tail risk (CAT) measure based on my model.

3.3.1. CDS portfolio construction

Three conditions have to be met in order to construct a CDS portfolio that is appropriate for my analyses. First of all, firms in the portfolio have to be investment grade entities. If the portfolio consists of junk grade entities, it is possible that there may be many defaults

[22]The approach of using beliefs as a common component of default risks among multiple firms is proposed by Benzoni, Collin-Dufresne, Goldstein, and Helwege (2014). The authors directly model investors' beliefs by assuming that investors learn about the hidden but fixed economic state from realized default events and continuous signals in the economy. As a result, at one firm's default, investors increase their beliefs about the bad economic state, creating jumps in default probabilities of all other firms. In contrast, I extract the time series of the two beliefs from the CDS data. For a detailed explanation, see Section 3.5.

[23]One may argue that a large number of defaults can be the result of many firms all having bad idiosyncratic shocks. However, this is very unlikely because idiosyncratic shocks are, by definition, orthogonal to each other. Moreover, even in the case where one firm's really bad idiosyncratic shock initiates a series of defaults in the economy (i.e. contagion), it is difficult to regard it as a bad idiosyncratic shock. Once an idiosyncratic shock becomes contagious, it becomes systematic.

in the portfolio even when there is no large systematic shock. Thus, it is difficult to link the collective defaults to catastrophic tail risk in the economy.

Second, CDS contracts in the portfolio have to be liquid. Extracting a firm's default probabilities from the CDS data (i.e. CDS spread curve) assumes that CDS spreads are solely determined by the firm's credit risk. Therefore, if contracts are not liquid, CDS spreads might contain a certain amount of liquidity premium, which can potentially distort my analyses.

Lastly, the portfolio should be constantly updated. There are cases in which firms start out investment graded, but over time become junk graded or, even worse, go bankrupt. Also, sometimes liquid CDS contracts become less liquid. For these reasons, updating members of the portfolio is crucial.

Fortunately, a portfolio that meets these criteria can be found in the market: Markit publishes the CDX North America Investment Grade (CDX NA IG) index which consists of 125 large U.S. investment grade firms that represent various industries.[24] This index is categorized into 5 sub-indices (Consumer; Energy; Financial; Industrial; Technology, Media, and Telecommunications), each of which occupies roughly 20%. This index is updated as a new series every 6 months, replacing firms that are not liquid or do not satisfy the investment grade. The first CDX NA IG index (series 1) started in October 2003, and ever since the index has been updated biannually in March and September. Each time a new (on-the-run) series is introduced, the previous indices become off-the-run. I construct my CDS portfolio so that it mimics on-the-run CDX NA IG index series.

By inheriting the properties of this index, my portfolio naturally meets all three conditions stated above. The portfolio is composed of relatively healthy firms; none of the firms have ever experienced a default during the last 10 years of the index's history.[25] Granted, some

[24] To be accurate, some of the firms are based in Canada.

[25] There have been two credit events in this portfolio: Fannie Mae and Freddie Mac. However, these events occurred since these two firms were placed into conservatorship by the government. As a consequence, these two firms became default-free entities.

firms that were at some point a part of the portfolio did default. However, since the portfolio is constantly updated every six months, these firms were eliminated from the portfolio before they went into default.

3.3.2. Catastrophic tail risk (CAT) measure

I define the CAT measure as the probability that my CDS portfolio experiences a large loss. Specifically, I construct it as

$$\mathrm{CAT}_t(T;h) = P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i > h\right). \tag{3.5}$$

Since N_{t+T}^i is the default indicator for firm i, the expression $\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i$ represents the loss rate of the CDS portfolio at time $t+T$. That is, this measure calculates the probability that the portfolio has a loss rate larger than a certain threshold h within T years.

I emphasize that the probability that defines the CAT measure is under the risk-neutral measure. In fact, the model itself does not depend on a specific probability measure. However, since extracted individual default probabilities (from the CDS spread data) are under the risk-neutral measure, their joint default probabilities also inherit this probability measure. Hence, the CAT measure represents risk-neutral probabilities.

Equation (3.5) reveals the three important features of the CAT measure that other tail risk measures do not have. First, the CAT measure has a term structure. By varying T of equation (3.5), I am able to obtain catastrophic tail risk measures across different horizons. For instance, if T=5, I can look at catastrophic tail risk for 5-year horizon, if T=10, 10-year horizon, and so on. Taking advantage of the rich information contained in this term structure data, a deeper understanding of the dynamics of catastrophic tail risk can be gained.

Second, with different values of h in equation (3.5), the CAT measure can capture the risk

of tail events with different extremities.[26] From this characteristic, it is possible to examine how extreme of an event investors take into consideration as a main source of risk when they trade assets.

Lastly, the CAT measure has the ability to pick up very long-term and extremely bad tail risks. This is possible because the measure is based on the CDS market, in contrast to other tail risk measures based on, most prevalently, the options market. In general, options with maturities longer than 1 year are fairly illiquid, making it difficult to extract long horizon tail risks from options data. Furthermore, since out-of-the-money options with moneyness smaller than 0.80 are usually not actively traded, tail risk measures based on options cannot distinguish economic catastrophes from relatively moderate market crashes.

I stress that the three properties listed above are unique to my measure. These novel characteristics of the CAT measure better equip me to gain further insights into catastrophic tail risk in a way that have not been achieved before.

3.3.2.1. Joint default probabilities

Computing the CAT measure using equation (3.5) requires developing a method to calculate joint default probabilities that reflect default correlations between firms. In the model, default correlations are generated not only through a belief shift, but also a regime shift. When investors change their beliefs about the true state (π_t) or severity (ξ_t), default probabilities of individual firms fluctuate together. Moreover, the future possibility of economic regime switches generates additional default correlations since default risks of all firms simultaneously rise or drop when regime shifts occur. Thus, calculating joint default probabilities requires considering both of these aspects.

[26] Of course, h has to be significantly large enough because if it is small, there is a higher possibility that the measure is affected not just by systematic risk but by idiosyncratic shocks. Given that the firms in the portfolio have never experienced a default in the past 10 years, a tail event in which more than, say, 15% of the firms go into default can be seen as extremely rare.

By the law of total probability, it follows that

$$P_t\left(\frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} > h\right) = \int_0^T P_t\left(\frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} > h \,\middle|\, U_{t,t+T} = u\right) f_{U_{t,t+T}}(u;\pi_t)du,$$

where $U_{t,t+T}$ represents the occupation time in the frailty state between time t and $t+T$ (Section 3.2.2.2) and $f_{U_{t,t+T}}(\cdot;\pi_t)$ is defined as the probability density function of this random variable when the probability of being in the frailty state at time t is π_t. According to Pedler (1971), this density function is derived in terms of the Dirac delta function and the modified Bessel function as provided in Appendix A3.2.2. By applying the law of total probability one more time with respect to the true severity of the frailty state, the equations become

$$\begin{aligned} P_t\left(\frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} > h\right) = \int_0^T \Bigg\{ & P_t\left(\frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} > h \,\middle|\, \lambda^i_f = \lambda^i_{Lf}, U_{t,t+T} = u\right)(1-\xi_t) \\ & + P_t\left(\frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} > h \,\middle|\, \lambda^i_f = \lambda^i_{Lf}, U_{t,t+T} = u\right)\xi_t \Bigg\} f_{U_{t,t+T}}(u)du. \quad (3.6) \end{aligned}$$

Now, it is important to observe that conditional on the true severity of the frailty state $\left(\lambda^i_f\right)$ and the occupation time in the frailty state $(U_{t,t+T})$, each firm's default risk becomes independent of others. Thus, conditional on those variables, each N^i_{t+T} can be viewed as an independent Bernoulli random variable with event probability $P^i_D(t,t+T|\lambda^i_f, U_{t,t+T}) = 1 - e^{\left\{-\lambda^i_n T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right)X^i_t - \lambda^i_f U_{t,t+T}\right\}}$. If firms are homogeneous, the sum of all these default processes $(\sum_{i=1}^{I} N^i_{t+T})$ follows a binomial distribution. However, I allow firms to be heterogeneous, so there is no simple way to calculate the exact distribution of this sum.[27] Therefore, instead of calculating the exact distribution, I approximate this distribution by using the Central Limit Theorem. This theorem can be used because default processes are independent (of course, conditionally) and $I = 125$ is a relatively large number. Note that I do not rely on the *identical* condition of the classical Central Limit Theorem. In

[27] Mortensen (2006) provides an algorithm to calculate this exact distribution based on a recursive relation. However, this algorithm takes too much computation time to be used in my setup.

fact, the sum still converges to a normal distribution even without this condition as long as some regularity conditions hold (the Lindeberg Central Limit Theorem).[28] The Lindeberg Central Limit Theorem suggests that

$$\sum_{i=1}^{I} N_{t+T}^{i} \,\Big|\, \lambda_f^i, U_{t,t+T} \sim N\left(\sum_{i=1}^{I} P_D^i, \sum_{i=1}^{I} P_D^i(1-P_D^i)\right),$$

where P_D^i is short for $P_D^i(t, t+T|\lambda_f^i, U_{t,t+T})$ for notational convenience. This implies

$$P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^{i} > h \,\Big|\, \lambda_f^i, U_{t,t+T}\right) = 1 - \Phi\left(\frac{h - \frac{1}{I}\sum_{i=1}^{I} P_D^i}{\frac{1}{I}\sqrt{\sum_{i=1}^{I} P_D^i(1-P_D^i)}}\right).$$

With this expression, the joint default probability $P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^{i} > h\right)$ (and subsequently, the CAT measure) is calculated from the 1-dimensional integral in equation (3.6).

3.4. Data

As described in Section 3.3.1, at every point in time, my CDS portfolio is comprised of 125 firms. However, since the list is updated regularly and some firms in the pool experience a spin-off or M&A, my final sample includes a total of 215 firms.[29] I collect the daily time series of CDS spread curves for each of these firms from the Markit database.[30] The sample consists of the data from October 2003 to September 2013.

Markit distinguishes the reference entities of CDS contracts using their RED (Reference Entity Database) code. Using this RED code can pose a problem because firms are given new codes when they change names or convert their legal status (e.g. limited liability company to corporation, or vice versa). Hence, I track the status of each and every firm over time and match the codes so that those representing essentially the same firm are

[28] In Appendix A3.4, I show that my setup indeed satisfies these regularity conditions (Lindeberg's condition).

[29] The sample also includes daily time series of LIBOR-Swap curves, which are conventionally used benchmark interest rates in the CDS market.

[30] Markit's CDS curve consists of 6-month, 1-, 2-, 3-, 4-, 5-, 7-, 10-, 15-, 20-, and 30-year maturities. In my analyses, I use the spreads with 3-, 5-, 7-, and 10-year maturities.

conjoined together. In addition, when firms go through a spin-off or M&A, I find the RED code of the legal successor and match it with that of the predecessor.

3.5. Model estimation procedure

Estimating the model using the CDS pricing data is a considerable challenge for three reasons. First, it requires estimating the dynamics of 215 firms. Since each firm exhibits meaningfully different dynamics, I allow firms to be heterogeneous. Second, I need to estimate 217 latent processes (2 beliefs and 215 idiosyncratic components). Standard linear Gaussian filters (such as the Kalman filter) are not applicable in my setup because CDS spreads are not linear in state variables nor are the state variables themselves Gaussian. Nonlinear filters such as the extended Kalman filter or the Unscented Kalman filter could potentially be used, but the high number of latent processes remains problematic. Lastly, the estimation should ideally fit a CDS term structure (not just a single spread), which puts more burden on the procedure.

I provide an estimation procedure that resolves these issues. The key is to first convert CDS curves into (risk-neutral) survival probability curves and then exploit benefits arising from having a large number of firms with term structure data. I briefly outline my estimation strategy.

1. For each CDS curve (for each date and each firm), I extract the term structure of survival probabilities ($S^i_{t,t+\tau}$). This can be done in an essentially model-free approach. (Section 3.5.1)

2. I consider a hypothetical firm that represents the average of the pool. Due to diversification, the survival probabilities of the firm are not affected by any latent idiosyncratic processes. (Thus, only two latent beliefs affect the survival probabilities.) After expressing two latent beliefs as functions of the average firm's 5-year and 10-year survival probabilities and the regime-switching parameters, I find the parameter values that minimize the root mean square error (RMSE) of survival probabilities. (Section 3.5.2)

3. Conditional on investors' beliefs and the regime-switching parameters, the estimation of each firm's dynamics becomes mutually independent because the only components left are idiosyncratic. With the assumption that 5-year survival probabilities are observed without error, I back out the latent idiosyncratic processes and estimate the dynamics of individual firms using maximum likelihood estimation (MLE). (Section 3.5.3)

I do not rely on a filtering method to estimate latent processes.[31] Instead, following the interest rate term structure literature, I assume that some data are exactly observed and others are observed with errors. (See, for example, Duffee (2002) and Aït-Sahalia and Kimmel (2010).) Specifically, I assume that the individual firms' 5-year survival probabilities are observed without error:

$$\underbrace{S^i_{t,t+5}}_{\text{Data}} = \underbrace{P^i_S(t,t+5)}_{\text{Model}}. \tag{3.7}$$

I choose 5-year maturity survival probabilities to be observed without error as 5-year maturity CDS spreads are the most common and actively traded in the market. Survival probabilities for other maturities are observed with iid Gaussian errors $\epsilon_{i,\tau}$:

$$\forall \tau \neq 5, \quad S^i_{t,t+\tau} = P^i_S(t,t+\tau) + \epsilon_{i,\tau}, \quad \text{where } \epsilon_{i,\tau} \sim N(0, {\sigma_{i,\epsilon}}^2).$$

In the following subsections, I go into greater detail explaining the three steps stated above.

3.5.1. Extraction of survival probability curve from CDS curve

A CDS contract is a bilateral swap agreement designed to transfer the credit risk of a certain reference entity from one party (protection buyer) to another (protection seller). If a credit event of the reference entity occurs, the protection seller compensates the loss

[31] Hence, I do not use belief updating equations in the estimation. As described above, I directly extract two beliefs from 5-year and 10-year survival probabilities of the average firm.

of the protection buyer. In return, the protection seller receives a series of coupons from the protection buyer as the insurance premium until the contract matures or a credit event occurs.

Thus, pricing a CDS contract requires calculating two expected present values: the *premium leg* and *protection leg*. The premium leg is the expected present value of future insurance premiums that the protection seller receives. The protection leg is the expected present value of future protection that the protection buyer pays.

If risk-neutral survival probabilities for all maturities $\{S_{0,t}, t \in [0,T]\}$ are provided, CDS contracts can be priced without relying on a specific model.[32] The inverse relationship can hold as well. That is, if CDS spreads for all maturities are observed, risk-neutral survival probabilities can be bootstrapped directly from spreads. Technically only a discrete set of maturities are observed in the market as I discuss in Section 3.4. Thus, to the extent that an interpolation method is used, this approach is model-free. Note that this conversion should be robust to various interpolation methods since the term structure of survival probabilities is always monotonically decreasing.[33] An unambiguously downward sloping term structure makes the interpolated curve insensitive to interpolation schemes; this conversion can be viewed as an essentially model-free approach.

In practice, market participants have been using a particular interpolation approach as a convention: they assume that survival probabilities are piecewise exponential functions. In other words, survival probabilities between two observed maturities are interpolated using an exponential function. In response to the popularity of this assumption, in February 2009, the International Swaps and Derivatives Association (ISDA) took this assumption and released it as the standard settlement method for CDS contracts.[34] Now, extracting

[32]To be precise, this equation is model-free as long as the recovery rate upon default is assumed to be constant. To prevent my analyses from becoming overly complicated, I take this assumption and use a 40% recovery rate following the market convention for investment grade CDS contracts. Appendix A3.5 provides further details about CDS pricing.

[33]Since the (cumulative) default probability is increasing with maturity, the survival probability is decreasing.

[34]This is called ISDA CDS Standard Model. For details, see www.cdsmodel.com/cdsmodel.

the survival probability curve from the CDS curve with this assumption is accepted as a market convention.

I implement this procedure and convert all CDS curves into their corresponding risk-neutral survival probability curves.[35] Instead of CDS curves themselves, I use these survival probability curves in the estimation.[36] This greatly simplifies the estimation procedure. Note that the T-year CDS spread is affected by all maturities of survival probabilities between 0 and T. I get rid of this complexity by directly using survival probabilities.[37] Since survival probabilities are explicitly expressed in terms of the factors and beliefs in the economy (as shown in equation (3.3)), estimation steps become much more straightforward.

I emphasize that this essentially model-free conversion does not imply that CDS pricing itself is entirely model-free. As I explained above, pricing CDS is model-free once the term structure of survival probabilities is obtained. However, this term structure of survival probabilities critically depends on the underlying model and its state variables. This point is summarized in Figure A.2.

3.5.2. Regime-switching parameters and latent beliefs

I consider a hypothetical firm that represents 125 investment grade firms in the pool. (At each time, my pool consists of 125 representative investment grade firms. Since this pool updates every 6 months, the entire time series data involves a total of 215 firms.) I assume that the survival probabilities of this firm take the average survival probabilities across the 125 firms in the pool. Given the large size of the pool, I assume that through diversification effects, each firm's idiosyncratic component has essentially no effect on this representative

[35] To make sure that this extraction procedure is indeed essentially model-free, I do the following experiments: using my model, I calculate the term structure of survival probabilities and implied CDS spreads. From this generated CDS curve, I extract the survival probabilities using the above model-free procedure and compare it with the original. This experiment shows that the two curves are almost identical.

[36] The model itself does not depend on a specific probability measure. However, since the model is estimated to match the extracted survival probabilities that are under the risk-neutral measure, all parameters and latent processes should also be interpreted as quantities under the risk-neutral measure.

[37] To see why using survival probabilities (instead of CDS spreads) is convenient in the estimation, it is helpful to consider the following analogy: (CDS spreads: survival probabilities) = (yield to maturity : spot rate). Note that yield to maturity is affected by all maturities but spot rate is only affected by a certain maturity.

firm. That is,

$$P_S^{\text{rep}}(t, t+T) \;=\; e^{-\lambda_n^{\text{rep}} T} F^{\text{rep}}(\pi_t, \xi_t; T). \tag{3.8}$$

Although individual survival probabilities are observed with errors (except for 5-year), I assume (by the law of large numbers) that the survival probabilities of the representative firm are observed without error. I use the most actively traded maturities, the 5-year and 10-year, to extract the two beliefs. That is, (π_t, ξ_t) solves

$$\begin{aligned} S_{t,t+5}^{\text{rep}} &= e^{-5\lambda_n^{\text{rep}}} F^{\text{rep}}(\pi_t, \xi_t; 5) \\ S_{t,t+10}^{\text{rep}} &= e^{-10\lambda_n^{\text{rep}}} F^{\text{rep}}(\pi_t, \xi_t; 10). \end{aligned} \tag{3.9}$$

In Appendix A3.3, I show that this system of equations can be transformed into two quadratic equations in π and ξ, which permit closed-form solutions. Mathematically, two sets of (π, ξ) exist due to the quadratic terms. Economically, however, only one set of (π, ξ) remains valid when a limit case of the model is considered. That is, Appendix A3.3 shows that the two latent processes are uniquely determined, and expressed as a function of other parameters and two observed survival probabilities (5-year and 10-year). Thus, I circumvent any need to use a complicated and slow filtering approach in the estimation procedure. To match the term structure of the representative firm's survival probabilities, I estimate the parameters $(\phi_0, \phi_1, \lambda_{Lf}^{\text{rep}}, \lambda_{Hf}^{\text{rep}}, \lambda_n^{\text{rep}})$ by minimizing RMSE.[38] Specifically, I find

$$\text{argmax} \sqrt{\frac{1}{N_t N_\tau} \sum_t \sum_\tau \left(S_{t,t+\tau} - P_S^{\text{rep}}(t, t+\tau) \right)^2}.$$

[38] To be consistent with my earlier assumption that the survival probabilities of the representative firms are observed without error, I do not use MLE. Instead, I minimize the RMSE to fit the entire term structure as closely as possible.

3.5.3. Idiosyncratic estimation

Conditional on investors' beliefs and the regime-switching parameters, the estimation of each individual firm's dynamics becomes mutually independent. Since all of the remaining random components are idiosyncratic (i.e. X_t^i's), I can treat each individual firm's estimation separately. For notational convenience, I stack all survival probabilities of i at time t as a vector and denote it as S_t^i. That is,

$$S_t^i = \left\{ S_{t,t+\tau}^i \quad | \quad \tau = 3, 5, 7, 10 \right\} \quad \forall i.$$

The conditional log-likelihood function regarding firm i is expressed as[39]

$$\log \mathcal{L}^i = \sum_{n=1}^{N} \log P(S_{t_n}^i | S_{t_{n-1}}^i).$$

Thus, calculating $\log \mathcal{L}^i$ requires deriving the expression for $\log P(S_{t+\Delta t}^i | S_t)$.

Although the values of X_t^i are not directly observed, I can extract them from 5-year survival probabilities because they are (assumed to be) observed without error. Equation (3.3) and (3.7) imply that

$$\begin{aligned} S_{t,t+5}^i &= P_S^i(t, t+5) \\ &= e^{\left\{-5\lambda_n^i + A\left(5;\Theta^i\right) + B\left(5;\Theta^i\right) X_t^i\right\}} F^i(\pi_t, \xi_t; 5), \end{aligned}$$

which can be rearranged as

$$X_t^i = \frac{1}{B\left(5;\Theta^i\right)} \left[\log S_{t,t+5}^i + 5\lambda_n^i - A\left(5;\Theta^i\right) - \log\left(F^i(\pi_t, \xi_t; 5) \right) \right]. \tag{3.10}$$

Not only does equation (3.10) provide the values of the latent process X_t^i extracted from 5-

[39] In the online appendix, I provide an expression for the conditional log-likelihood function that also considers observing default processes (N_t^i). Since few firms went into default in the pool and those firms had very high implied default probabilities before default events, including default processes does not change the estimation results much.

year survival probabilities but it also facilitates in deriving the expression for the transition density of $S^i_{t+\Delta t}$. Denoting the vector of survival probabilities of maturities other than 5-year as $S^{i,-5Y}_t$, Bayes' rule implies

$$\begin{aligned} P(S^i_{t+\Delta t}|S^i_t) &= P(\underbrace{S^i_{t+\Delta t,t+\Delta t+5}}_{\text{5-year}}, \underbrace{S^{i,-5}_{t+\Delta t}}_{\text{others}}|S^i_t) \\ &= P(S^i_{t+\Delta t,t+\Delta t+5}|S^i_t) \cdot P(S^{i,-5}_{t+\Delta t}|S^i_{t+\Delta t,t+\Delta t+5}, S^i_t). \end{aligned}$$

The first transition density $P(S^i_{t+\Delta t,t+\Delta t+5}|S^i_t)$ is straightforward to calculate: since observing a 5-year spread is equivalent to observing the value of X^i_t, equation (3.10) implies that

$$P(S^i_{t+\Delta t,t+\Delta t+5}|S^i_t) = P(X^i_{t+\Delta t}|X^i_t) \underbrace{\left| \frac{1}{S^{i,5Y}_{t+\Delta t} B\left(T;\Theta^i\right)} \right|}_{\text{Jacobian}},$$

where $P(X^i_{t+\Delta t}|X^i_t)$ is the transition density of X^i_t.[40] Furthermore, the second transition density $P(S^{i,-5}_{t+\Delta t}|S^i_{t+\Delta t,t+\Delta t+5}, S^i_t)$ is also simply obtained by computing the likelihood of implied measurement errors. Note that conditional on the 5-year survival probability, the value of $X^i_{t+\Delta t}$ is extracted, so the model implied survival probabilities of other maturities can be obtained. Thus, observing $S^{i,-5}_{t+\Delta t}$ is equivalent to observing their pricing (or measurement) errors $\epsilon_{i,\tau} = S^i_{t,t+\tau} - P^i_S(t, t+\tau)$. Since measurement errors are iid by assumption,

$$P(S^{i,-5}_{t+\Delta t}|S^i_{t+\Delta t,t+\Delta t+5}, S^i_t) = \prod_{i=1}^{I} \prod_{\tau \neq 5} \frac{1}{\sigma_{i,\epsilon}\sqrt{2\pi}} \exp\left(-\frac{\epsilon^2_{i,\tau}}{2\sigma_{i,\epsilon}}\right).$$

This approach of deriving the transition equation is analogous to Duffee (2002) and Aït-Sahalia and Kimmel (2010) for interest rate term structure estimation.

[40] In Appendix A3.1, I derive this transition density of X^i_t in exact (A3.1.3) and approximate (A3.1.4) forms.

3.6. Interpretation of estimation results

As described in Section 3.5.2, I first estimate the regime-switching parameters and latent beliefs by using the term structure of the average firm's survival probabilities. The estimated transition rate from the normal state to the frailty state (ϕ_0) is 0.0824 and the one from the frailty state to the normal state (ϕ_1) is 0.0884. The time series of latent beliefs are also displayed in Figure A.3: the red solid line represents the probability that the economy is in the extreme frailty state ($\pi_t \xi_t$) and the green dashed line represents the probability that the economy is in the moderate frailty state ($\pi_t(1 - \xi_t)$).

3.6.1. Extracted time series of investors' beliefs

As can be seen in Figure A.3, the model together with the CDS data extracts and interprets investors beliefs in a very intuitive, yet intriguing way. Not so surprisingly, the figure shows that before the financial crisis, investors believed that the probability that the economy was in the frailty state was near 0. This probability started to rise in mid-2007 and sharply increased from 2008 to 2009. After 2010, the probability of the frailty state returned to a lower level.

Figure A.3 provides a fuller account of investors' beliefs. Between 2007 and 2008, accumulated bad news and concerns about subprime mortgages raised investors beliefs about the frailty state, which led to an increase in both lines. An interesting shift in patterns can be observed in March of 2008 when one of the largest investment banks in the U.S., Bear Stearns, collapsed. At this event, investors realized that there was a higher chance that not only was the economy in the frailty state but also that this frailty was very extreme. This shift in beliefs is reflected in the increase in the red line and the decrease in the green line. Subsequently, Bear Stearns was purchased by JP Morgan with the support of the Federal Reserve. This allowed investors to update their beliefs accordingly by abruptly decreasing the probability of the extreme frailty and in turn increasing the probability of the moderate frailty. In sum, at the risk of Bear Stearns' default, investors increased their beliefs about

the extreme frailty, but after observing its rescue, they revised those beliefs by putting more weight on the moderate frailty.

A second interesting pattern can be observed in September of 2008 when Lehman Brothers filed for Chapter 11 bankruptcy protection. After experiencing this bad event, investors started to believe with higher probability that the economy was in the frailty state. Furthermore, they started to believe with conviction that the frailty they were experiencing was very extreme. Thus, from this time point, the red line shows a sharp increase while the green line shows a steep decline.

The red line representing the extreme frailty peaks at December of 2008 and starts to drop from March of 2009 when S&P 500 hit its 13-year low. On the other hand, the green line representing the moderate frailty dips in December of 2008 and begins to rise up until mid-2009. This pattern has a clear interpretation: after the failure of Lehman Brothers and the subsequent credit freeze, investors strongly believed that the economy is in the extreme frailty state in which there is a very high chance of many firms collectively going into default. (This is evidenced by the increase in the red line and decrease in the green line.) However, due to the Fed's and government's series of actions, investors gradually reduced their beliefs about the extreme frailty. Although investors still believed that the economy was in the frailty state, they began to think that this frailty was not as severe as they thought. (This translates into the decrease in the red line and increase in the green line.)

The last interesting pattern is observed during the post-crisis period (after 2010). Compared to the crisis period, this period exhibits much lower (implied) default probabilities. However, it can be observed that the red line still maintains a significant level compared to the green line. This can be interpreted as the following: investors think that the economy is no longer in the frailty state but rather is most likely in the normal state. Nevertheless, having experienced the financial crisis, investors now believe that if something bad happens in the future, that event is going to be very severe. In short, after the crisis, investors put much

more weight on the extreme frailty.

What aspects of the model and data make it possible to separately identify the two beliefs? Clearly, it is impossible to do so relying on a single time series, such as stock market valuation ratios. I emphasize that the term structure of the CDS market plays a crucial role in disentangling the two beliefs. To illustrate this point, I calculate the implied CDS curves of the hypothetical firm representing the average firm in the portfolio, and plot the level (5-year) and the slope (10-year minus 3-year) of these curves in Figure A.4.

At first glance, the level and slope appear to have an inverse relationship: during the crisis period (2008-2010), the level surged and the slope plunged.[41] However, this relation fails to hold in the pre-crisis and the post-crisis (after 2010) periods. Specifically, both the level and slope were higher in the post-crisis period. These different relations between the level and slope signify that the term structure data is driven by at least two factors. In my model, two latent beliefs π_t and ξ_t act as such factors and do an excellent job in explaining the entire term structure series because π_t and ξ_t have differential impacts on the CDS term structure as described in Section 3.2.3. Due to this capability, the estimated time series of the two latent beliefs inherit the rich information contained in the CDS term structure. Furthermore, as can be seen in Figure A.5, the estimated model fits the data extremely well for all maturities.

3.6.2. Individual firm dynamics and cross section of CDS spreads

Each individual firm's survival probabilities (equivalently, default probabilities or CDS spreads) exhibit its distinctive patterns in the data. These different behaviors between various firms are mainly captured by two components in the model: (1) different frailty levels in the frailty state (λ^i_{Lf} and λ^i_{Hf}) and (2) different realizations of idiosyncratic components (X^i_t). Figure A.6 explains the role of these two components using two firms as examples: ConocoPhillips and Transocean.

[41] Actually, the curve was inverted during this period, which made the slope negative. That is, shorter maturity spreads were higher than longer maturity spreads.

The top left panel of Figure A.6 displays the time series of ConocoPhillips' 5-year survival probabilities (red dashed line) together with the time series of the average firm's survival probabilities (blue solid line). This comparison shows that ConocoPhillips has a smaller-than-average sensitivity towards the frailty risk in the economy: the survival probabilities not only maintain a lower level, but also fluctuate with smaller amplitudes compared to the average firm. In my model, this pattern is captured by the small values of λ^i_{Lf} and λ^i_{Hf}. The values of λ^i_{Lf} and λ^i_{Hf} for ConocoPhillips are estimated as 0.0028 and 0.0642, which are roughly 19% and 53% of $\lambda^{\text{rep}}_{Lf}$ and $\lambda^{\text{rep}}_{Hf}$ for the average firm. Since these small values of two frailty levels explain almost all the variations in the firm's survival probabilities, the path of its idiosyncratic shocks (X^i_t) is extracted as small values, sticking around zero (bottom left panel).

On the other hand, the time series of Transocean's 5-year survival probabilities (top right panel) shows relatively high sensitivity towards the frailty risk: the values of λ^i_{Lf} and λ^i_{Hf} are 0.026 and 0.1472, which are roughly 170% and 120% of $\lambda^{\text{rep}}_{Lf}$ and $\lambda^{\text{rep}}_{Hf}$ for the average firm. However, it is difficult to explain the entire pattern of Transocean's survival probabilities solely based on the two frailty levels. This is because this firm exhibits exceptionally low survival probabilities (i.e. very high default probabilities) during the post-crisis period (after 2010). For example, around mid-2010, Transocean's 5-year survival probability decreased to 58%, which is even lower than the level at the peak of the crisis. Since the market, on average, shows much higher survival probabilities during the post-crisis period compared to the crisis period, such a pattern is clearly attributable to the firm's idiosyncratic shocks: the bottom right panel shows that the firm has experienced large idiosyncratic shocks during the post-crisis period. In fact, during this time period, Transocean - an offshore drilling contractor - was involved in a series of lawsuits because one of its drilling rigs caught fire, resulting in an oil spill off the Gulf of Mexico.

3.7. Implications for catastrophic tail risk

Equation (3.5) implies that the CAT measure has two dimensions: horizon (T) and extremity (h). By varying T, the CAT measure represents catastrophic tail risks of different horizons. In addition, when different values of h are chosen, the CAT measure captures various extremities of catastrophic tails. These two dimensions allow the CAT measure to have a surface structure for each given date. That is, the CAT measure consists of a time series of *CAT surfaces*, each of which contains daily information about catastrophic tail risks with various horizons and extremities.

Due to these two dimensions, each CAT surface has two characteristics. First, given a fixed horizon, a CAT surface is always downward sloping in extremity:

$$\mathrm{CAT}_t(T; h_1) \geq \mathrm{CAT}_t(T; h_2), \quad \text{if} \quad h_1 < h_2.$$

This property is straightforward because the CAT measure represents the upper tail of the joint default probability. For instance, CAT 15 (the CAT measure with 15% extremity) represents the probability of having more than 15% defaults in the pool. That is, no matter how many defaults occur, the probability of these defaults is captured by CAT 15 as long as they exceed the 15% threshold. Thus, CAT 15 is always greater than or equal to, say, CAT 25 (of course, given that they have the same horizons). Note that CAT 25 is equal to the sum of CAT 25 and the probability of having defaults between 15% and 25%.

Second, given a fixed level of extremity, a CAT surface is always upward sloping in horizon:

$$\mathrm{CAT}_t(T_1; h) \leq \mathrm{CAT}_t(T_2; h), \quad \text{if} \quad T_1 < T_2.$$

Since the (cumulative) default probability with a longer maturity is always higher than that with a shorter maturity, this property is also straightforward.

Hence, due to these two properties, all CAT surfaces are monotonically increasing in the

directions of increasing horizon and decreasing extremity. This is not to say that CAT surfaces always look the same. In fact, despite this monotonicity, the patterns of CAT surfaces look quite different over time. In Figure *A*.7, I provide CAT surfaces at different dates, each of which represents one of the three periods: (1) pre-crisis period (March 1st, 2004), (2) crisis period (March 2nd, 2009), and (3) post-crisis period (March 1st, 2014).[42]

As can be seen in the first panel, the CAT surface during the pre-crisis period was relatively flat and maintained a low level. However, the financial crisis lifted the entire surface to a very high level (second panel). Specifically, out of the entire surface, the part with low extremity and long horizon was noticeably raised with greater magnitude while the part with high extremity and low maturity was not raised as much. This results in a somewhat S-shaped CAT surface. After the crisis, the surface came back to a lower level resembling the pattern in the pre-crisis period. However, compared to the pre-crisis period, the surface is much steeper, especially in the direction of the horizon axis.

To see how CAT surfaces have evolved over time, I plot the time series of CAT 15, CAT 20, and CAT 25 with 3-year, 5-year, 7-year, and 10-year horizons in Figure A.8. In general, all the time series seem to show similar patterns in that they maintained a low level before the crisis, peaked during the crisis, and came back to a lower level after the crisis. However, careful observation reveals that these time series with different horizons and different extremities exhibit meaningful differences in their patterns, which provide useful information for understanding catastrophic tail risk in the economy. In the following two subsections, based on this information, I investigate the distribution and the term structure of catastrophic tail risk during the pre-crisis, crisis, and post-crisis periods.

[42]In this section, when I mention the pre-crisis, crisis, and post-crisis periods, I am referring to these specific dates. This specific choice of dates does not affect the results of my analyses.

3.7.1. Distribution of catastrophic tail risk

Based on the CAT measure, it is possible to extract the (risk-neutral) distribution of catastrophic tail events. According to the definition of the CAT measure, it follows that

$$\mathrm{CAT}_t(T; h_1) - \mathrm{CAT}_t(T; h_2) = P_t\left(h_1 \leq \frac{1}{I}\sum_{i=1}^{I} N^i_{t+T} < h_2\right),$$

which is the probability mass of catastrophic events having defaults between h_1% and h_2%. This implies that the probability mass function (or histogram) of catastrophic tail events can be obtained by simply taking the difference between two adjacent CAT values. Hence, for each day, I depict the distribution of catastrophic tail risk using probabilities of 10-15%, 15-20%, 20-25%, 25-30%, and 30-35% defaults. Although all intervals represent extremely rare and disastrous events, the first interval (with 15-20% defaults) and the last interval (with 30-50%) can be regarded as the most moderate state and the most extreme state, respectively. Figure A.9 displays such distributions (with 10-year horizon) in the pre-crisis, crisis, and post-crisis periods.

The pre-crisis period (on the top left panel) presents a fairly standard-shaped catastrophic tail distribution. The most moderate state occupies the highest probability mass and this probability mass diminishes as the state gets more extreme, which generates an inverted J-shape. During the crisis period (on the bottom panel), this inverted J-shaped pattern becomes more prominent: the extreme states (such as 30-35% and 35-40% intervals) still maintain the pre-crisis period level even after the crisis, but the probabilities of moderate states (such as 15-20% and 20-25% intervals) double, which makes the distribution steeper compared to the pre-crisis period. From this observation, it is possible to presume that the escalated catastrophic tail risk perceived by investors during the recent financial crisis is attributable to the increased possibility of relatively moderate catastrophic events, at least around March, 2009.

Another intriguing finding can be obtained when the distributions of the pre-crisis and

post-crisis periods are compared. In the post-crisis period, the overall catastrophic tail risk (measured by summing the probabilities of all intervals) has returned to roughly the same level as the pre-crisis period. However, the shape of the distributions differs significantly between these two periods. In contrast to the inverted J-shape pattern in the pre-crisis period, the post-crisis period exhibits a flat distribution.[43] This implies that investors put much less weight on the possibility of relatively moderate catastrophic events, but instead, put much more weight on the possibility of even more extreme events. Recall that this result is consistent with what can be learned from the time series of the two latent beliefs: after the financial crisis, investors started to believe that if something bad happens in the future, that event is going to be very severe (Section 3.6.1). I emphasize that this insight cannot be acquired from other tail risk measures based on the stock or options markets because by their very nature they are not able to identify these extremely bad and rare events.

3.7.2. Term structure of catastrophic tail risk

Figure A.10 displays the term structures of catastrophic tail risk based on CAT 15 (top left panel), CAT 20 (bottom panel), and CAT 25 (top right panel). In each panel, I put the three term structures that represent the pre-crisis (green dot-dashed line with triangle markers), crisis (red dashed line with asterisk markers), and post-crisis (blue solid line with circle markers) periods.

In all three panels, the curves corresponding to the three time periods show the same order patterns in terms of level. That is, the crisis period curve is always positioned at the top, the post-crisis curve at the middle, and the pre-crisis curve at the bottom. This implies the somewhat obvious fact that for all horizons, the level of catastrophic tail risk was highest during the crisis period and lowest during the pre-crisis period.

[43] It can be seen that the probability mass in the post-crisis period is increasing as the state becomes more extreme. This is because these probabilities are under the risk-neutral measure. Although the actual (or physical) probability decreases as the state becomes extreme, the risk-neutral probability can increase due to risk-adjustments.

Beyond a simple comparison between the levels of the curves, more can be learned from their shapes. First of all, the term structure exhibits a very steep slope for the post-crisis, compared to that of the pre-crisis. Thus, although in the short horizon there seems to be little difference between pre-crisis and post-crisis periods, a much higher catastrophic risk is expected in the long run since catastrophic risk increases with much greater speed as the horizon increases. This is also consistent with the results mentioned in Section 3.6.1 and 3.7.1. After the crisis, investors no longer believe that the economy is in the frailty state, which lowers short horizon catastrophic risk to the pre-crisis level. Nevertheless, investors tend to believe that if something bad occurs in the future, that event will be very severe. Since the economic state is subject to regime-switching, even though the economy may currently be in the normal state, there is a significant probability that it may shift to the frailty state in the long run. Thus, taking this into account, long horizon catastrophic risk maintains a higher level in the post-crisis compared to the pre-crisis period.

While the pre-crisis and post-crisis curves always resemble a straight line, the curve for the crisis period shows a somewhat distinctive shape that varies based on extremities. After the 3-year horizon, the crisis period exhibits a convex curve for CAT 15 and CAT 20, but for CAT 25, the curve is concave. From these observations, it can be inferred that the term structure of catastrophic risk not only experiences changes in level and slope, but in curvature as well.

To investigate the evolution of term structures in more detail, I perform a principal component analysis (PCA) on each of CAT 15, CAT 20, and CAT 25. Figure A.11 shows the loadings of the CAT measure on three principal components.

The blue solid line with circle markers, the red dashed line with asterisk markers, and the green dot-dashed line with triangle markers each represent the loadings on the first, second, and third principal components respectively. What each principal component represents can be inferred from the patterns of loadings across different horizons. First of all, the blue line shows a flat curve with positive values. This means that a positive innovation to the first

principal component raises the catastrophic risk for all horizons by a similar amount. Thus, the first factor can be seen as the *level* factor that causes a parallel shift in the term structure. Next, the red line associated with the second principal component is monotonically upward sloping. Hence, a positive innovation to the second component decreases short horizon catastrophic risk but increases long horizon risk. In other words, this component can be seen as the *slope* factor that determines the steepness of the term structure. Lastly, the green line representing the third principal component is tent-shaped. Therefore, a positive innovation to this component increases the risk for intermediate horizons, but decreases the risk for short and long horizons. Due to this characteristic, the third principal component can be seen as the *curvature* factor influencing how bent the curve is.

As expected, a principal component analysis confirms that the three main factors affecting the term structure are level, slope, and curvature. In fact, these three factors explain nearly all of the variation in the term structure. In the case of CAT 15, for example, the level factor alone accounts for 96.06% of the total variance, and when the slope factor is added, 98.95%, and with all three factors, 99.95%. In sum, the principal component analysis shows that all changes in the term structure can be parsimoniously abbreviated into three factors.

3.8. Implications for asset returns

In this section, using the rich information contained in the CAT measure, I investigate asset pricing implications of catastrophic tail risk. As results, I show that catastrophic tail risk (1) predicts future returns in various markets, and (2) is negatively priced, generating significant dispersion in the cross section of stock returns. In the following subsections, I discuss each aspect in detail.

3.8.1. Properties of the CAT measure

Before exploring return predictabilities based on the CAT measure, I take a look at how individual CAT measures with various horizons and extremities are correlated with one another as well as how these measures are related with other financial variables. Table 3.1

presents a correlation matrix of 10-year horizon CAT 15, CAT 20, CAT 25, and CAT 30. In general, each individual CAT measure is highly positively correlated with each other, and especially, CAT measures with adjacent extremities show even stronger correlations.

Table 3.2 displays the correlations of individual CAT measures with the same extremities but different horizons (CAT 15 and CAT 25). This too shows that CAT measures of different horizons are highly correlated and that the more adjacent the horizon, the stronger the correlations. In addition, correlations among multiple horizons generally seem to be higher for CAT measures with lower extremities (here, CAT 15).

The correlations between individual CAT measures and three classic stock return predictors are displayed in Table 3.3. Predictors include the log Price-Dividend ratio (log P/D; a well known long horizon return predictor) as well as the variance premium (a well known short horizon return predictor). Since the CAT measure is constructed using credit default probabilities, for comparative purposes, the default spread (also a long horizon return predictor) is included as well.[44]

First, the individual CAT measures tend to exhibit relatively high negative correlations with the log P/D ratio but high positive correlations with the default spread. This makes sense since high levels of the CAT measure imply bad economic times, which corresponds to low pricing in the stock market and high level of corporate default probabilities. On the other hand, the individual CAT measures have very low correlations with the variance premium.

Table 3.4 shows the persistence (monthly AR(1) coefficients) of individual CAT measures together with other stock return predictors. In my sample period (Oct. 2003 to Sep. 2013), the log P/D ratio and default spread display high levels of persistence, 0.971 and 0.958 respectively. Individual CAT measures are also fairly persistent; the persistence of (10-year) CAT 30 is 0.908 and as tail extremity decreases, persistence increases such that at

[44]For detailed discussions about these classic predictors, refer to Fama and French (1988) for the log P/D ratio; Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011) for the variance premium; and Fama and French (1989) for the default spread.

CAT 15, it is highest at 0.947. In contrast, the variance premium, a short horizon predictor, has a low persistence of 0.133 during the sample period.

I emphasize that although the CAT measure is constructed from default probabilities of firms in the economy, it is not perfectly correlated with the default spread. Going back to Table 3.3, even looking at CAT 15, which has the highest correlation with the default spread, the coefficient is 0.856, which is not close to 1 nor that much different from the correlation between the CAT measure and the log P/D ratio. This is because the CAT measure is not only capturing investment grade firms' average level of default probabilities like the default spread, but is also capturing probabilities of catastrophic events. As a result, the CAT measure can capture a portion of systematic risk that the default spread cannot separately identify. These aspects can be confirmed by looking at the correlation patterns between different extremities of the CAT measure and the default spread. As extremities increase, the CAT measure becomes less correlated with the default spread.

Lastly, the CAT measure is compared with different tail risk measures that are based on options data. Table 3.5 presents three tail risk measures: at-the-money (ATM) implied volatility, out-of-the-money (OTM) implied volatility with moneyness 0.90, and risk-neutral third moment of log stock returns (Bakshi, Kapadia, and Madan (2003)).

These tail risk measures have relatively high correlations with individual CAT measures. An increase in implied volatilities is positively associated with an increase in tail risk, so it makes intuitive sense that implied volatilities are positively correlated with the CAT measures. In addition, since a more negatively skewed stock return distribution means higher tail risk, the risk-neutral third moment of log stock returns has negative correlations with the CAT measures.

What is important to note here is that the magnitude of correlation depends on the CAT measure's tail extremity; as tail extremity increases, the correlations decrease. For instance, while the risk-neutral third moment has a correlation of -0.598 with CAT 15, with a higher

extremity, say, with CAT 30, the magnitude decreases to -0.446.

This fact signifies that the CAT measure covers tail risks that the options market cannot capture. While options-based measures have been developed to estimate the tail risk in the stock market, they cannot measure tail risks that are extremely deep. In general, when moneyness becomes smaller than, say, 0.80, options become very illiquid, making it difficult to distinguish between small market crashes and catastrophic tail events. On the other hand, since the CAT measure uses joint default probabilities implied by the CDS market to capture catastrophic risk, it can complement options data. That is, when extremity decreases, the CAT measure more closely resembles options-based measures, and vice versa as can be seen in Table 3.5.

3.8.2. Return predictability

3.8.2.1. Predictability of stock returns

I start with return predictability of stock returns based on the CAT measure. Among various maturities and tail extremities of the CAT measure, here I choose 10-year CAT 15, which covers the broadest area of the catastrophic tail distribution. For consistency, I use this measure in all my predictability regressions not only for stock returns but also for government and corporate bond returns.[45]

The following equation shows a typical predictive regression equation for stock returns:

$$\frac{12}{h}\sum_{n=1}^{h}\log(R^e_{t+n}) - \log(R^f_{t+n}) = \text{const.} + \beta X_t + \epsilon_{t+h}, \quad h = 3, 6, 12, 24,$$

where R^e_t represents the 1-month return on the aggregate stock market and R^f_t represents the 1-month Treasury bill rate. The return horizon h is in monthly terms (e.g. $h = 12$ when the horizon is 1-year) and X_t represents stock return predictors used in the analysis.

[45] While 10-year CAT 15 tends to generate the highest R^2 in many cases, other CAT measures also have significant predictive power.

From the first panel of Table 3.6, one can observe that the CAT measure is a long horizon stock return predictor: the predictive power of the CAT measure in terms of R^2 increases in horizon. Looking at the predictive regressions for 1-year and 2-year horizons, the coefficients of the CAT measure are significant for both cases.[46] Furthermore, the CAT measure generates a fairly high R^2. For the 1-year and 2-year horizons, R^2s are 10.56% and 34.90% respectively.

Also, note that the coefficients of the CAT measure are positive. This is an expected result: higher levels of the CAT measure predict higher future excess returns. High levels of catastrophic risk, which is translated into a bad economic state, imply low stock prices. Since the CAT measure is a persistent variable, such low pricing is associated with a high expected return in the long run. Specifically, a 1% point increase in 10-year CAT 15 predicts 0.529% higher excess returns over the next year, and 0.656% higher (annualized) excess returns over the next 2 years.

To compare the CAT measure's performance on return predictability with that of other long horizon return predictors, the second and third panels of Table 3.6 show univariate predictive regression results based on the log P/D and default spread. For the 1-year and 2-year horizon return predictive regressions, coefficients of these classic predictors are all significant and the regressions generate fairly sizable R^2. Importantly, however, the CAT measure explains a larger portion of the variation in expected returns in univariate regressions.

Furthermore, the CAT measure drives out these predictors when they are placed into the same regression. The second and third columns of Table 3.7 present the results of bivariate predictive regressions with the CAT measure and one of the two other predictors (log P/D or default spread). In these two cases, the coefficient of the CAT measure is still significant while the coefficients of other predictors not only become insignificant, but their signs are reversed. Referring back to Table 3.6, in the univariate predictive regression, the coefficient

[46] All standard errors are Newey-West corrected with lags equal to h+3.

for the default spread was positive, and the coefficient for the log P/D was negative.

The last column of Table 3.7 shows the results of the multivariate predictive regression when all three return predictors are entered into the regression. As can be seen in this table, the CAT measure is still highly significant and drives out the two predictors altogether. From this, it is apparent that the CAT measure outperforms other classic long horizon return predictors.

Furthermore, I show the results are robust to finite sample bias and issues with overlapping returns. The details are provided in Appendix A3.6.

3.8.2.2. Predictability of government bond returns

Following the standard notation, let $r_{t+1}^{(n)}$ be the log holding period return from investing in an n-year maturity government bond at time t and selling it as an $n-1$ year maturity bond at time $t+1$. The excess log return on the n-year maturity government bond $rx_{t+1}^{(n)}$ is defined as,

$$rx_{t+1}^{(n)} = r_{t+1}^{(n)} - y_t^{(1)},$$

where $y_t^{(1)}$ represents the log yield on the 1-year zero-coupon government bond at time t.

With this definition of excess log returns on government bonds, I run the following predictive regression:

$$rx_{t+1}^{(n)} = \text{Const.} + \beta X_t + \epsilon_{t+1}^{(n)}, \quad n = 2, 3, 4, 5,$$

where X_t represents bond return predictors used in the analysis.[47] The first panel of Table 3.9 shows that the CAT measure predicts returns on government bonds extremely well. Across all maturities (from 2-year to 5-year), the coefficients of the CAT measure are all highly significant (with t-statistics greater than 4).[48] Furthermore, R^2 is greater than 36% for all maturities.

[47] Since $rx_{t+1}^{(n)}$ is the 1-year excess log holding period return, return horizons considered in the predictive regressions are always 1 year, regardless of bond maturities n.

[48] All standard errors are Newey-West corrected with lags equal to $h+3$ (here $h=12$).

To assess the performance of the CAT measure in predicting government bond returns, I compare the CAT measure with a classic bond return predictor, the Cochrane-Piazzesi (CP) factor. Cochrane and Piazzesi (2005) construct a single factor by calculating a tent-shaped linear combination of forward rates and show that it predicts excess returns on government bonds with high R^2. Since Cochrane and Piazzesi (2005) used a data sample from 1964 to 2003, I first extend the data sample to September 2013 (1964 to 2013; extended sample), re-estimate their CP factor, and run predictive regressions. The results are shown in the second panel of Table 3.9. Even with the extended data sample, the coefficients are still significant, but the CP factor has less predictive power in terms of R^2. Cochrane and Piazzesi (2005) report R^2 that are roughly 31-37%, but in the extended sample, R^2 shrinks to 19-24%.

Additionally, I run predictive regressions using the CP factor constructed from my sample period (October 2003 to September 2013). The results from this regression are reported in the third panel of Table 3.9. In this case, R^2 increases for all maturities, compared to the extended sample CP factor (especially for long maturities). However even these heightened R^2 values are still smaller than those obtained using the CAT measure. Especially for short maturity bonds, the CAT measure generates much higher R^2 (36.79% vs 24.28% for $n = 2$).

I also compare the CAT measure with the term spread, another classic bond market predictor (the fourth panel of Table 3.9). Across all maturities of bonds, the CAT measure performs better than the term spread in terms of R^2.

Table 3.10 presents multivariate predictive regressions of 5-year bond returns ($n = 5$). The second column shows the results when the CAT measure and the CP factor constructed from the extended sample are put into the same regression. As the table shows, the CAT measure is still significant and drives out the CP factor so that it is no longer significant. When the CAT measure and the CP factor constructed from my sample period are entered into the regression (the third column), both the CAT measure and the CP factor are significant and the predictive regression generates a high R^2 (53.98%). Moreover, the CAT measure drives

out the term spread when entered into the same regression (the fourth column). Lastly, as seen in the fifth and sixth columns, the CAT measure is always significant in multivariate regressions with the CAT measure, term spread and one of the CP factors.

The set of results shows that high catastrophic tail risk predicts high future government bond returns. For example, the univariate regression implies that a 1% point increase in 10-year CAT 15 predicts 0.191% higher excess returns on the 5-year maturity government bond over the next year. This suggests that government bond returns are also subject to catastrophic tail risk, which might be interpreted as the possibility of a government default or a significant inflation at the event of catastrophic tail events.

3.8.2.3. Predictability of corporate bond returns

To investigate the predictability of corporate bond returns, I use the Barclays U.S. Corporate Investment Grade index to obtain the time series of corporate bond returns (R_t^{IGB}) and run the following predictive regression:

$$\frac{12}{h}\sum_{n=1}^{h}\log(R_{t+n}^{IGB}) - \log(R_{t+n}^{f}) = \text{const.} + \beta X_t + \epsilon_{t+h}, \quad h = 3, 6, 12, 24.$$

Table 3.11 presents univariate and multivariate predictive regressions. The first panel shows that high CAT risk predicts high corporate bond returns across all horizons.[49] I compare the predictive power of the CAT measure with that of the default spread (a predictor directly from the credit market). In terms of R^2, the CAT measure produces slightly higher R^2, but both measures seem to have similar predictive power as shown in the second panel. However, when both are entered into the same regression, results show that the CAT measure drives out the default spread for long-horizon returns (2-year).

In sum, although the CAT measure is solely constructed from CDS data, it robustly predicts future excess returns in multiple markets, outperforming or driving out classic return

[49] All standard errors are Newey-West corrected with lags equal to h+3.

predictors. Classic predictors for one market are not necessarily strong predictors for others. For example, the log P/D is not a good predictor for bond returns, and the CP factor is not a good predictor for stock returns. The joint predictability of the CAT measure across different asset classes suggests that it captures important systematic risk shared by various markets.

3.8.3. Cross section of stock returns

To examine whether catastrophic tail risk is indeed an important risk factor for asset pricing, I turn to the cross section of stock returns. The main focus here is to test if the sensitivity to the CAT measure generates significant return dispersion in the cross section of stock returns. Similar approaches have been taken by Pastor and Stambaugh (2003) for aggregate liquidity risk and Ang, Hodrick, Xing, and Zhang (2006) for aggregate volatility risk.

First, I download the individual stock return time series from the CRSP database. Then, for each month and each stock, I calculate the *pre-ranking CAT beta*. Specifically, I run the following rolling regressions with 10-year CAT 25 and the Fama and French (1993) three factors based on past 24 months data:[50]

$$\log(R_t^{ie}) - \log(R_t^f) = \text{const.} + \beta^{\text{CAT}} \Delta\text{CAT}_t + \beta^{MKT}\text{MKT}_t + \beta^{SMB}\text{SMB}_t + \beta^{HML}\text{HML}_t + \epsilon_t.$$

The CAT beta refers the loading (or coefficient) on ΔCAT_t, which is a proxy for the innovation to the CAT measure. I emphasize that I control for the Fama-French three factors to make sure that the CAT beta picks up the sensitivity of stock returns to catastrophic tail risk that is not captured by the exposures to MKT, SMB, and HML factors. Once the CAT betas are obtained, I sort individual stocks into quintile portfolios at the end of each month based on those betas, and investigate value-weighted returns in the next month.

[50]Note that in this cross sectional analysis, I use CAT 25. If I use another CAT measure with different extremities (say, CAT 15), I get less significant results. This potentially suggests that the tail component with extremities greater than 25% (not between 15 and 25%) is the one that produces meaningful cross sectional dispersion in the stock market.

Table 3.12 shows the statistics of the resulting five portfolios. The pre-ranking CAT betas suggest that the first quintile portfolio is the most sensitive to catastrophic tail risk ($\beta^{\text{CAT}} = -1.47$). That is, this portfolio is the most exposed to catastrophic tail risk. As a result, this quintile portfolio has the highest mean return (1.11% monthly, or equivalently 13.32% annually). Furthermore, the mean return, CAPM alpha, and Fama-French three-factor alpha of the zero-cost portfolio that goes long the first quintile portfolio and short the fifth quintile portfolio are significantly large, as shown in the last two rows.[51] Specifically, the mean return of this long-short portfolio is 0.45% per month, which corresponds to an annual return of 5.4%. In other words, the portfolios that are the most sensitive and least sensitive to catastrophic tail risk show a significant return dispersion, which implies that catastrophic tail risk is negatively priced in the cross section of stock returns.

It should be noted that the pattern of the mean returns across the five quintile portfolios is not perfectly monotonic.[52] One of the possible explanations for this pattern is that betas are very noisy measures of sensitivity or that betas and returns are time-varying and quickly mean-reverting. To check for this possibility, I calculate the full-sample *post-ranking CAT betas* using the same regression equation above but with the time series of the constructed portfolio returns in the full sample. This shows that the pattern of post-ranking CAT betas is similar to that of mean returns.[53]

In conclusion, catastrophic tail risk can explain the cross section of stock returns controlling for exposures to the Fama-French three factors. The results obtained are robust to adding other control factors such as the momentum factor of Carhart (1997) or the liquidity factor of Pastor and Stambaugh (2003).

[51] All standard errors are Newey-West corrected with lags equal to 12.

[52] This is not driven by the concentration of small stocks or value stocks in the first quintile portfolio.

[53] It should be emphasized that except for the first portfolio, the difference between any two portfolios is not statistically significant. Only the difference between the first portfolio and any other portfolio is highly significant.

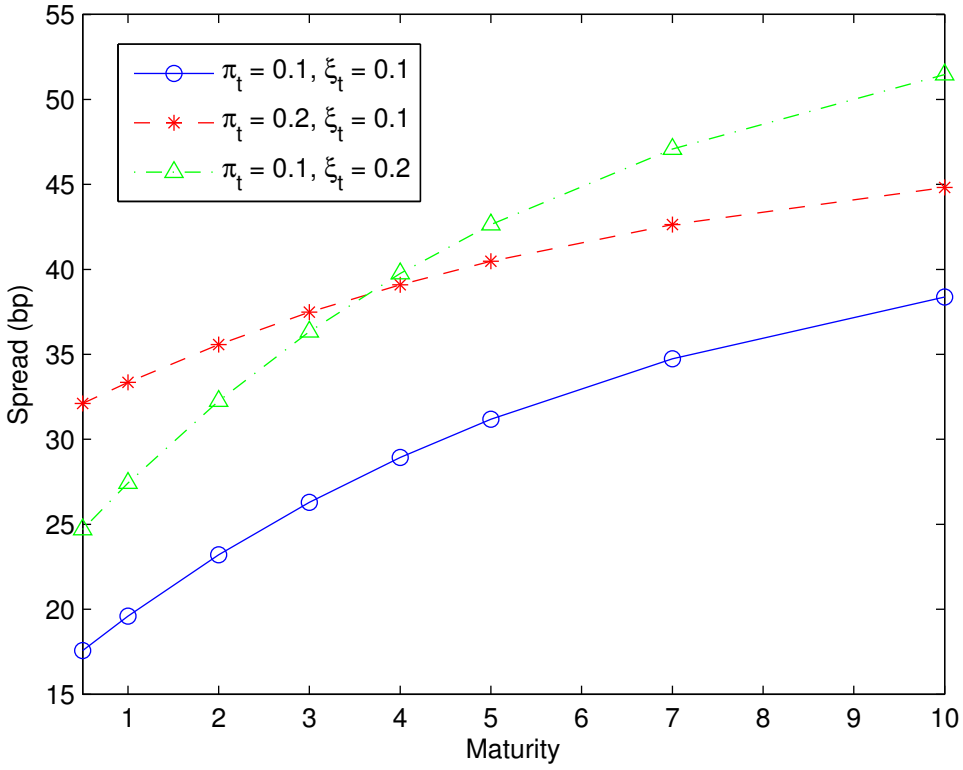
3.9. Conclusion

Since the financial crisis, more attention has been devoted to understanding the significance of tail risk in asset markets. I study tail risk from a completely different angle by looking at the risk of massive correlated defaults in the economy. By its definition, this risk exclusively reflects extreme events that cannot be separated out using traditional methods. To capture this risk, I build a model that incorporates the co-movements between credit risks of multiple firms that arise from regime and belief shifts. Based on the estimated model, I develop a measure of catastrophic tail risk implied by the CDS data.

Do investors incorporate these extremely unlikely tail events? My results say yes. The empirical results consistently indicate that they are an important source of systematic risk for asset pricing. The CAT measure jointly predicts stock and bond returns, which suggests that catastrophic tail risk is pervasive across various markets. Investors account for this risk, evidenced by the significant return dispersion between stocks with high and low exposures on catastrophic tail risk.

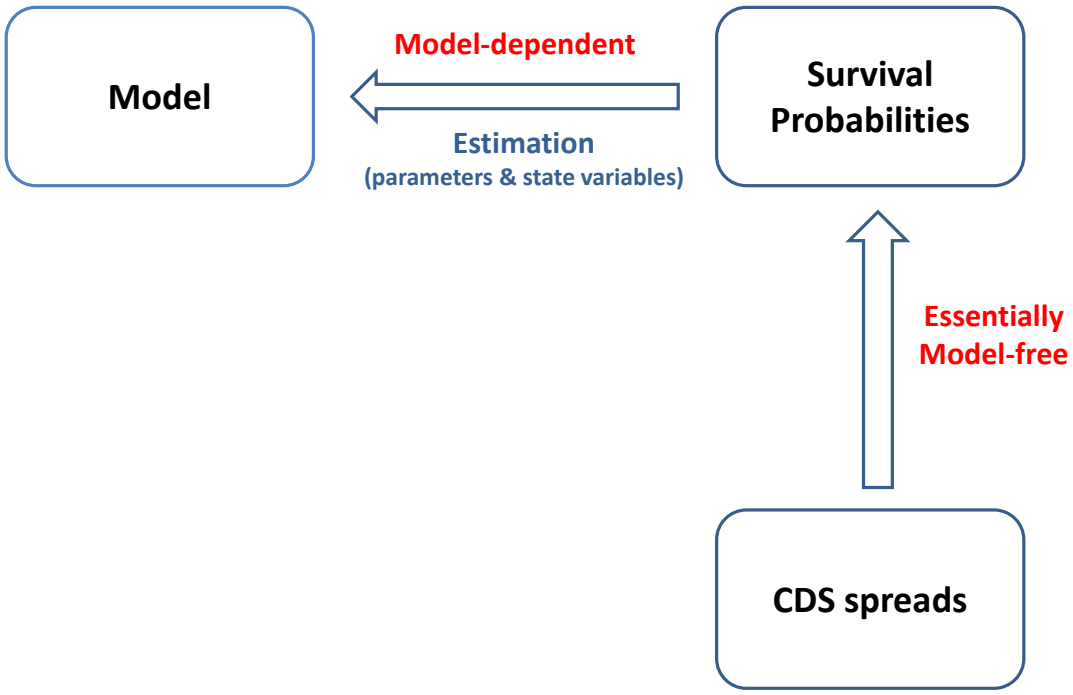
This paper shows that the term structure and cross section of the CDS market contain valuable information about the macroeconomy and investors' beliefs. This information is filtered into the CAT measure. This measure can be exploited to deepen understanding about the foundations of economic catastrophes and their impact on the aggregate economy.

Figure A.1: Impact of changes in beliefs on CDS term structure



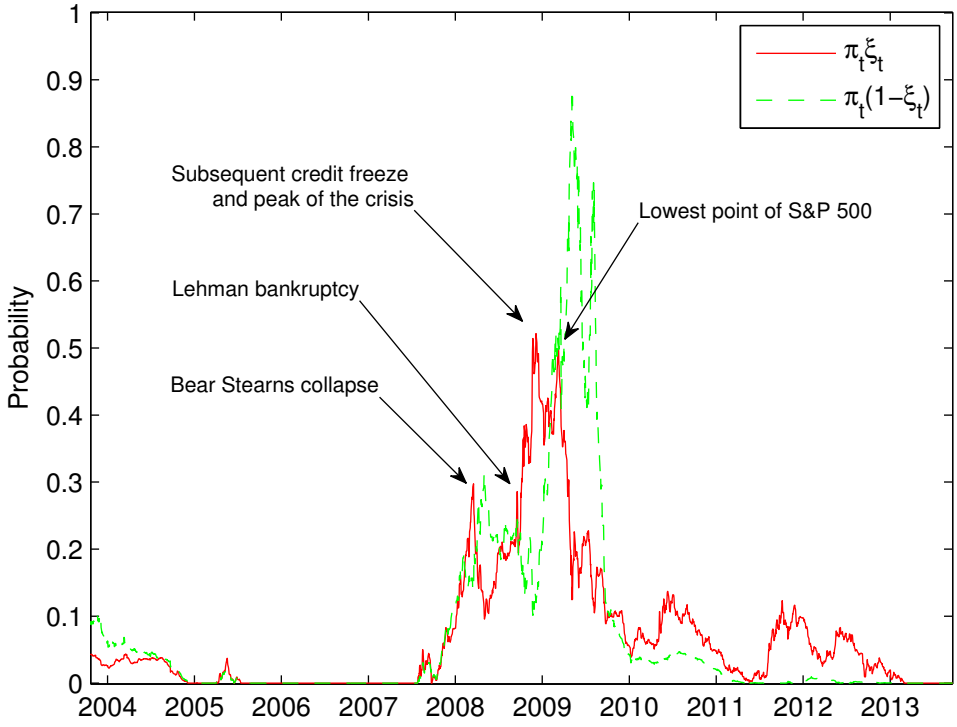
Note: This figure shows the comparative statistics of the model when the two beliefs change. The blue line with circle markers represents the model-implied CDS term structure when the two beliefs (π_t and ξ_t) are all 10%. The red line with asterisk markers fixes ξ_t but increases π_t to 20%. The green line with triangle markers fixes π_t but increases ξ_t to 20%. π_t represents investors' beliefs that the economy is currently in the frailty state at time t, and ξ_t represents investors' beliefs that the true severity of the frailty state is high.

Figure A.2: Estimation of the model



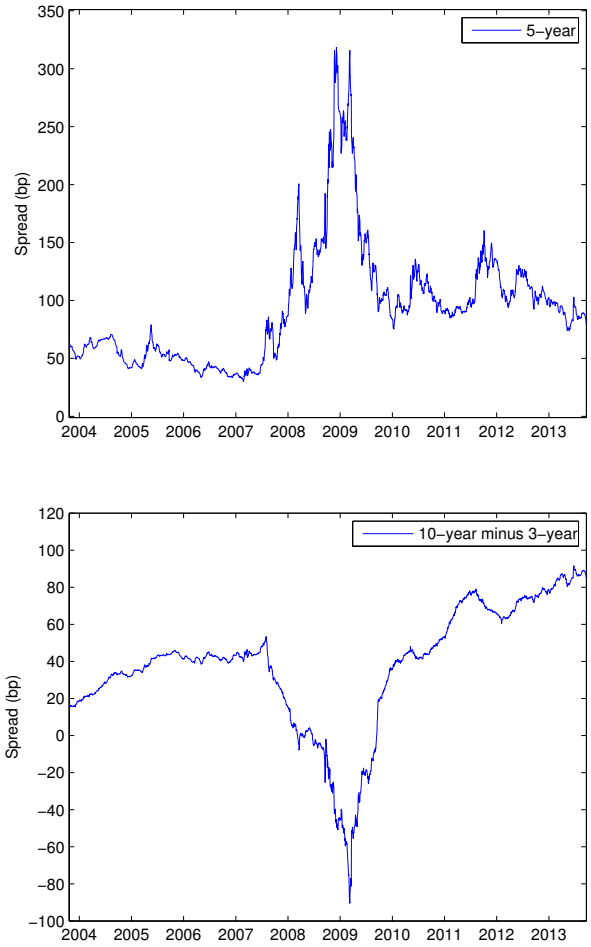
Note: The diagram describes the relationship between the model, survival probabilities, and CDS spreads. Pricing CDS is model-free once the term structure of survival probabilities is obtained (also the reverse direction is also essentially model-free). However, the term structure of survival probabilities critically depends on the underlying model and its state variables.

Figure A.3: Implied beliefs of investors



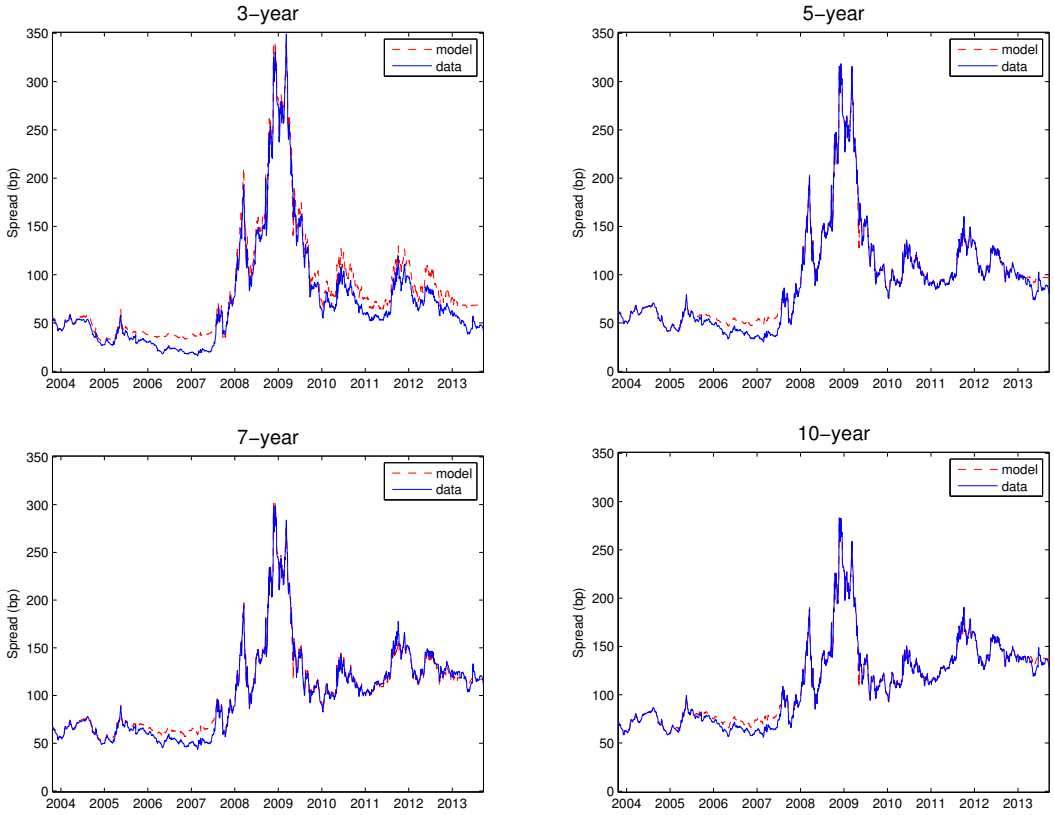
Note: This figure displays the extracted time series of investors' two beliefs. The red solid line represents the probability that the economy is in the severe frailty state ($\pi_t\xi_t$). The green dashed line represents the probability that the economy is in the moderate frailty state ($\pi_t(1-\xi_t)$).

Figure A.4: Time series of the average firm's CDS curves: level and slope



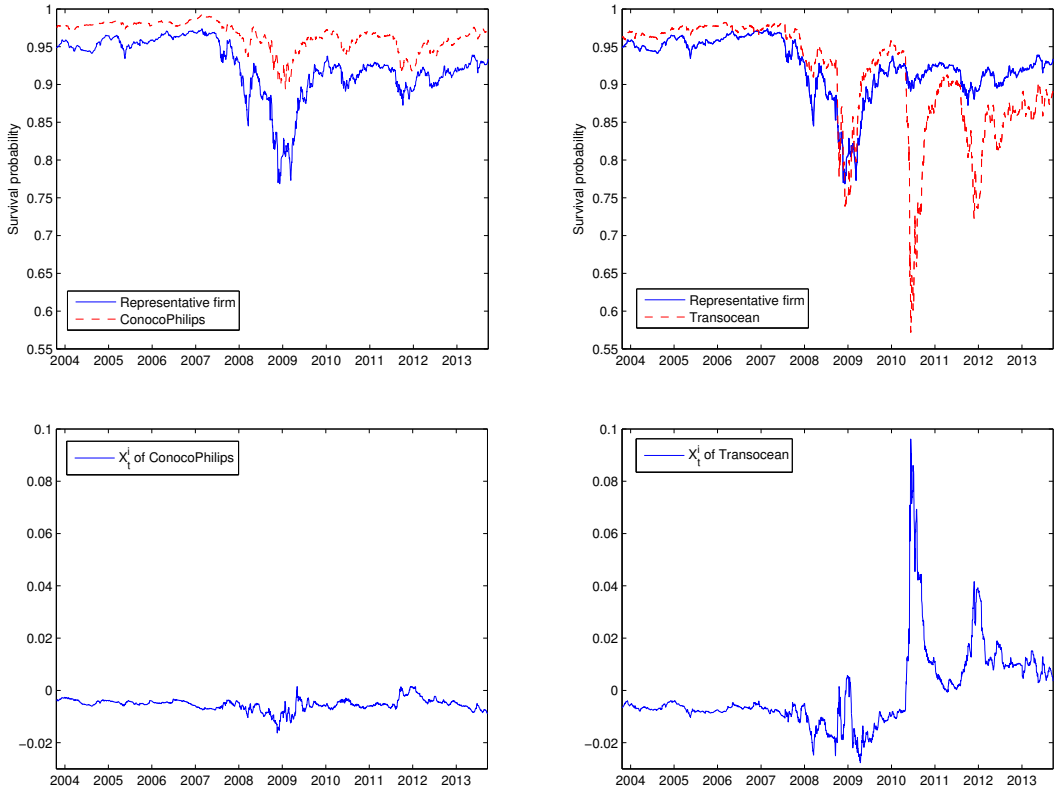
Note: Two panels present the implied CDS curves of the hypothetical firm representing the average firm in the portfolio. The top panel shows the time series of 5-year CDS spreads, which depict the level of the CDS curves. The bottom panel shows the time series of the difference between 10-year spreads and 3-year spreads, which depict the slope of the CDS curves.

Figure A.5: CDS term structure of the average firm in the model and data



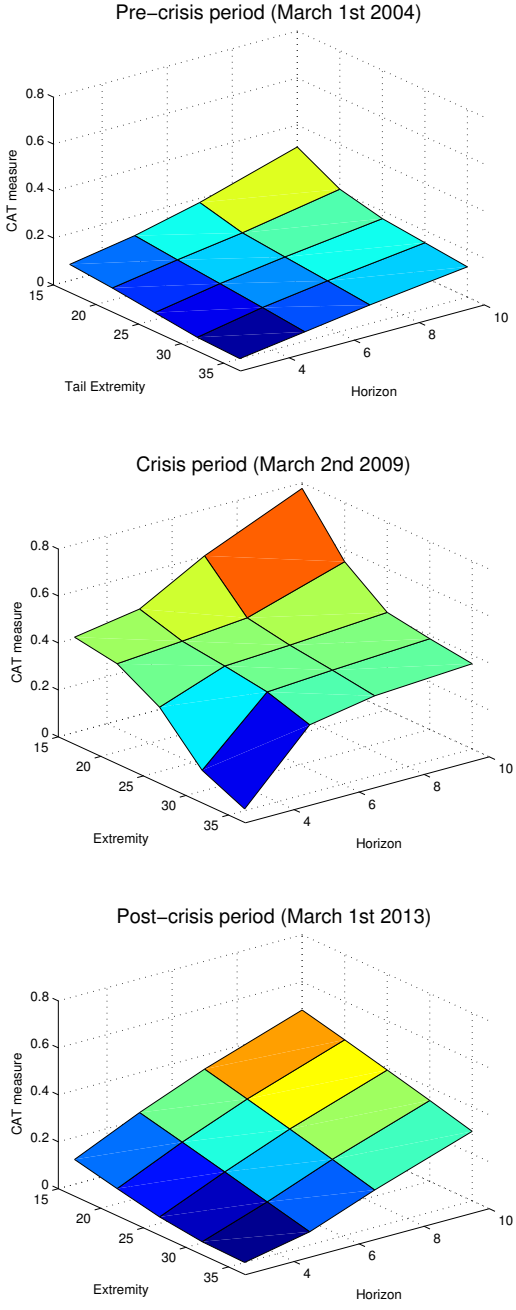
Note: The figure shows the CDS term structure of the average firm in the model and data. The red dashed lines represent the time series of the model implied CDS term structure (3-year, 5-year, 7-year, 10-year maturities). The blue lines represent their data counterparts.

Figure A.6: Individual firm dynamics: two examples



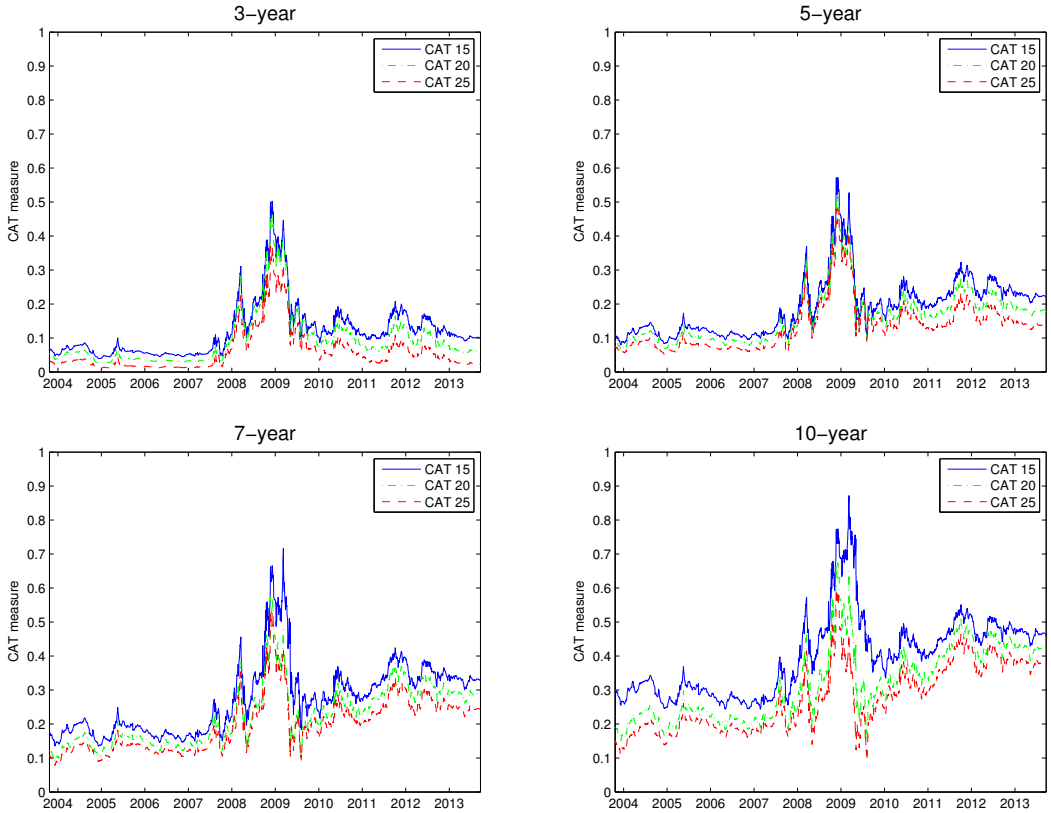
Note: This figure provides two examples of individual firm dynamics. The two left panels are about ConocoPhillips and the two right panels are about Transocean. The top two panels display each firm's time series of 5-year implied survival probabilities (red dashed line) together with the average firm's corresponding time series (blue solid line). The bottom two panels present the realizations of each firm's idiosyncratic component (X_t^i) of default intensity.

Figure A.7: CAT surfaces in the pre-crisis, crisis, and post-crisis period



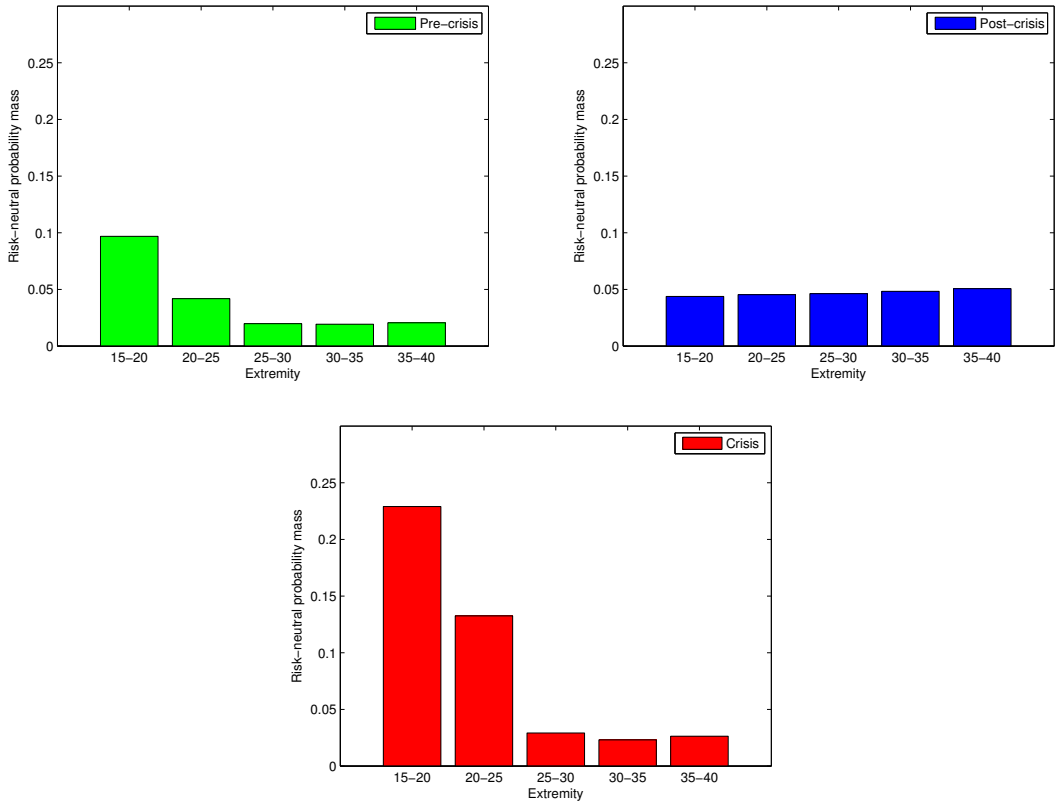
Note: The figure shows three CAT surfaces representing the three periods: the pre-crisis, crisis, and post-crisis periods. Each CAT surface has two dimensions - tail extremity and horizon.

Figure A.8: Time series of CAT surfaces



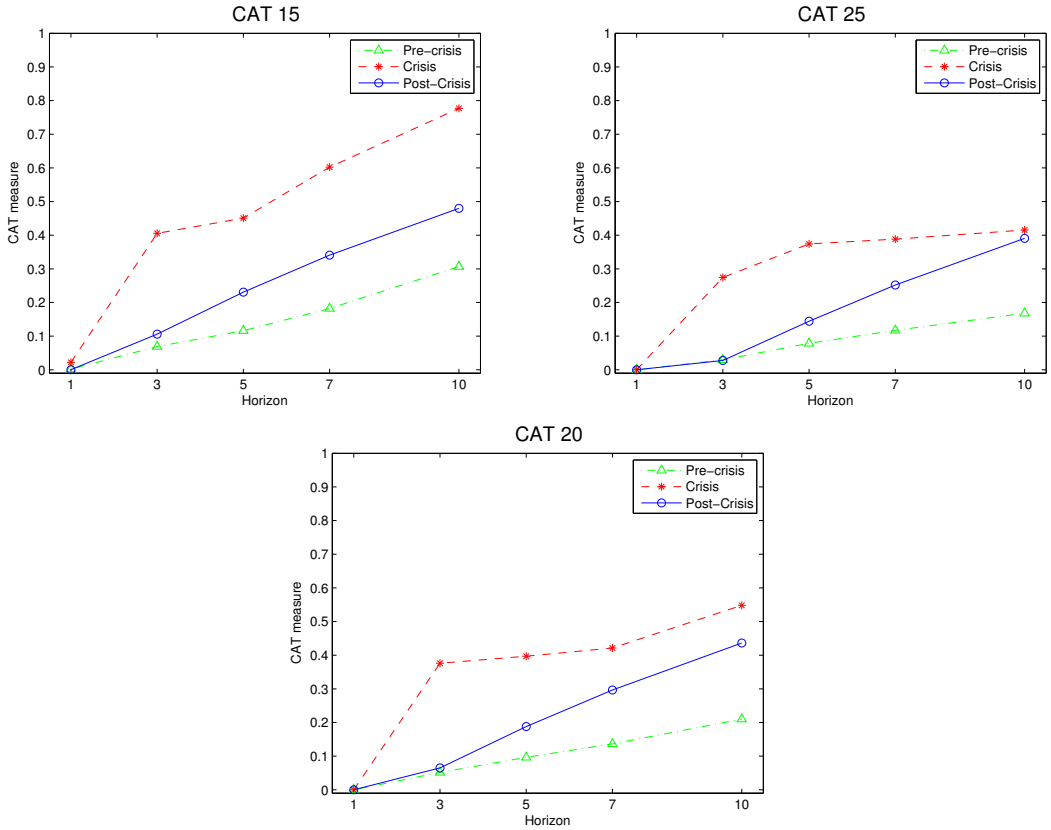
Note: The figure shows the time series of CAT surfaces. Tail extremities 15%, 20%, and 25% are considered at 3-year, 5-year, 7-year, and 10-year horizons. The blue solid lines represent CAT 15, the green dot-dashed lines represent CAT 20, and the red dashed lines represent CAT 25.

Figure A.9: 10-year catastrophic tail distribution in the pre-crisis, crisis, and post-crisis period



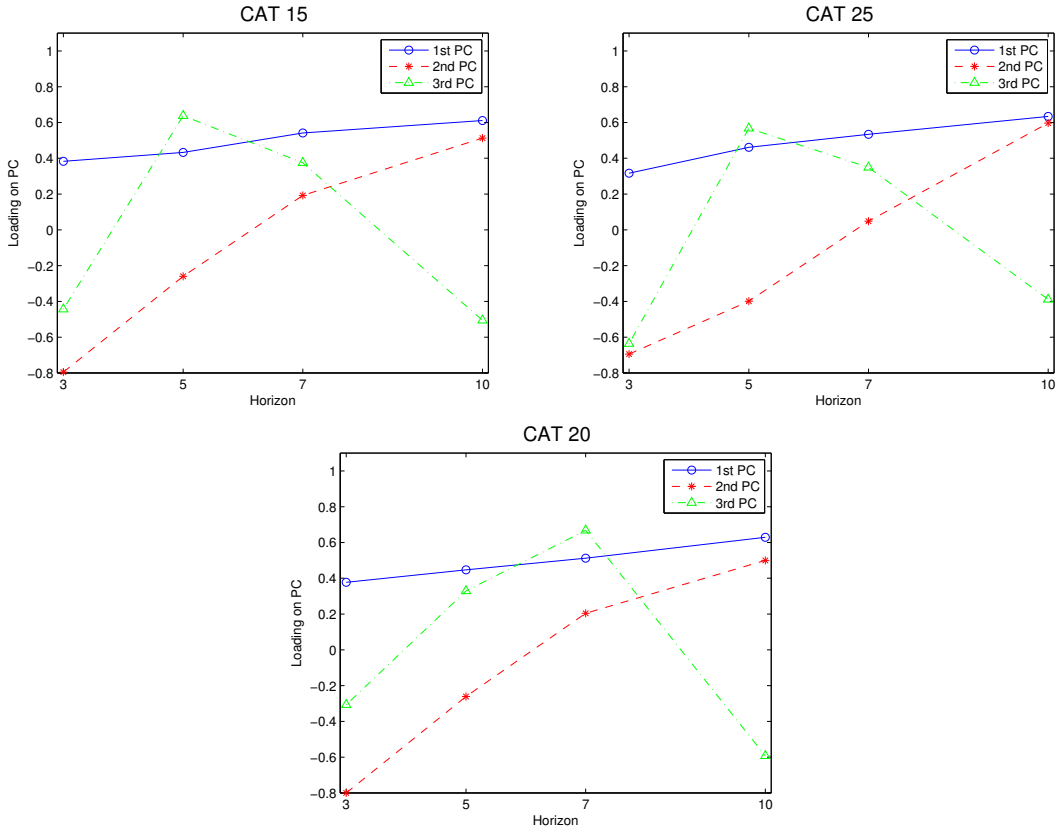
Note: The figure presents three histograms, each of which represents the risk-neutral distribution of catastrophic tail events during one of the three periods: the pre-crisis, crisis, and post-crisis periods. All distributions are under the 10-year horizon. Each histogram has 5 bars representing the probability masses between 15-20%, 20-25%, 25-30%, 30-35%, and 35-40%. The green histogram (top left panel) represents the pre-crisis period, the red histogram (bottom panel) represents the crisis period, and the blue histogram (top right panel) represents the post-crisis period.

Figure A.10: Term structure of the CAT measure in the pre-crisis, crisis, and post-crisis period



Note: The figure presents the term structure of the CAT measure with three different extremities: CAT 15 (top left panel), CAT 20 (bottom panel), and CAT 25 (top right panel). In each panel, I put the three term structures that represent the pre-crisis (green dot-dashed line with triangle markers), crisis (red dashed line with asterisk markers), and post-crisis (blue solid line with circle markers) periods.

Figure A.11: Principal component analysis



Note: This figure shows the results of the principal component analysis based on the term structures of CAT 15 (top left panel), CAT 20 (bottom panel), and CAT 25 (top right panel). The blue solid line with circle markers, the red dashed line with asterisk markers, and the green dot-dashed line with triangle markers each represent the loadings on the first, second, and third principal components respectively.

Table 3.1: Correlation matrix of CAT measures with different extremities

	CAT 15	CAT 20	CAT 25	CAT 30
CAT 15	—			
CAT 20	0.907	—		
CAT 25	0.800	0.975	—	
CAT 30	0.815	0.977	0.999	—

Notes: The table calculates the correlations between 10-year horizon CAT measures with different extremities. Specifically, CAT 15, CAT 20, CAT 25, and CAT 30 are considered.

Table 3.2: Correlation matrix of CAT measures with same extremities but different horizons

		CAT 15			
		3-year	5-year	7-year	10-year
CAT 15	3-year	—			
	5-year	0.937	—		
	7-year	0.902	0.982	—	
	10-year	0.883	0.943	0.984	—

		CAT 25			
		3-year	5-year	7-year	10-year
CAT 25	3-year	—			
	5-year	0.927	—		
	7-year	0.737	0.936	—	
	10-year	0.491	0.780	0.950	—

Notes: The table presents the correlations between different horizons of CAT measures. The correlations between 3-year, 5-year, 7-year, and 10-year horizon CAT measures are computed for CAT 15 and CAT 25.

Table 3.3: Correlation matrix of CAT measures and other classic stock return predictors

	CAT 15	CAT 20	CAT 25	CAT 30
log(P/D)	-0.885	-0.694	-0.549	-0.574
Default spread	0.856	0.657	0.498	0.524
Variance premium	-0.019	-0.067	-0.095	-0.103

Notes: The table presents the correlations between 10-year horizon CAT measures (CAT 15, CAT 20, CAT 25, and CAT 30) and other classic stock return predictors, including the log P/D, default spread, and variance premium.

Table 3.4: Persistence of CAT measures and other classic stock return predictors

log(P/D)	0.971
Default spread	0.958
Variance premium	0.133
CAT 15	0.947
CAT 20	0.925
CAT 25	0.915
CAT 30	0.908

Notes: The table shows the persistence of 10-year horizon CAT measures together with the persistence of other classic stock return predictors.

Table 3.5: Correlation matrix of individual CAT measures and tail risk measures

	CAT 15	CAT 20	CAT 25	CAT 30
ATM implied volatility	0.789	0.660	0.565	0.594
OTM implied volatility	0.785	0.620	0.508	0.541
RN 3rd moment of log returns	-0.598	-0.493	-0.419	-0.446

Notes: The table presents the correlations between 10-year horizon CAT measures and three tail risk measures: at-the-money (ATM) implied volatility, out-of-the-money (OTM) implied volatility with moneyness 0.90, and risk-neutral third moment of log stock returns.

Table 3.6: Univariate predictive regressions of stock returns

	Horizon h (in months)			
	3	6	12	24
CAT	0.246	0.498	0.529	0.656
	(0.46)	(1.62)	(2.72)	(3.50)
R^2	0.66%	4.41%	10.56%	34.90%
log(P/D)	-0.185	-0.354	-0.338	-0.343
	(-0.42)	(-1.39)	(-2.95)	(-2.87)
R^2	0.78%	4.62%	9.08%	21.05%
Default spread	1.626	8.786	10.332	11.306
	(0.11)	(1.01)	(3.11)	(3.78)
R^2	0.05%	2.45%	7.22%	19.51%

Notes: The table reports the results from the following predictive regression:

$$\frac{12}{h}\sum_{n=1}^{h}\log(R^e_{t+n}) - \log(R^f_{t+n}) = \text{const.} + \beta X_t + \epsilon_{t+h}, \quad h = 3, 6, 12, 24,$$

where R^e_t represents the 1-month return on the aggregate stock market and R^f_t represents the 1-month Treasury bill rate. The return horizon h is in monthly terms and X_t represents stock return predictors used in the analysis. Since predictive regressions here are univariate, X_t represents either one of the CAT measure (first panel), log P/D (second panel), or default spread (third panel). All standard errors are Newey-West corrected with lags equal to h+3.

Table 3.7: Multivariate predictive regressions of 2-year horizon stock returns

CAT	0.936	0.879	0.945
	(3.15)	(3.04)	(3.22)
log(P/D)	0.214		0.157
	(1.11)		(0.48)
Default spread		-6.036	-2.397
		(-1.42)	(-0.33)
Adjusted R^2	35.43%	35.12%	34.21%

Notes: The table reports the results from the following predictive regression:

$$\frac{12}{h}\sum_{n=1}^{h}\log(R^e_{t+n}) - \log(R^f_{t+n}) = \text{const.} + \beta X_t + \epsilon_{t+h}, \quad h = 24,$$

where R^e_t represents the 1-month return on the aggregate stock market, R^f_t represents the 1-month Treasury bill rate. Since predictive regressions here are multivariate, X_t is a vector of predictors that includes the CAT measure and at least one of the two classic predictors (log P/D or default spread). All standard errors are Newey-West corrected with lags equal to $h + 3$ (here $h = 24$).

Table 3.8: VAR-implied R^2

	Horizon h (in months)			
	3	6	12	24
Data	0.66%	4.41%	10.56%	34.90%
5%	0.02%	0.14%	0.67%	1.39%
Median	1.67%	6.58%	16.26%	27.32%
95%	10.46%	24.54%	43.98%	60.01%
Pop (Hodrick)	1.18%	4.77%	11.48%	18.59%

Notes: The table presents VAR-implied R^2 values of univariate stock return predictive regressions based on the CAT measure. The first row reports the R^2 values in the data and the next three rows report the median values and 90% confidence intervals obtained by simulating the estimated VAR(1) system. The last row reports the population R^2 computed using the formula in Hodrick (1992).

Table 3.9: Univariate predictive regressions of government bond returns

	n (maturity of bond at time t)			
	2	3	4	5
CAT	0.052	0.099	0.146	0.191
	(4.33)	(4.41)	(4.81)	(5.03)
R^2	36.79%	36.94%	39.03%	40.92%
CP factor (1964-2013)	0.436	0.835	1.252	1.477
	(5.51)	(5.29)	(5.63)	(5.31)
R^2	19.57%	21.34%	24.72%	22.57%
CP factor (2003-2013)	0.372	0.759	1.214	1.643
	(2.83)	(3.25)	(4.09)	(4.46)
R^2	24.28%	28.14%	34.66%	39.24%
Term spread	0.514	0.892	1.138	1.343
	(4.06)	(3.42)	(2.90)	(2.61)
R^2	34.49%	28.98%	22.73%	19.54%

Notes: The table reports the results from the following predictive regression:

$$rx_{t+1}^{(n)} = \text{Const.} + \beta X_t + \epsilon_{t+1}^{(n)}, \quad n = 2, 3, 4, 5,$$

where $rx_{t+1}^{(n)}$ is the excess log holding period return from investing in an n year maturity government bond at time t and selling it as an $n-1$ year maturity bond at time $t+1$, and X_t represents bond return predictors used in the analysis. Since predictive regressions are univariate, X_t presents either one of the CAT measure (first panel), extended sample CP factor (second panel), my sample CP factor (third panel), or term spread (fourth panel). All standard errors are Newey-West corrected with lags equal to $h+3$ (here $h = 12$).

Table 3.10: Multivariate predictive regressions of 5-year maturity government bond returns

CAT	0.205	0.133	0.165	0.190	0.122
	(4.87)	(3.80)	(5.64)	(4.07)	(2.95)
CP factor (1964-2013)	0.650			0.545	
	(1.45)			(1.02)	
CP factor (2003-2013)		1.102			1.039
		(3.12)			(2.83)
Term spread			0.588	0.275	0.317
			(1.83)	(0.61)	(0.93)
Adjusted R^2	45.47%	53.98%	42.94%	45.53%	54.43%

Notes: The table reports the results from the following predictive regression:

$$rx_{t+1}^{(n)} = \text{Const.} + \beta X_t + \epsilon_{t+1}^{(n)}, \quad n = 5,$$

where $rx_{t+1}^{(n)}$ is the excess log holding period return from investing in an n year maturity government bond at time t and selling it as an $n-1$ year maturity bond at time $t+1$. Since predictive regressions here are multivariate, X_t is the vector of predictors that includes the CAT measure and at least one other predictors (one of the two CP factors or term spread). All standard errors are Newey-West corrected with lags equal to $h+3$ (here $h = 12$).

Table 3.11: Predictive regressions of corporate bond returns

	Horizon h (in months)			
	3	6	12	24
CAT	0.423	0.453	0.422	0.350
	(3.02)	(3.59)	(5.40)	(8.46)
R^2	13.05%	29.34%	46.75%	66.17%
Default spread	9.880	10.491	9.573	7.190
	(2.71)	(4.38)	(10.45)	(6.01)
R^2	12.77%	28.20%	43.10%	51.02%
CAT	0.245	0.271	0.272	0.322
	(1.42)	(1.88)	(2.51)	(3.56)
Default spread	5.232	5.350	4.420	0.764
	(0.94)	(1.29)	(2.03)	(0.49)
Adjusted R^2	12.87%	30.79%	49.17%	65.68%

Note: The table reports the results from the following predictive regression:

$$\frac{12}{h}\sum_{n=1}^{h}\log(R_{t+n}^{IGB}) - \log(R_{t+n}^{f}) = \text{const.} + \beta X_t + \epsilon_{t+h},$$

where R_t^{IGB} is the 1-month return on corporate bonds calculated using the Barclays U.S. Corporate Investment Grade index, and R_t^f represents the 1-month Treasury bill rate. The return horizon h is in monthly terms and X_t represents corporate bond return predictors used in the analysis. The first two panels show the results from the univariate predictive regressions where X_t is either the CAT measure (first panel) or default spread (second panel). The third panel shows the results from the multivariate predictive regressions where X_t is a vector of predictors that include both the CAT measure and default spread. All standard errors are Newey-West corrected with lags equal to h+3.

Table 3.12: The CAT measure and cross section of stock returns

Quintile Portfolios	Mean return	Std return	CAPM alpha	FF-3 alpha	Pre-ranking CAT beta	Post-ranking CAT beta
1	1.11	6.33	0.45	0.47	-1.47	-0.16
2	0.40	4.77	-0.12	-0.12	-0.56	-0.01
3	0.40	4.54	-0.10	-0.10	-0.11	-0.04
4	0.46	4.50	-0.03	-0.03	0.32	-0.08
5	0.66	5.81	0.06	0.02	1.07	-0.12
1 minus 5	0.45		0.40	0.45		
	(2.26)		(2.25)	(2.36)		

Notes: The table reports the statistics of the five portfolios constructed by sorting individual stocks based on the pre-ranking CAT betas (sixth column), which are calculated using the following rolling regressions with past 24 months data:

$$\log(R_t^{ie}) - \log(R_t^f) = \text{const.} + \beta^{\text{CAT}}\Delta\text{CAT}_t + \beta^{MKT}\text{MKT}_t + \beta^{SMB}\text{SMB}_t + \beta^{HML}\text{HML}_t + \epsilon_t.$$

All portfolio returns are monthly and value-weighted. The means and standard deviations of the five portfolios are shown in the second and third columns. Also, the CAPM alphas and Fama-French three-factor alphas are reported in the fourth and fifth columns. The post-ranking CAT betas (last column) are calculated using the same regression equation above but with the time series of the constructed portfolio returns in the full sample. The last two rows report the mean return, CAPM alpha, and Fama-French three-factor alpha of the zero-cost portfolio that goes long the first quintile portfolio and short the fifth quintile portfolio together with their t-statistics. All standard errors are Newey-West corrected with lags equal to 12.

APPENDIX

A1. Appendix for Chapter 1

A1.1. Data construction

Our sample consists of daily data on option prices, volume and open interest for European put options on the S&P 500 index from OptionMetrics. Data are from 1996 to 2012. Options expire on the Saturday that follows the third Friday of the month. We extract monthly observations using data from the Wednesday of every option expiration week. We apply standard filters to ensure that the contracts on which we base our analyses trade sufficiently often for prices to be meaningful. That is, we exclude observations with bid price smaller than 1/8 and those with zero volume and open interest smaller than one hundred contracts (Shaliastovich (2009)).

OptionMetrics constructs implied volatilities using the formula of Black and Scholes (1973) (generalized for an underlying that pays dividends), with LIBOR as the short-term interest rate. The dividend-yield is extracted from the put-call parity relation. We wish to construct a data set of implied volatilities with maturities of 1, 3 and 6 months across a range of strike prices. Of course, there will not be liquid options with maturity precisely equal to, say, 3 months, at each date. For this reason, we use polynomial interpolation across strike prices and times to expiration.[1] Specifically, at each date in the sample, we regress implied volatilities on a polynomial in strike price K and maturity T:

$$\sigma(K,T) = \theta_0 + \theta_1 K + \theta_2 K^2 + \theta_3 T + \theta_4 T^2 + \theta_5 KT + \theta_6 KT^2 + \epsilon_{K,T}$$

We run this regression on options with maturities ranging from 30 to 247 days, and with moneyness below 1.1. The implied volatility surface is generated by the fitted values of this regression.

[1] See Dumas, Fleming, and Whaley (1998), Christoffersen and Jacobs (2004) and Christoffersen, Heston, and Jacobs (2009).

A1.2. Details of calculations for the single-factor model

A1.2.1. The state-price density

Duffie and Skiadas (1994) show that the state-price density π_t equals

$$\pi_t = \exp\left\{\int_0^t \frac{\partial}{\partial V} f\left(C_s, V_s\right) ds\right\} \frac{\partial}{\partial C} f\left(C_t, V_t\right). \tag{A.1}$$

Equation (A.1) shows the state-price density can be expressed in terms of a locally deterministic term and a term that is locally stochastic. To obtain (1.12), we require both to be expressed in terms of C_t and λ_t. We derive the expression for the stochastic term first.

It follows from (1.4) that

$$\frac{\partial}{\partial C} f\left(C_t, V_t\right) = \beta\left(1-\gamma\right)\frac{V_t}{C_t}. \tag{A.2}$$

Wachter (2013) shows that continuation utility V_t can be expressed in terms of C_t as follows:[2]

$$V_t = J(C_t, \lambda_t), \tag{A.3}$$

where

$$J(C_t, \lambda_t) = \frac{C_t^{1-\gamma}}{1-\gamma} e^{a+b\lambda_t}, \tag{A.4}$$

and

$$a = \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + b\frac{\kappa\bar{\lambda}}{\beta} \tag{A.5}$$

$$b = \frac{\kappa+\beta}{\sigma_\lambda^2} - \sqrt{\left(\frac{\kappa+\beta}{\sigma_\lambda^2}\right)^2 - 2\frac{E_\nu\left[e^{(1-\gamma)Z}-1\right]}{\sigma_\lambda^2}}. \tag{A.6}$$

[2]Wachter (2013) expresses the value function in terms of wealth rather than consumption. Because the ratio of wealth to consumption is β^{-1}, it is straightforward to go from one expression to the other. The difference between the expressions is found in the definition of a. In the earlier paper, this expression has an extra term given by $(1-\gamma)\log\beta$.

For future reference, we note that b is a solution to the quadratic equation

$$\frac{1}{2}\sigma_\lambda^2 b^2 - (\kappa + \beta)b + E_\nu \left[e^{(1-\gamma)Z} - 1\right] = 0. \tag{A.7}$$

Substituting (A.3) and (A.4) into (A.2) implies that

$$\frac{\partial}{\partial C} f\left(C_t, V_t\right) = \beta C_t^{-\gamma} e^{a+b\lambda_t}. \tag{A.8}$$

It also follows from (1.4) that

$$\frac{\partial}{\partial V} f(C_t, V_t) = \beta(1-\gamma)\left(\log C_t - \frac{1}{1-\gamma}\log\left((1-\gamma)V_t\right)\right) + \beta.$$

Substituting in for V_t from (A.3) and (A.4) implies

$$\frac{\partial}{\partial V} f(C_t, V_t) = -\beta(a + b\lambda_t) - \beta. \tag{A.9}$$

Finally, we collect constant terms:

$$\eta = -\beta a - \beta \tag{A.10}$$

so that

$$\frac{\partial}{\partial V} f(C_t, V_t) = \eta - \beta b \lambda_t.$$

Therefore, from (A.1) it follows that the state-price density can be written as

$$\pi_t = \exp\left(\eta t - \beta b \int_0^t \lambda_s ds\right) \beta C_t^{-\gamma} e^{a+b\lambda_t}.$$

A1.2.2. The iid limit

In this section we compute the limit of the state price density as σ_λ approaches zero. Note that b in equation (A.6) can be rewritten as

$$b = \frac{1}{\sigma_\lambda^2}\left(\kappa + \beta - \sqrt{(\kappa+\beta)^2 - 2E_\nu\left[e^{(1-\gamma)Z} - 1\right]\sigma_\lambda^2}\right).$$

L'Hopital's rule implies

$$\begin{aligned}\lim_{\sigma_\lambda \to 0} b &= \lim_{\sigma_\lambda \to 0} \frac{1}{2}\left((\kappa+\beta)^2 - 2E_\nu\left[e^{(1-\gamma)Z} - 1\right]\sigma_\lambda^2\right)^{-\frac{1}{2}} 2E_\nu\left[e^{(1-\gamma)Z} - 1\right] \\ &= \frac{E_\nu\left[e^{(1-\gamma)Z} - 1\right]}{\kappa + \beta}.\end{aligned}$$

It follows from the equation for a, (A.5), that

$$\begin{aligned}\lim_{\sigma_\lambda \to 0}(a + b\lambda_t) &= \lim_{\sigma_\lambda \to 0}(a + b\bar{\lambda}) \\ &= \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + (\kappa+\beta)\frac{\bar{\lambda}}{\beta}\lim_{\sigma_\lambda \to 0} b \\ &= \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + \frac{E_\nu\left[e^{(1-\gamma)Z} - 1\right]\bar{\lambda}}{\beta},\end{aligned}$$

where we assume that $\lambda_0 = \bar{\lambda}$ and therefore that $\lambda_t = \bar{\lambda}$ for all t.

We now apply these results to calculate the limit of π_t/π_0. It follows from (A.1), (A.8) and (A.9) that

$$\begin{aligned}\lim_{\sigma_\lambda \to 0}\frac{\pi_t}{\pi_0} &= \exp\left\{\left(-\beta - \beta\lim_{\sigma_\lambda \to 0}(a + b\bar{\lambda})\right)t\right\}\left(\frac{C_t}{C_0}\right)^{-\gamma} \\ &= \exp\left\{\left(-\beta - (1-\gamma)(\mu - \frac{1}{2}\gamma\sigma^2) - E_\nu\left[e^{(1-\gamma)Z} - 1\right]\bar{\lambda}\right)t\right\}\left(\frac{C_t}{C_0}\right)^{-\gamma},\end{aligned}$$

which is equivalent to the result one obtains by calculating the state price density in the iid case. Note that this result is not automatic, but rather holds because we choose the lower

of the two roots of (A.7).[3]

A1.2.3. An isomorphism with power preferences under the iid assumption

In this section we show that, in an iid model, ratios of the state price density at different times implied by power utility are the same as those implied by recursive utility assuming the discount rate is adjusted appropriately. Thus the power utility model and the recursive utility model are isomorphic when the endowment process is iid.

Let $\pi_{p,t}$ be the state price density assuming power utility with discount rate β_p and relative risk aversion γ. Then

$$\frac{\pi_{p,t}}{\pi_{p,0}} = e^{-\beta_p t}\left(\frac{C_t}{C_0}\right)^{-\gamma}.$$

For convenience, let π_t be the state price density for recursive utility (with EIS equal to one). As shown in Appendix A1.2.2,

$$\frac{\pi_t}{\pi_0} = e^{\left((1-\gamma)\left(-\mu+\frac{1}{2}\gamma\sigma^2\right)-\bar{\lambda}E_\nu\left[e^{(1-\gamma)Z}-1\right]-\beta\right)t}\left(\frac{C_t}{C_0}\right)^{-\gamma}.$$

It follows that, for β given by

$$\beta = \beta_p + (1-\gamma)\left(-\mu+\frac{1}{2}\gamma\sigma^2\right) - \bar{\lambda}E_\nu\left[e^{(1-\gamma)Z}-1\right],$$

ratios of the state price densities are the same.

A1.2.4. Approximating the price-dividend ratio

The formula for the price-dividend ratio in the SDR model is derived by Wachter (2013) and is given by

$$G(\lambda_t) = \int_0^\infty \exp\left\{a_\phi(\tau) + b_\phi(\tau)\lambda_t\right\} d\tau,$$

where $a_\phi(\tau)$ and $b_\phi(\tau)$ have closed-form expressions given in that paper. The algorithm for computing option prices that we use requires that $\log G(\lambda)$ be linear in λ. Define

[3]This point is also made by Tauchen (2005) for a model with stochastic volatility.

$g(\lambda) = \log G(\lambda)$. For a given λ^*, note that for λ_t close to λ^*,

$$g(\lambda) \simeq g(\lambda^*) + (\lambda - \lambda^*)g'(\lambda^*). \tag{A.11}$$

Moreover,

$$\begin{aligned} g'(\lambda^*) &= \frac{G'(\lambda^*)}{G(\lambda^*)} \\ &= \frac{1}{G(\lambda^*)} \int_0^\infty b_\phi(\tau) \exp\left\{a_\phi(\tau) + b_\phi(\tau)\lambda^*\right\} d\tau. \end{aligned} \tag{A.12}$$

The expression (A.12) has an interpretation: it is a weighted average of the coefficients $b_\phi(\tau)$, where the average is over τ, and the weights are proportional to $\exp\left\{a_\phi(\tau) + b_\phi(\tau)\lambda^*\right\}$. With this in mind, we define the notation

$$b_\phi^* = \frac{1}{G(\lambda^*)} \int_0^\infty b_\phi(\tau) \exp\left\{a_\phi(\tau) + b_\phi(\tau)\lambda^*\right\} d\tau \tag{A.13}$$

and the log-linear function

$$\hat{G}(\lambda) = G(\lambda^*) \exp\left\{b_\phi^*(\lambda - \lambda^*)\right\}. \tag{A.14}$$

It follows from exponentiating both sides of (A.11) that

$$G(\lambda) \simeq \hat{G}(\lambda).$$

When we apply this method in this paper, we choose λ^* equal to the long-run mean $\bar{\lambda}$.

This log-linearization method differs from the more widely-used method of Campbell (2003), applied in continuous time by Chacko and Viceira (2005). However, in this application it is more accurate over the relevant range. This is not surprising, since we are able to exploit the fact that the true solution for the price-dividend ratio is known. In dynamic models with the EIS not equal to one, the solution is typically unknown.

Figure A.1 shows implied volatilities from option prices computed using the loglinear approximation described above, and from option prices computed by solving the expectation in (1.10) directly, using by averaging over simulated sample paths. To keep the computation tractable, we assume a single jump size of -30%. The implied volatilities are extremely close in the two cases.

A1.2.5. Transform analysis

The normalized put option price is given as

$$P^n(\lambda_t, T-t; K^n) = E_t\left[\frac{\pi_T}{\pi_t}\left(K^n - \frac{F_T}{F_t}\right)^+\right]. \tag{A.15}$$

It follows from (1.11), (1.12), and (A.14) that

$$\begin{aligned}\frac{\pi_T}{\pi_t} &= \exp\left\{-\int_t^T(\beta b\lambda_s - \eta)ds - \gamma\log\left(\frac{C_T}{C_t}\right) + b(\lambda_T - \lambda_t)\right\}\\ \frac{F_T}{F_t} &= \exp\left\{\phi\log\left(\frac{C_T}{C_t}\right) + b^*_\phi(\lambda_T - \lambda_t)\right\},\end{aligned}$$

where η, b and b^*_ϕ are constants defined by (A.10), (1.7) and (A.13), respectively. Then (A.15) can be rewritten as

$$\begin{aligned}P^n(\lambda_t, T-t; K^n) = {} & E_t\left[e^{-\int_t^T(\beta b\lambda_s-\eta)\,ds-\gamma(\log C_T-\log C_t)+b(\lambda_T-\lambda_t)}K^n\,1_{\left\{\frac{F_T}{F_t}\le K^n\right\}}\right]\\ & - E_t\left[e^{-\int_t^T(\beta b\lambda_s-\eta)\,ds+(\phi-\gamma)(\log C_T-\log C_t)+(b+b^*_\phi)(\lambda_T-\lambda_t)}1_{\left\{\frac{F_T}{F_t}\le K^n\right\}}\right].\end{aligned} \tag{A.16}$$

Note that

$$1_{\left\{\frac{F_T}{F_t}\le K^n\right\}} = 1_{\left\{b^*_\phi(\lambda_T-\lambda_t)+\phi(\log C_T-\log C_t)\le\log K^n\right\}}.$$

Equation (A.16) characterizes the put option in terms of expectations that can be computed using the transform analysis of Duffie, Pan, and Singleton (2000). This analysis requires only the solution of a system of ordinary differential equations and a one-dimensional numerical integration. Below, we describe how we use their analysis.

To use the method of Duffie, Pan, and Singleton (2000), it is helpful to write down the following stochastic process, which, under our assumptions, is well-defined for a given λ_t.

$$X_\tau = \begin{bmatrix} \log C_{t+\tau} - \log C_t \\ \lambda_{t+\tau} \end{bmatrix}.$$

Note that the $\{X_\tau\}$ process is defined purely for mathematical convenience. Further define

$$\begin{aligned} R(X_\tau) &= [0,\, \beta b] X_\tau - \eta = \beta b \lambda_{t+\tau} - \eta \\ d_1 &= [-\gamma,\, b]^\top \\ d_2 &= [\phi,\, b_\phi^*]^\top, \end{aligned}$$

and let

$$\mathcal{G}_{p,q}(y; X_0, T-t) = E\left[e^{-\int_0^{T-t} R(X_\tau) d\tau} e^{p^\top X_{T-t}} 1_{\{q^\top X_{T-t} \le y\}}\right]. \tag{A.17}$$

Note that $\{X_\tau\}$ is an affine process in the sense defined by Duffie et al. It follows that

$$\begin{aligned} P^n(\lambda, T-t; K^n) = e^{-b\lambda} K^n E\left[e^{-\int_0^{T-t} R(X_\tau) d\tau + d_1^\top X_{T-t}}\, 1_{d_2^\top X_{T-t} \le \log K^n + b_\phi^* \lambda}\right] \\ - e^{-(b+b_\phi^*)\lambda} E\left[e^{-\int_0^{T-t} R(X_\tau) d\tau + (d_1+d_2)^\top X_{T-t}}\, 1_{d_2^\top X_{T-t} \le \log K^n + b_\phi^* \lambda}\right], \end{aligned}$$

and therefore

$$\begin{aligned} P^n(\lambda, T-t; K^n) = e^{-b\lambda} \left(K^n \mathcal{G}_{d_1,d_2}\left(\log K^n + b_\phi^* \lambda, X_0, T-t\right)\right. \\ \left. - e^{-b_\phi^* \lambda}\, \mathcal{G}_{d_1+d_2,d_2}\left(\log K^n + b_\phi^* \lambda, X_0, T-t\right)\right), \end{aligned}$$

where $X_0 = [0,\, \lambda]$. The terms written using the function $\mathcal{G}$ can then be computed tractably using the transform analysis of Duffie et al.

A1.3. Solution to the multifactor model

A1.3.1. Utility

We conjecture that the value function is given by

$$J(C,\lambda,\xi) = \frac{C^{1-\gamma}}{1-\gamma} e^{a+b_\lambda \lambda + b_\xi \xi}. \tag{A.18}$$

It follows from the form of $f(C,V)$ that

$$\begin{aligned} f(C,V) &= \beta(1-\gamma)V\left(\log C - \frac{1}{1-\gamma}\log\left[(1-\gamma)V\right]\right) \\ &= \beta(1-\gamma)V\log C - \beta V \log\left[(1-\gamma)V\right] \\ &= \beta V \log\left(\frac{C^{1-\gamma}}{(1-\gamma)V}\right) \\ &= -\beta V(a + b_\lambda\lambda + b_\xi\xi), \end{aligned}$$

where the last equation follows from the equilibrium condition that the utility process is equal to the value function under the optimal policies: $V_t = J(C_t, \lambda_t, \xi_t)$.

By differentiating $J(C,\lambda,\xi)$, we obtain

$$\begin{aligned} \frac{\partial J}{\partial C} &= (1-\gamma)\frac{J}{C}, & \frac{\partial^2 J}{\partial C^2} &= -\gamma(1-\gamma)\frac{J}{C^2}, \\ \frac{\partial J}{\partial \lambda} &= b_\lambda J, & \frac{\partial^2 J}{\partial \lambda^2} &= b_\lambda^2 J, \\ \frac{\partial J}{\partial \xi} &= b_\xi J, & \frac{\partial^2 J}{\partial \xi^2} &= b_\xi^2 J. \end{aligned} \tag{A.19}$$

Applying Ito's Lemma to $J(C,\lambda,\xi)$ with conjecture (A.18) and derivatives (A.19):

$$\begin{aligned} \frac{dV_t}{V_{t^-}} = (1-\gamma)(\mu dt + \sigma dB_t) - \frac{1}{2}\gamma(1-\gamma)\sigma^2 dt \\ + b_\lambda\left(\kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t}\right) + \frac{1}{2}b_\lambda^2\sigma_\lambda^2\lambda_t dt \\ + b_\xi\left(\kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t}\right) + \frac{1}{2}b_\xi^2\sigma_\xi^2\xi_t dt + (e^{(1-\gamma)Z_t} - 1)dN_t. \end{aligned}$$

Under the optimal consumption path, it must be that

$$V_t + \int_0^t f(C_s, V_s)ds = E_t\left[\int_0^\infty f(C_s, V_s)ds\right] \tag{A.20}$$

(see Duffie and Epstein (1992)). By the law of iterative expectations, the left-hand side of (A.20) is a martingale. Thus, the sum of the drift and the jump compensator of $(V_t + \int_0^t f(C_s, V_s)ds)$ equals zero. That is,

$$\begin{aligned} 0 = (1-\gamma)\mu - \frac{1}{2}\gamma(1-\gamma)\sigma^2 + b_\lambda\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2}b_\lambda^2\sigma_\lambda^2\lambda_t + b_\xi\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2}b_\xi^2\sigma_\xi^2\xi_t \\ + \lambda_t E_\nu\left[e^{(1-\gamma)Z_t} - 1\right] - \beta(a + b_\lambda\lambda_t + b_\xi\xi_t). \end{aligned} \tag{A.21}$$

By collecting terms in (A.21), we obtain

$$\begin{aligned} 0 = \underbrace{\left[(1-\gamma)\mu - \frac{1}{2}\gamma(1-\gamma)\sigma^2 + b_\xi\kappa_\xi\bar{\xi} - \beta a\right]}_{=0} \\ + \lambda_t \underbrace{\left[-b_\lambda\kappa_\lambda + \frac{1}{2}b_\lambda^2\sigma_\lambda^2 + E_\nu\left[e^{(1-\gamma)Z_t} - 1\right] - \beta b_\lambda\right]}_{=0} \\ + \xi_t \underbrace{\left[b_\lambda\kappa_\lambda - b_\xi\kappa_\xi + \frac{1}{2}b_\xi^2\sigma_\xi^2 - \beta b_\xi\right]}_{=0}. \end{aligned}$$

Solving these equations gives us

$$\begin{aligned} a &= \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + \frac{b_\xi\kappa_\xi\bar{\xi}}{\beta} \\ b_\lambda &= \frac{\kappa_\lambda + \beta}{\sigma_\lambda^2} - \sqrt{\left(\frac{\kappa_\lambda + \beta}{\sigma_\lambda^2}\right)^2 - 2\frac{E_\nu\left[e^{(1-\gamma)Z_t} - 1\right]}{\sigma_\lambda^2}} \\ b_\xi &= \frac{\kappa_\xi + \beta}{\sigma_\xi^2} - \sqrt{\left(\frac{\kappa_\xi + \beta}{\sigma_\xi^2}\right)^2 - 2\frac{b_\lambda\kappa_\lambda}{\sigma_\xi^2}}, \end{aligned}$$

where we have chosen the negative root based on the economic consideration that when there are no disasters, λ_t and ξ_t should not appear in the value function. Namely, for

$Z_t = 0$, $b_\lambda = b_\xi = 0$. Lastly, note that these results verify the conjecture (A.18).

A1.3.2. State-price density

Following the same steps for the SDR model, we can show that

$$\begin{aligned}\frac{\partial}{\partial C} f(C_t, V_t) &= \beta C_t^{-\gamma} e^{a + b_\lambda \lambda_t + b_\xi \xi_t} \\ \frac{\partial}{\partial V} f(C_t, V_t) &= \eta - \beta b_\lambda \lambda_t - \beta b_\xi \xi_t.\end{aligned}$$

Therefore, it follows that the state-price density can be written as

$$\pi_t = \exp\left(\eta t - \beta b_\lambda \int_0^t \lambda_s ds - \beta b_\xi \int_0^t \xi_s ds\right) \beta C_t^{-\gamma} e^{a + b_\lambda \lambda_t + b_\xi \xi_t}. \tag{A.22}$$

By applying Ito's Lemma to (A.22), we derive the following stochastic differential equation:

$$\begin{aligned}\frac{d\pi_t}{\pi_{t^-}} = \left\{-\beta - \mu + \gamma\sigma^2 - \lambda_t E_\nu\left[e^{(1-\gamma)Z_t} - 1\right]\right\} dt \\ - \gamma\sigma dB_t + b_\lambda \sigma_\lambda \sqrt{\lambda_t} dB_{\lambda,t} + b_\xi \sigma_\xi \sqrt{\xi_t} dB_{\xi,t} + (e^{-\gamma Z_t} - 1) dN_t.\end{aligned}$$

In equilibrium, the sum of the drift and the jump compensator of the state-price density growth must equal the negative of the riskfree rate. The riskfree rate therefore equals

$$r_t = \beta + \mu - \gamma\sigma^2 + \lambda_t E_\nu\left[e^{(1-\gamma)Z_t} - e^{-\gamma Z_t}\right].$$

A1.3.3. Dividend claim price

Let F_t denote the price of the dividend claim. The pricing relation implies

$$\begin{aligned}F_t &= E_t\left[\int_t^\infty \frac{\pi_s}{\pi_t} D_s ds\right] \\ &= \int_t^\infty E_t\left[\frac{\pi_s}{\pi_t} D_s\right] ds.\end{aligned}$$

Let $H(D_t, \lambda_t, \xi_t, s-t)$ denote the price of the asset that pays the aggregate dividend at time s, namely,

$$H(D_t, \lambda_t, \xi_t, s-t) = E_t\left[\frac{\pi_s}{\pi_t} D_s\right].$$

By the law of iterative expectations, it follows that $\pi_t H_t$ is a martingale:

$$\pi_t H(D_t, \lambda_t, \xi_t, s-t) = E_t[\pi_s D_s].$$

Conjecture that

$$H(D_t, \lambda_t, \xi_t, \tau) = D_t \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda_t + b_{\phi\xi}(\tau)\xi_t\right). \tag{A.23}$$

Ito's Lemma implies

$$\begin{aligned}\frac{dH_t}{H_{t-}} = \Big\{ & \mu_D + b_{\phi\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2} b_{\phi\lambda}(\tau)^2 \sigma_\lambda^2 \lambda_t + b_{\phi\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2} b_{\phi\xi}(\tau)^2 \sigma_\xi^2 \xi_t \\ & - a'_\phi(\tau) - b'_{\phi\lambda}(\tau)\lambda_t - b'_{\phi\xi}(\tau)\xi_t \Big\} dt \\ & + \phi\sigma dB_t + b_{\phi\lambda}(\tau)\sigma_\lambda \sqrt{\lambda_t} dB_{\lambda,t} + b_{\phi\xi}(\tau)\sigma_\xi \sqrt{\xi_t} dB_{\xi,t} + (e^{\phi Z_t} - 1) dN_t.\end{aligned}$$

Combining the SDE for H_t with the one for π_t derived in the previous sections, we can derive the SDE for $\pi_t H_t$:

$$\begin{aligned}\frac{d(\pi_t H_t)}{\pi_{t^-} H_{t^-}} = \Big\{ &-\beta - \mu + \gamma\sigma^2 - \lambda_t E_\nu \left[e^{(1-\gamma)Z_t} - 1\right] \\ &+ \mu_D + b_{\phi\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2} b_{\phi\lambda}(\tau)^2 \sigma_\lambda^2 \lambda_t \\ &+ b_{\phi\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2} b_{\phi\xi}(\tau)^2 \sigma_\xi^2 \xi_t \\ &- a'_\phi(\tau) - b'_{\phi\lambda}(\tau)\lambda_t - b'_{\phi\xi}(\tau)\xi_t \\ &- \gamma\phi\sigma^2 + b_\lambda b_{\phi\lambda}(\tau)\sigma_\lambda^2 \lambda_t + b_\xi b_{\phi\xi}(\tau)\sigma_\xi^2 \xi_t \Big\} dt \\ &+ (\phi - \gamma)\sigma dB_t + (b_\lambda + b_{\phi\lambda}(\tau))\sigma_\lambda \sqrt{\lambda_t} dB_{\lambda,t} + (b_\xi + b_{\phi\xi}(\tau))\sigma_\xi \sqrt{\xi_t} dB_{\xi,t} \\ &+ (e^{(\phi-\gamma)Z_t} - 1) dN_t.\end{aligned}$$

Since $\pi_t H_t$ is a martingale, the sum of the drift and the jump compensator of $\pi_t H_t$ equals zero. Thus:

$$\begin{aligned}0 = &-\beta - \mu + \gamma\sigma^2 - \lambda_t E_\nu \left[e^{(1-\gamma)Z_t} - 1\right] \\ &+ \mu_D + b_{\phi\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2} b_{\phi\lambda}(\tau)^2 \sigma_\lambda^2 \lambda_t \\ &+ b_{\phi\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2} b_{\phi\xi}(\tau)^2 \sigma_\xi^2 \xi_t \\ &- a'_\phi(\tau) - b'_{\phi\lambda}(\tau)\lambda_t - b'_{\phi\xi}(\tau)\xi_t \\ &- \gamma\phi\sigma^2 + b_\lambda b_{\phi\lambda}(\tau)\sigma_\lambda^2 \lambda_t + b_\xi b_{\phi\xi}(\tau)\sigma_\xi^2 \xi_t + \lambda_t E_\nu \left[e^{(\phi-\gamma)Z_t} - 1\right]. \quad \text{(A.24)}\end{aligned}$$

Collecting terms of (A.24) results in the following equation:

$$
\begin{aligned}
0 = & \underbrace{\left[-\beta - \mu + \gamma\sigma^2 + \mu_D + b_{\phi\xi}(\tau)\kappa_\xi\bar{\xi} - \gamma\phi\sigma^2 - a'_\phi(\tau)\right]}_{=0} \\
& + \lambda_t \underbrace{\left[-b_{\phi\lambda}(\tau)\kappa_\lambda + \frac{1}{2}b_{\phi\lambda}(\tau)^2\sigma_\lambda^2 + b_\lambda b_{\phi\lambda}(\tau)\sigma_\lambda^2 + E_\nu\left[e^{(\phi-\gamma)Z_t} - e^{(1-\gamma)Z_t}\right] - b'_{\phi\lambda}(\tau)\right]}_{=0} \\
& + \xi_t \underbrace{\left[b_{\phi\lambda}(\tau)\kappa_\lambda - b_{\phi\xi}(\tau)\kappa_\xi + \frac{1}{2}b_{\phi\xi}(\tau)^2\sigma_\xi^2 + b_\xi b_{\phi\xi}(\tau)\sigma_\xi^2 - b'_{\phi\xi}(\tau)\right]}_{=0}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
a'_\phi(\tau) &= \mu_D - \mu - \beta + \gamma\sigma^2(1-\phi) + \kappa_\xi\bar{\xi}b_{\phi\xi}(\tau) \\
b'_{\phi\lambda}(\tau) &= \frac{1}{2}\sigma_\lambda^2 b_{\phi\lambda}(\tau)^2 + (b_\lambda\sigma_\lambda^2 - \kappa_\lambda)b_{\phi\lambda}(\tau) + E_\nu\left[e^{(\phi-\gamma)Z_t} - e^{(1-\gamma)Z_t}\right] \qquad \text{(A.25)} \\
b'_{\phi\xi}(\tau) &= \frac{1}{2}\sigma_\xi^2 b_{\phi\xi}(\tau)^2 + (b_\xi\sigma_\xi^2 - \kappa_\xi)b_{\phi\xi}(\tau) + \kappa_\lambda b_{\phi\lambda}(\tau).
\end{aligned}
$$

This establishes that H satisfies the conjecture (A.23). We note that by no-arbitrage,

$$H(D_t, \lambda_t, \xi_t, 0) = D_t.$$

This condition provides the boundary conditions for the system of ODEs (A.25):

$$a_\phi(0) = b_{\phi\lambda}(0) = b_{\phi\xi}(0) = 0.$$

Recall that once we get $a_\phi(\tau)$, $b_{\phi\lambda}(\tau)$, and $b_{\phi\xi}(\tau)$,

$$
\begin{aligned}
F_t &= \int_t^\infty E_t\left[\frac{\pi_s}{\pi_t}D_s\right]ds \\
&= \int_t^\infty H(D_t, \lambda_t, \xi_t, s-t)ds \\
&= D_t\int_t^\infty \exp\left(a_\phi(s-t) + b_{\phi\lambda}(s-t)\lambda_t + b_{\phi\xi}(s-t)\xi_t\right)ds \\
&= D_t\int_0^\infty \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda_t + b_{\phi\xi}(\tau)\xi_t\right)d\tau.
\end{aligned}
$$

That is, the price-dividend ratio can be written as

$$G(\lambda_t, \xi_t) = \int_0^\infty \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda_t + b_{\phi\xi}(\tau)\xi_t\right) d\tau.$$

A1.3.4. Approximating the price-dividend ratio

Let $g(\lambda, \xi) = \log G(\lambda, \xi)$. For given λ^* and ξ^*, the two-dimensional Taylor approximation implies

$$g(\lambda, \xi) \simeq g(\lambda^*, \xi^*) + \left.\frac{\partial g}{\partial \lambda}\right|_{\lambda^*, \xi^*} (\lambda - \lambda^*) + \left.\frac{\partial g}{\partial \xi}\right|_{\lambda^*, \xi^*} (\xi - \xi^*). \tag{A.26}$$

We note that

$$\begin{aligned}
\left.\frac{\partial g}{\partial \lambda}\right|_{\lambda^*, \xi^*} &= \frac{1}{G(\lambda^*, \xi^*)} \left.\frac{\partial G}{\partial \lambda}\right|_{\lambda^*, \xi^*} \\
&= \frac{1}{G(\lambda^*, \xi^*)} \int_0^\infty b_{\phi\lambda}(\tau) \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda^* + b_{\phi\xi}(\tau)\xi^*\right) d\tau.
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
\left.\frac{\partial g}{\partial \xi}\right|_{\lambda^*, \xi^*} &= \frac{1}{G(\lambda^*, \xi^*)} \left.\frac{\partial G}{\partial \xi}\right|_{\lambda^*, \xi^*} \\
&= \frac{1}{G(\lambda^*, \xi^*)} \int_0^\infty b_{\phi\xi}(\tau) \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda^* + b_{\phi\xi}(\tau)\xi^*\right) d\tau.
\end{aligned}$$

We define the notation

$$\begin{aligned}
b^*_{\phi\lambda} &= \frac{1}{G(\lambda^*, \xi^*)} \int_0^\infty b_{\phi\lambda}(\tau) \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda^* + b_{\phi\xi}(\tau)\xi^*\right) d\tau \\
b^*_{\phi\xi} &= \frac{1}{G(\lambda^*, \xi^*)} \int_0^\infty b_{\phi\xi}(\tau) \exp\left(a_\phi(\tau) + b_{\phi\lambda}(\tau)\lambda^* + b_{\phi\xi}(\tau)\xi^*\right) d\tau,
\end{aligned}$$

and the log-linear function

$$\hat{G}(\lambda, \xi) = G(\lambda^*, \xi^*) \exp\left\{b^*_{\phi\lambda}(\lambda - \lambda^*) + b^*_{\phi\xi}(\xi - \xi^*)\right\}.$$

It follows from exponentiating both sides of (A.26) that

$$G(\lambda, \xi) \simeq \hat{G}(\lambda, \xi).$$

In our analysis, we pick λ^* and ξ^* as the stationary mean of λ_t and ξ_t, respectively.

A2. Appendix for Chapter 2

A2.1. Value function

We note that the state variables in the economy are C_t, λ_t, and ξ_t. Since the value function is homogeneous of degree $(1-\gamma)$ in consumption and the EIS is equal to 1, we conjecture the value function $J(C, \lambda, \xi)$ is given by

$$J(C, \lambda, \xi) = \frac{C^{1-\gamma}}{1-\gamma} e^{a + b_\lambda \lambda + b_\xi \xi}. \tag{A.27}$$

At equilibrium, the continuation utility is equal to the value function under the optimal policies:

$$V_t = J(C_t, \lambda_t, \xi_t).$$

This optimality condition implies that

$$\begin{aligned} f(C, V) &= \beta(1-\gamma)V\left(\log C - \frac{1}{1-\gamma}\log\left[(1-\gamma)V\right]\right) \\ &= \beta V \log\left(\frac{C^{1-\gamma}}{(1-\gamma)V}\right) \\ &= -\beta V(a + b_\lambda \lambda + b_\xi \xi), A \end{aligned}$$

By differentiating $J(C, \lambda, \xi)$, we obtain

$$\begin{aligned} \frac{\partial J}{\partial C} &= (1-\gamma)\frac{J}{C}, & \frac{\partial^2 J}{\partial C^2} &= -\gamma(1-\gamma)\frac{J}{C^2}, \\ \frac{\partial J}{\partial \lambda} &= b_\lambda J, & \frac{\partial^2 J}{\partial \lambda^2} &= b_\lambda^2 J, \\ \frac{\partial J}{\partial \xi} &= b_\xi J, & \frac{\partial^2 J}{\partial \xi^2} &= b_\xi^2 J. \end{aligned} \tag{A.28}$$

Applying Ito's Lemma to $J(C, \lambda, \xi)$ with conjecture (A.27) and derivatives (A.28):

$$\begin{aligned}\frac{dV_t}{V_{t^-}} = (1-\gamma)(\mu_C dt + \sigma_C dB_{C,t}) - \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 dt \\ + b_\lambda\left(\kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t}\right) + \frac{1}{2}b_\lambda^2\sigma_\lambda^2\lambda_t dt \\ + b_\xi\left(\kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dB_{\xi,t}\right) + \frac{1}{2}b_\xi^2\sigma_\xi^2\xi_t dt + (e^{(1-\gamma)Z_{C,t}} - 1)dN_{C,t}. \quad \text{(A.29)}\end{aligned}$$

By adding $\int_0^t f(C_s, V_s)ds$ on both sides of (A.29), we obtain

$$V_t + \int_0^t f(C_s, V_s)ds = E_t\left[\int_0^\infty f(C_s, V_s)ds\right]. \quad \text{(A.30)}$$

By the law of iterative expectations, the left-hand side of (A.30) is a martingale. Thus, the sum of the drift and the jump compensator of $(V_t + \int_0^t f(C_s, V_s)ds)$ equals zero. That is,

$$\begin{aligned}0 = (1-\gamma)\mu_C - \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 + b_\lambda\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2}b_\lambda^2\sigma_\lambda^2\lambda_t + b_\xi\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2}b_\xi^2\sigma_\xi^2\xi_t \\ + \lambda_t E\left[e^{(1-\gamma)Z_{C,t}} - 1\right] - \beta(a + b_\lambda\lambda_t + b_\xi\xi_t). \quad \text{(A.31)}\end{aligned}$$

By collecting terms in (A.31), we obtain

$$\begin{aligned}0 = \underbrace{\left[(1-\gamma)\mu_C - \frac{1}{2}\gamma(1-\gamma)\sigma_C^2 + b_\xi\kappa_\xi\bar{\xi} - \beta a\right]}_{=0} \\ + \lambda_t\underbrace{\left[-b_\lambda\kappa_\lambda + \frac{1}{2}b_\lambda^2\sigma_\lambda^2 + E\left[e^{(1-\gamma)Z_{C,t}} - 1\right] - \beta b_\lambda\right]}_{=0} \\ + \xi_t\underbrace{\left[b_\lambda\kappa_\lambda - b_\xi\kappa_\xi + \frac{1}{2}b_\xi^2\sigma_\xi^2 - \beta b_\xi\right]}_{=0}.\end{aligned}$$

Solving these equations gives us

$$\begin{aligned} a &= \frac{1-\gamma}{\beta}\left(\mu_C - \frac{1}{2}\gamma\sigma_C^2\right) + \frac{b_\xi \kappa_\xi \bar{\xi}}{\beta} \\ b_\lambda &= \frac{\kappa_\lambda + \beta}{\sigma_\lambda^2} - \sqrt{\left(\frac{\kappa_\lambda + \beta}{\sigma_\lambda^2}\right)^2 - 2\frac{E\left[e^{(1-\gamma)Z_{C,t}} - 1\right]}{\sigma_\lambda^2}} \\ b_\xi &= \frac{\kappa_\xi + \beta}{\sigma_\xi^2} - \sqrt{\left(\frac{\kappa_\xi + \beta}{\sigma_\xi^2}\right)^2 - 2\frac{b_\lambda \kappa_\lambda}{\sigma_\xi^2}}, \end{aligned}$$

where we have chosen the negative root based on the economic consideration that when there are no disasters, λ and ξ should not appear in the value function. Namely, for $Z_{C,t} = 0$, $b_\lambda = b_\xi = 0$. Lastly, note that these results verify the conjecture (A.27).

A2.2. The state-price density

By Duffie and Skiadas (1994), the state-price density π_t can be written as

$$\pi_t = \exp\left(\int_0^t \frac{\partial}{\partial V} f\left(C_t, V_t\right) ds\right) \frac{\partial}{\partial C} f\left(C_t, V_t\right). \tag{A.32}$$

By differentiating equation (2.1) with respect to C and V, we can show that

$$\begin{aligned} \frac{\partial}{\partial C} f\left(C_t, V_t\right) &= \beta(1-\gamma)\frac{V_t}{C_t} = \beta C_t^{-\gamma} e^{a + b_\lambda \lambda_t + b_\xi \xi_t} \\ \frac{\partial}{\partial V} f\left(C_t, V_t\right) &= \beta \log\left[\frac{C^{1-\gamma}}{(1-\gamma)V}\right] - \beta = \eta - \beta b_\lambda \lambda_t - \beta b_\xi \xi_t, \end{aligned}$$

where

$$\eta = -\beta a - \beta.$$

Therefore, it follows that the state-price density can be expressed as

$$\pi_t = \exp\left(\eta t - \beta b_\lambda \int_0^t \lambda_s ds - \beta b_\xi \int_0^t \xi_s ds\right) \beta C_t^{-\gamma} e^{a + b_\lambda \lambda_t + b_\xi \xi_t}.$$

Furthermore, by applying Ito's Lemma to equation (2.2), we derive the following stochastic differential equation:

$$\frac{d\pi_t}{\pi_{t^-}} = \left\{-\beta - \mu_C + \gamma\sigma_C^2 - \lambda_t E\left[e^{(1-\gamma)Z_{C,t}} - 1\right] - \nu E\left[e^{b_\lambda Z_{\lambda,t}} - 1\right]\right\} dt$$
$$- \gamma\sigma_C dB_{C,t} + b_\lambda\sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} + b_\xi\sigma_\xi\sqrt{\xi_t}dB_{\xi,t}$$
$$+ (e^{-\gamma Z_{C,t}} - 1)dN_{C,t} + (e^{b_\lambda Z_{\lambda,t}} - 1)dN_{\lambda,t}. \quad \text{(A.33)}$$

At equilibrium, the sum of the drift and the jump compensator of the state-price density growth must equal the riskfree rate multiplied by -1. It thus follows that the riskfree rate is given by

$$r_t = \beta + \mu - \gamma\sigma^2 + \lambda_t E\left[e^{(1-\gamma)Z_t} - e^{-\gamma Z_t}\right].$$

A2.3. Individual firm value dynamics

Let $H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, s-t)$ denote the time-t value of firm i's payoff at time s. That is,

$$H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, s-t) = E_t\left[\frac{\pi_s}{\pi_t}D_{i,s}\right].$$

We conjecture that $H_i(\cdot)$ has the following functional form:

$$H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, \tau) = D_{i,t}\exp\left(a_i(\tau) + b_{i\lambda}(\tau)\lambda_t + b_{i\xi}(\tau)\xi_t + b_{i\chi}(\tau)\chi_t\right).$$

To verify this conjecture, I apply Ito's lemma to the process $\pi_t H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, s-t)$ and derive its conditional mean (which is the sum of its drift and jump compensator). The conditional mean of this process always equals zero because the process is a martingale.[4]

[4]Equation (A.34) implies that $\pi_t H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, s-t) = E_t[\pi_s D_{i,s}]$. It is straightforward that $E_t[\pi_s D_{i,s}]$ is a martingale by the law of iterative expectations.

This zero mean condition provides the system of ODEs for $a_i(\tau)$, $b_{i\lambda}(\tau)$, $b_{i\xi}(\tau)$, and $b_{i\chi}(\tau)$:

$$\begin{aligned}
a_i'(\tau) &= -\beta - \mu_C - \gamma(\phi_i - 1)\sigma_C^2 + \mu_i + b_{i\xi}(\tau)\kappa_\xi\bar{\xi} + b_{i\chi}(\tau)\kappa_\chi\bar{\chi} \\
&\qquad + \nu E\left[e^{(b_\lambda + b_{i\lambda}(\tau))Z_{\lambda,t} + b_{i\chi}(\tau)Z_{\chi,t}} - e^{b_\lambda Z_{\lambda,t}}\right] \\
b_{i\lambda}'(\tau) &= -b_{i\lambda}(\tau)\kappa_\lambda + \frac{1}{2}b_{i\lambda}(\tau)^2\sigma_\lambda^2 + b_\lambda b_{i\lambda}(\tau)\sigma_\lambda^2 + E\left[e^{(\phi_i-\gamma)Z_{C,t}} - e^{(1-\gamma)Z_{C,t}}\right] \\
b_{i\xi}'(\tau) &= b_{i\lambda}(\tau)\kappa_\lambda - b_{i\xi}(\tau)\kappa_\xi + \frac{1}{2}b_{i\xi}(\tau)^2\sigma_\xi^2 + b_\xi b_{i\xi}(\tau)\sigma_\xi^2 \\
b_{i\chi}'(\tau) &= -b_{i\chi}(\tau)\kappa_\chi + \frac{1}{2}b_{i\chi}(\tau)^2\sigma_\chi^2 + E\left[e^{Z_t^i} - 1\right].
\end{aligned}$$

Therefore, equation (2.3) can be written as

$$\begin{aligned}
A_{i,t} &= \int_t^\infty H_i(D_{i,t}, \lambda_t, \xi_t, \chi_t, s-t)ds \\
&= D_{i,t}\int_t^\infty \exp\left(a_i(s-t) + b_{i\lambda}(s-t)\lambda_t + b_{i\xi}(s-t)\xi_t + b_{i\chi}(s-t)\chi_t\right)ds \\
&= D_{i,t}\int_0^\infty \exp\left(a_i(\tau) + b_{i\lambda}(\tau)\lambda_t + b_{i\xi}(\tau)\xi_t + b_{i\chi}(\tau)\chi_t\right)d\tau.
\end{aligned}$$

This subsequently implies that that the asset-payout ratio ($G_i = A_{i,t}/D_{i,t}$) is expressed as a function of three state variables as can be seen in equation (2.4).

A2.4. Pricing of CDX products based on the model

In this appendix, we provide our approach for pricing CDX products (the CDX index and its tranche products) based on our equilibrium model. To make the computation tractable, we discretize the premium and protection legs of each product by assuming that defaults occur on average in the middle of premium payment dates (see, e.g., Mortensen (2006)) (Section A2.4.1). This discretization is relatively accurate because premium payment interval $\Delta_m = t_m - t_{m-1}$ is relatively small ($\Delta_m = 0.25$) since premium payments are quarterly. Then, we express these discretized legs ($\mathbf{Prot}_{\text{CDX}}$, $\mathbf{Prem}_{\text{CDX}}$, $\mathbf{Prot}_{\text{Tran},j}$, and $\mathbf{Prem}_{\text{Tran},j}$) in terms of four expectations under the physical measure (Section A2.4.2). We calculate these four expectations using a simulation-based approach. (Section A2.4.3).

A2.4.1. Discretization of pricing formulas

First consider the discretization of the premium leg ($\mathbf{Prot}_{\mathrm{CDX}}$) and the premium leg ($\mathbf{Prem}_{\mathrm{CDX}}$) for CDX pricing. of a CDX contract derived in Section 2.2.3. With the discretization assumption explained above, it follows that

$$\begin{aligned}
\mathbf{Prot}_{\mathrm{CDX}} &= \tilde{E}\left[\int_0^T e^{-\int_0^t r_s ds} dL_t\right] \\
&\simeq \sum_{m=1}^{M} \tilde{E}\left[e^{-\int_0^{t_m-\Delta_m/2} r_s ds}(L_{t_m} - L_{t_{m-1}})\right] \\
&= \sum_{m=1}^{M} \tilde{E}\left[e^{-\int_0^{t_m-\Delta_m/2} r_s ds} L_{t_m}\right] - \sum_{m=1}^{M} \tilde{E}\left[e^{-\int_0^{t_m-\Delta_m/2} r_s ds} L_{t_{m-1}}\right].
\end{aligned}$$

Since Δ_m is small, the risk-free rate is unlikely to change much between time $t_m - \Delta_m/2$ and t_m. Therefore, we can approximate

$$\int_{t_m-\Delta_m/2}^{t_m} r_s ds \simeq \frac{1}{2}\Delta_m r_{t_m}.$$

which allows us to further simply the protection leg ($\mathbf{Prot}_{\mathrm{CDX}}$) as

$$\mathbf{Prot}_{\mathrm{CDX}} \simeq \sum_{m=1}^{M} \left\{\tilde{E}\left[e^{\frac{1}{2}\Delta_m r_{t_m}} e^{-\int_0^{t_m} r_s ds} L_{t_m}\right] - \tilde{E}\left[e^{-\frac{1}{2}\Delta_m r_{t_{m-1}}} e^{-\int_0^{t_{m-1}} r_s ds} L_{t_{m-1}}\right]\right\}.$$

We discretize the premium leg ($\mathbf{Prem}_{\mathrm{CDX}}(S)$) using the same assumption:

$$\begin{aligned}
\mathbf{Prem}_{\mathrm{CDX}}(S) &= S\tilde{E}\left[\sum_{m=1}^{M} e^{-\int_0^{t_m} r_s ds}(1 - n_{t_m})\Delta_m + \int_{t_{m-1}}^{t_m} e^{-\int_0^u r_s ds}(u - t_{m-1}) dn_s\right] \\
&\simeq S\sum_{m=1}^{M} \Delta_m \tilde{E}\left[e^{-\int_0^{t_m} r_s ds}(1 - n_{t_m}) + e^{-\int_0^{t_m-\Delta_m/2} r_s ds}\frac{n_{t_m} - n_{t_{m-1}}}{2}\right].
\end{aligned}$$

Since $\Delta_m/2$ is small, we further approximate this expression as

$$\begin{aligned}
\mathbf{Prem}_{\mathrm{CDX}}(S) &\simeq S\sum_{m=1}^{M}\Delta_m\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}\left(1-\frac{1}{2}n_{t_m}-\frac{1}{2}n_{t_{m-1}}\right)\right] \\
&= S\sum_{m=1}^{M}\Delta_m\left\{H_0(t_m)-\frac{1}{2}\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}n_{t_m}\right]-\frac{1}{2}\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}n_{t_{m-1}}\right]\right\} \\
&\simeq S\sum_{m=1}^{Tf}\Delta_m\left\{H_0(t_m)-\frac{1}{2}\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}n_{t_m}\right]-\frac{1}{2}\tilde{E}\left[e^{-\Delta_m r_{t_{m-1}}e^{-\int_0^{t_{m-1}} r_s ds}}n_{t_{m-1}}\right]\right\}.
\end{aligned}$$

where $H_0(\tau)=H_0(\lambda,\xi,\tau)$ is the price of the default-free zero-coupon bond with maturity τ, which we derive in Appendix A2.5.

While nearly all pricing formulas for credit derivatives under a reduced-form setup assume that interest rates are uncorrelated with defaults, we cannot make this simplifying assumption. This is because the systematic risk under our equilibrium model simultaneously affects both interest rates and the likelihood of defaults.

With the same discretization approach, we also obtain the discretization of the premium leg ($\mathbf{Prot}_{\mathrm{Tran},j}$) and the premium leg ($\mathbf{Prem}_{\mathrm{Tran},j}(U,S)$) for CDX tranche pricing. It is straightforward to show that

$$\begin{aligned}
\mathbf{Prot}_{\mathrm{Tran},j} &\simeq \sum_{m=1}^{Tf}\tilde{E}\left[e^{-\int_0^{t_m-\Delta_m/2} r_s ds}(T^L_{j,t_m}-T^L_{j,t_{m-1}})\right] \\
&\simeq \sum_{m=1}^{Tf}\left\{\tilde{E}\left[e^{\frac{1}{2}\Delta_m r_{t_m}}e^{-\int_0^{t_m} r_s ds}T^L_{j,t_m}\right]-\tilde{E}\left[e^{-\frac{1}{2}\Delta_m r_{t_{m-1}}}e^{-\int_0^{t_{m-1}} r_s ds}T^L_{j,t_{m-1}}\right]\right\}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{Prem}_{\mathrm{Tran},j}(U,S) &\simeq U+S\sum_{m=1}^{Tf}\Delta_m\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}\frac{(1-T^L_{j,t_m}-T^R_{j,t_m})+(1-T^L_{j,t_{m-1}}-T^R_{j,t_{m-1}})}{2}\right] \\
&\simeq U+S\sum_{m=1}^{Tf}\Delta_m\left\{H_0(t_m)-\frac{1}{2}\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}T^L_{j,t_m}\right]-\frac{1}{2}\tilde{E}\left[e^{-\int_0^{t_m} r_s ds}T^R_{j,t_m}\right]\right. \\
&\quad \left.-\frac{1}{2}\tilde{E}\left[e^{-\Delta_m r_{t_{m-1}}}e^{-\int_0^{t_{m-1}} r_s ds}T^L_{j,t_{m-1}})\right]-\frac{1}{2}\tilde{E}\left[e^{-\Delta_m r_{t_{m-1}}}e^{-\int_0^{t_{m-1}} r_s ds}T^R_{j,t_{m-1}}\right]\right\}.
\end{aligned}$$

A2.4.2. Pricing formulas in terms of four expectations

For notational convenience, let X_0 denote the state vector at time 0, namely,

$$X_0 = \begin{bmatrix} \lambda_0 \\ \xi_0 \\ \chi_0 \end{bmatrix}.$$

We also define the following four expectations:

$$\begin{aligned}
\mathrm{EDR}(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} e^{-\int_0^t r_s ds} n_t\right] \\
\mathrm{ELR}(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} e^{-\int_0^t r_s ds} L_t\right] \\
\mathrm{ETLR}_j(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} e^{-\int_0^t r_s ds} T^L_{j,t}\right] \\
\mathrm{ETRR}_j(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} e^{-\int_0^t r_s ds} T^R_{j,t}\right] \quad \forall u \in \mathbb{R}.
\end{aligned} \tag{A.34}$$

Using these notations, we can re-write the pricing formulas for the CDX index and its tranches as the followings:

$$\begin{aligned}
\mathbf{Prot}_{\mathrm{CDX}}(X_0,T) &= \sum_{m=1}^{Tf} \left\{ \mathrm{ELR}\left(\frac{\Delta_m}{2}, t_m, X_0\right) - \mathrm{ELR}\left(-\frac{\Delta_m}{2}, t_{m-1}, X_0\right) \right\} \\
\mathbf{Prem}_{\mathrm{CDX}}(X_0,T) &= U_{\mathrm{CDX}} + S_{\mathrm{CDX}} \sum_{m=1}^{Tf} \Delta_m \left\{ H_0(t_m) - \frac{1}{2}\mathrm{EDR}\left(0, t_m, X_0\right) \right. \\
&\qquad \left. - \frac{1}{2}\mathrm{EDR}\left(\Delta_m, t_{m-1}, X_0\right) \right\} \\
\mathbf{Prot}_{\mathrm{Tran},j}(X_0,T) &= \sum_{m=1}^{Tf} \left\{ \mathrm{ETLR}_j\left(\frac{\Delta_m}{2}, t_m, X_0\right) - \mathrm{ETLR}_j\left(-\frac{\Delta_m}{2}, t_{m-1}, X_0\right) \right\} \\
\mathbf{Prem}_{\mathrm{Tran},j}(X_0,T) &= U_{\mathrm{Tran},j} + S_{\mathrm{Tran},j} \sum_{m=1}^{Tf} \Delta_m \left\{ H_0(t_m) - \frac{\left[\mathrm{ETLR}_j + \mathrm{ETRR}_j\right]\left(0, t_m, X_0\right)}{2} \right. \\
&\qquad \left. - \frac{\left[\mathrm{ETLR}_j + \mathrm{ETRR}_j\right]\left(\Delta_m, t_{m-1}, X_0\right)}{2} \right\}.
\end{aligned}$$

That is, the CDX index and its tranches are priced if we are able to calculate four expectations above. Since there are multiple firms in the pool, it is impossible to compute these expectations in closed-form. We therefore use Monte Carlo simulation. This approach is especially relevant because we are interested in multiple maturities of the CDX index and

their multiple tranches. Using simulation approach, we can price all these products together in one set of simulations, so computation is fairly fast.

Note that if we use the measure changing technique before we use Monte Carlo simulation, computation can be even simpler and faster. Recall that the Radon-Nikodym derivative of the physical measure ($\mathbb{P}$) with respect to the risk-neutral measure ($\tilde{\mathbb{P}}$) is

$$\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}} = e^{-\int_0^t r_s ds}\frac{\pi_0}{\pi_t},$$

which is equivalent to

$$e^{-\int_0^t r_s ds} = \frac{\pi_t}{\pi_0}\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}. \tag{A.35}$$

By plugging (A.35) into our four expectations, we obtain

$$\begin{aligned}
\mathrm{EDR}(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} n_t \frac{\pi_t}{\pi_0}\frac{dP}{dQ}\right] = E^P\left[e^{u\cdot r_t} n_t \frac{\pi_t}{\pi_0}\right] \\
\mathrm{ELR}(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} L_t \frac{\pi_t}{\pi_0}\frac{dP}{dQ}\right] = E^P\left[e^{u\cdot r_t} L_t \frac{\pi_t}{\pi_0}\right] \\
\mathrm{ETLR}_j(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} T^L_{j,t} \frac{\pi_t}{\pi_0}\frac{dP}{dQ}\right] = E^P\left[e^{u\cdot r_t} T^L_{j,t} \frac{\pi_t}{\pi_0}\right] \\
\mathrm{ETRR}_j(u,t,X_0) &= \tilde{E}\left[e^{u\cdot r_t} T^R_{j,t} \frac{\pi_t}{\pi_0}\frac{dP}{dQ}\right] = E^P\left[e^{u\cdot r_t} T^R_{j,t} \frac{\pi_t}{\pi_0}\right].
\end{aligned} \tag{A.36}$$

There are two advantages of using equations (A.36) instead of equations (A.34) in simulation. First, we do not need to derive the risk-neutral dynamics of the model. (Note that the model is specified under the physical measure.) Second, integral expression ($e^{-\int_0^t r_s ds}$) disappears in equations (A.36) because the Radon-Nikodym derivative absorbs it when the probability measure is changed. Therefore, we do not need to use numerical integration when we implement equations (A.36).

A2.4.3. Simulation

Based on the values of $A_{i,t}$'s for all $i = 1, \cdots, N$, equation (2.5), (2.7), and (2.9) enable us to compute n_t, L_t, $T^L_{j,t}$, and $T^R_{j,t}$. Thus, equation (A.36) suggests that if we simulate (1) the state-price density (π_t) and (2) the firm value of each firm ($A_{i,t}$), we are able to price the CDX index and its tranches.

First, we start with the state-price density. From equation (2.2), it follows that

$$\frac{\pi_{t+\Delta t}}{\pi_t} = \exp\left[\eta\Delta t - \beta b_\lambda \int_t^{t+\Delta t} \lambda_s ds - \beta b_\xi \int_t^{t+\Delta t} \xi_s ds \right. \\ \left. -\gamma \log\left(\frac{C_{t+\Delta t}}{C_t}\right) + b_\lambda(\lambda_{t+\Delta t} - \lambda_t) + b_\xi(\xi_{t+\Delta t} - \xi_t)\right]. \quad \text{(A.37)}$$

Using the approximations

$$\begin{aligned} \lambda_{t+\Delta t}\Delta t &\simeq \int_t^{t+\Delta t} \lambda_s ds \\ \xi_{t+\Delta t}\Delta t &\simeq \int_t^{t+\Delta t} \xi_s ds, \end{aligned}$$

we can show that

$$\frac{\pi_{t+\Delta t}}{\pi_t} \simeq \exp\left[\eta\Delta t - \beta b_\lambda \lambda_{t+\Delta t}\Delta t - \beta b_\xi \xi_{t+\Delta t}\Delta t \right. \\ \left. -\gamma \log\left(\frac{C_{t+\Delta t}}{C_t}\right) + b_\lambda(\lambda_{t+\Delta t} - \lambda_t) + b_\xi(\xi_{t+\Delta t} - \xi_t)\right]. \quad \text{(A.38)}$$

Since Ito's Lemma implies

$$d\log C_t = \left(\mu_c - \frac{1}{2}\sigma_c^2\right) dt + \sigma_c dB_{C,t} + Z_{C,t} N_{C,t},$$

we are able to simulate log consumption growth, $\log\left(\frac{C_{t+\Delta t}}{C_t}\right)$ by applying the Euler scheme to the above SDE. Once we compute the simulation paths for two state variables (λ_t, ξ_t) and log consumption growth, equation (A.38) delivers the simulation path for the state-price

density, π_t.

Now, we consider how we simulate the firm value of firm i. We note that

$$\begin{aligned}\frac{A_{i,t+\Delta t}}{A_{i,t}} &= \frac{D_{i,t+\Delta t}}{D_{i,t}}\frac{G_i(\lambda_{t+\Delta t},\xi_{t+\Delta t})}{G_i(\lambda_t,\xi_t)} \\ &= \exp\left[\log\left(\frac{D_{i,t+\Delta t}}{D_{i,t}}\right)\right]\frac{G_i(\lambda_{t+\Delta t},\xi_{t+\Delta t})}{G_i(\lambda_t,\xi_t)}. \end{aligned} \tag{A.39}$$

We can simulate the log payout growth, $\log\left(\frac{D_{i,t+\Delta t}}{D_{i,t}}\right)$ based on the following SDE:

$$d\log D_{i,t} = \left(\mu_i - \frac{1}{2}\phi_i^2 - \frac{1}{2}\sigma_i^2\right)dt + \phi_i\sigma_c dB_{C,t} + \sigma_i dB_t^i + \phi_i Z_{C,t}dN_{C,t} + Z_t^i dN_t^i.$$

Since we can compute G_i using equation (2.4), the simulation path of $A_{i,t}$is generated based on equation (A.39).

A2.5. Default-free zero-coupon bond price

Let $H_0(\lambda_t,\xi_t,s-t)$ denote the time-t price of the default-free zero-coupon bond maturing at time $s>t$. By the pricing relation,

$$H_0(\lambda_t,\xi_t,s-t) = E_t\left[\frac{\pi_s}{\pi_t}\right]. \tag{A.40}$$

By multiplying π_t on both sides of (A.40), we obtain a martingale:

$$\pi_t H_0(\lambda_t,\xi_t,s-t) = \underbrace{E_t\left[\pi_s\right]}_{\text{martingale}}.$$

We conjecture that

$$H_0(\lambda_t,\xi_t,\tau) = \exp\left(a_0(\tau) + b_{0\lambda}(\tau)\lambda_t + b_{0\xi}(\tau)\xi_t\right). \tag{A.41}$$

By Ito's Lemma,

$$\frac{dH_{0,t}}{H_{0,t^-}} = \Big\{ b_{0\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2}b_{0\lambda}(\tau)^2\sigma_\lambda^2\lambda_t + b_{0\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2}b_{0\xi}(\tau)^2\sigma_\xi^2\xi_t$$
$$- a_0'(\tau) - b_{0\lambda}'(\tau)\lambda_t - b_{0\xi}'(\tau)\xi_t \Big\} dt$$
$$+ b_{0\lambda}(\tau)\sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} + b_{0\xi}(\tau)\sigma_\xi\sqrt{\xi_t}dB_{\xi,t} + (e^{b_{0\lambda}(\tau)Z_{\lambda,t}} - 1)N_{\lambda,t}. \quad \text{(A.42)}$$

Furthermore, we also derive the stochastic differential equation for $\pi_t H_{0,t}$ by combining equation (A.42) and (A.33) using Ito's Lemma:

$$\frac{d(\pi_t H_{0,t})}{\pi_{t^-}H_{0,t^-}} = \Big\{ -\beta - \mu_C + \gamma\sigma_C^2 - \lambda_t E\left[e^{(1-\gamma)Z_{\lambda,t}} - 1\right] - \nu E\left[e^{b_\lambda Z_{\lambda,t}} - 1\right]$$
$$+ b_{0\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2}b_{0\lambda}(\tau)^2\sigma_\lambda^2\lambda_t$$
$$+ b_{0\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2}b_{0\xi}(\tau)^2\sigma_\xi^2\xi_t$$
$$- a_0'(\tau) - b_{0\lambda}'(\tau)\lambda_t - b_{0\xi}'(\tau)\xi_t$$
$$+ b_\lambda b_{0\lambda}(\tau)\sigma_\lambda^2\lambda_t + b_\xi b_{0\xi}(\tau)\sigma_\xi^2\xi_t \Big\} dt$$
$$- \gamma\sigma_C dB_{C,t} + (b_\lambda + b_{0\lambda}(\tau))\sigma_\lambda\sqrt{\lambda_t}dB_{\lambda,t} + (b_\xi + b_{0\xi}(\tau))\sigma_\xi\sqrt{\xi_t}dB_{\xi,t}$$
$$+ (e^{-\gamma Z_{C,t}} - 1)dN_{C,t} + (e^{(b_\lambda + b_{0\lambda}(\tau))Z_{\lambda,t}} - 1)dN_{\lambda,t}.$$

Since $\pi_t H_{0,t}$ is a martingale, the sum of the drift and the jump compensator of $\pi_t H_{0,t}$ equals zero. That is,

$$\begin{aligned}
0 = &-\beta - \mu_C + \gamma\sigma_C^2 - \lambda_t E\left[e^{(1-\gamma)Z_{C,t}} - 1\right] - \nu E\left[e^{b_\lambda Z_{\lambda,t}} - 1\right] \\
&+ b_{0\lambda}(\tau)\kappa_\lambda(\xi_t - \lambda_t) + \frac{1}{2}b_{0\lambda}(\tau)^2\sigma_\lambda^2\lambda_t \\
&+ b_{0\xi}(\tau)\kappa_\xi(\bar{\xi} - \xi_t) + \frac{1}{2}b_{0\xi}(\tau)^2\sigma_\xi^2\xi_t \\
&- a_0'(\tau) - b_{0\lambda}'(\tau)\lambda_t - b_{0\xi}'(\tau)\xi_t \\
&+ b_\lambda b_{0\lambda}(\tau)\sigma_\lambda^2\lambda_t + b_\xi b_{0\xi}(\tau)\sigma_\xi^2\xi_t \\
&+ \lambda_t E\left[e^{-\gamma Z_{C,t}} - 1\right] + \nu E\left[e^{(b_\lambda + b_{0\lambda}(\tau))Z_{\lambda,t}} - 1\right]. \quad \text{(A.43)}
\end{aligned}$$

By collecting terms of (A.43),

$$\begin{aligned}
0 = &\underbrace{\left[-\beta - \mu_C + \gamma\sigma_C^2 + b_{0\xi}(\tau)\kappa_\xi\bar{\xi} + \nu E\left[e^{(b_\lambda + b_{0\lambda}(\tau))Z_{\lambda,t}} - e^{b_\lambda Z_{\lambda,t}}\right] - a_0'(\tau)\right]}_{=0} \\
&+ \lambda_t \underbrace{\left[-b_{0\lambda}(\tau)\kappa_\lambda + \frac{1}{2}b_{0\lambda}(\tau)^2\sigma_\lambda^2 + b_\lambda b_{0\lambda}(\tau)\sigma_\lambda^2 + E\left[e^{-\gamma Z_{C,t}} - e^{(1-\gamma)Z_{C,t}}\right] - b_{0\lambda}'(\tau)\right]}_{=0} \\
&+ \xi_t \underbrace{\left[b_{0\lambda}(\tau)\kappa_\lambda - b_{0\xi}(\tau)\kappa_\xi + \frac{1}{2}b_{0\xi}(\tau)^2\sigma_\xi^2 + b_\xi b_{0\xi}(\tau)\sigma_\xi^2 - b_{0\xi}'(\tau)\right]}_{=0}.
\end{aligned}$$

These conditions provide a system of ODEs:

$$\begin{aligned}
a_0'(\tau) &= -\beta - \mu_C + \gamma\sigma_C^2 + b_{0\xi}(\tau)\kappa_\xi\bar{\xi} + \nu E\left[e^{(b_\lambda + b_{0\lambda}(\tau))Z_{\lambda,t}} - e^{b_\lambda Z_{\lambda,t}}\right] \\
b_{0\lambda}'(\tau) &= -b_{0\lambda}(\tau)\kappa_\lambda + \frac{1}{2}b_{0\lambda}(\tau)^2\sigma_\lambda^2 + b_\lambda b_{0\lambda}(\tau)\sigma_\lambda^2 + E\left[e^{-\gamma Z_{C,t}} - e^{(1-\gamma)Z_{C,t}}\right] \\
b_{0\xi}'(\tau) &= b_{0\lambda}(\tau)\kappa_\lambda - b_{0\xi}(\tau)\kappa_\xi + \frac{1}{2}b_{0\xi}(\tau)^2\sigma_\xi^2 + b_\xi b_{0\xi}(\tau)\sigma_\xi^2. \quad \text{(A.44)}
\end{aligned}$$

This shows that H_0 satisfies the conjecture (A.41). We can obtain the boundary conditions for the system of ODEs (A.44) because

$$H_0(\lambda_t, \xi_t, 0) = 1,$$

which is equivalent to

$$a_0(0) = b_{0\lambda}(0) = b_{0\xi}(0) = 0.$$

A3. Appendix for Chapter 3

A3.1. EJ-OU process

In the model, I assume that X_t^i follows an EJ-OU process. Note that this process is in the affine jump-diffusion class of Duffie, Pan, and Singleton (2000). This provides some useful mathematical properties that are necessary for calculating important model quantities.

A3.1.1. Default probability

The following result is used for calculating the expression for the idiosyncratic risk component:

$$E_t\left[e^{-\int_t^{t+T} X_s^i ds}\right] \;=\; e^{\left\{A\left(T;\Theta^i\right)+B\left(T;\Theta^i\right)X_t^i\right\}},$$

where

$$\begin{aligned} B\left(T;\Theta^i\right) &= -\frac{1}{\kappa^i}\left(1-e^{-\kappa^i T}\right) \\ A\left(T;\Theta^i\right) &= \frac{{\sigma^i}^2}{4{\kappa^i}^3}\left(1-e^{-2\kappa^i T}\right)+\left[\frac{\bar{X}^i}{\kappa^i}-\frac{{\sigma^i}^2}{{\kappa^i}^3}\right]\left(1-e^{-\kappa^i T}\right) \\ &\quad +\left[\frac{{\sigma^i}^2}{2{\kappa^i}^2}-\bar{X}^i-\ell^i\right]T+\frac{\ell^i}{\kappa^i+\nu^i}\log\left|\left(1+\frac{\nu^i}{\kappa^i}\right)e^{\kappa^i T}-\frac{\nu^i}{\kappa^i}\right|. \end{aligned}$$

A3.1.2. Characteristic function

According to Das and Foresi (1996), the conditional characteristic function of $X_{t+\Delta t}^i$ given X_t^i (namely, $\varphi(\omega;\Delta t, X_t^i)$) is derived as an exponentially affine function of X_t^i:

$$\varphi(\omega;\Delta t, X_t^i) = E\left[\exp\left(i\omega X_{t+\Delta t}^i\right)|X_t^i\right] = e^{\left\{\tilde{A}\left(\omega;\Delta t,\Theta^i\right)+\tilde{B}\left(\omega;\Delta t,\Theta^i\right)X_t^i\right\}}$$

where

$$\begin{aligned}
\tilde{B}\left(\omega;\Delta t,\Theta^i\right) &= i\omega e^{-\kappa^i\Delta t} \\
\tilde{A}\left(\omega;\Delta t,\Theta^i\right) &= i\omega\bar{X}^i\left(1-e^{-\kappa^i\Delta t}\right)-\omega^2{\sigma^i}^2\left(\frac{1-e^{-2\kappa^i\Delta t}}{4\kappa^i}\right) \\
&\quad -\frac{i\ell^i}{\kappa}\left[\arctan\left(\omega\nu^i e^{-\kappa^i\Delta t}\right)-\arctan\left(\omega\nu^i\right)\right] \\
&\quad +\frac{\ell^i}{2\kappa}\log\left(\frac{1+\omega^2{\nu^i}^2e^{-2\kappa^i\Delta t}}{1+\omega^2{\nu^i}^2}\right).
\end{aligned}$$

A3.1.3. Transition density - exact

To construct the likelihood function in the estimation stage, the transition density of X_t^i is required. Unfortunately, the transition density of an EJ-OU process does not allow a closed-form expression due to the jump term. However, this can be calculated semi-analytically using the transform approach (see Singleton (2001)). Note that the density function is obtained by taking the inverse Fourier transform of the characteristic function derived in Appendix A3.1.2:

$$P(X_{t+\Delta t}^i|X_t^i) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\omega X_{t+\Delta t}}\varphi(\omega;\Delta t,X_t^i)d\omega.$$

Specifically, since the transition density is real-valued, the above integral can be re-written as

$$P(X_{t+\Delta t}|N_{t+\Delta t},S_t,X_t^c,N_t)=\frac{1}{\pi}\int_0^{\infty}Re\left[e^{-i\omega X_{t+\Delta t}}\phi(\omega;X_t^c)\right]d\omega$$

where $Re[\cdot]$ denotes the real part of a complex number. Thus, the derivation of the conditional transition density of $X_{t+\Delta t}$ reduces to computing a 1-dimensional integral.

A3.1.4. Transition density - approximate

The transformation approach in A3.1.3 provides exact expressions for the transition density and, subsequently, the likelihood function. However, this procedure puts tremendous

computational burden on the estimation due to the large number of firms and daily time series. Specifically, it requires computing (2574×215) number of numerical integrals on each estimation iteration. Thus, instead of this transformation approach, I approximate the transition of X_t^i in discrete time with Δt[5]:

$$X_{t+\Delta t}^i = X_t^i + \kappa^i(\bar{X}^i - X_t^i)\Delta t + \sigma^i\sqrt{\Delta t}\ \epsilon_{i,t} + Z_{t+\Delta t}(N_{t+\Delta t}^i - N_t^i)$$

where ϵ_t^i follows an iid standard normal distribution. Since I use daily data (which implies a very small Δt), it is reasonable to assume that there is at most one jump within each Δt period. This assumption reduces the jump component to a Bernoulli random variable with probability $\ell^i\Delta t$. That is,

$$X_{t+\Delta t}^i = X_t^i + \kappa^i(\bar{X}^i - X_t^i)\Delta t + \sigma^i\sqrt{\Delta t}\ \epsilon_{i,t} + Z_{t+\Delta t}1_{t+\Delta t}$$

where

$$1_{t+\Delta t} \sim \text{Bernoulli}(\ell^i\Delta t).$$

Conditional on no jump, the transition density of X_t^i follows a normal distribution with mean $\left(X_t^i + \kappa^i(\bar{X}^i - X_t^i)\Delta t\right)$ and variance $\left({\sigma^i}^2\Delta t\right)$. This implies

$$P(X_{t+\Delta t}^i | X_t^i, 1_{t+\Delta t} = 0) = \frac{1}{\sigma^i\sqrt{2\pi\Delta t}} \exp\left\{-\frac{(X_{t+\Delta t}^i - X_t^i - \kappa^i(\bar{X}^i - X_t^i)\Delta t))}{2{\sigma^i}^2\Delta t}\right\}.$$

On the other hand, conditional on a jump, the transition density follows the sum of a normal and exponential distribution. The density of the sum of two independent random variables can be derived from the convolution of their densities. The convolution between a normal and exponential distribution is known as an exponentially modified Gaussian distribution.

[5]The comparison with the result from the transformation approach shows that this approximation is extremely accurate.

Thus, I can show that

$$
\begin{aligned}
&P(X^i_{t+\Delta t}|X^i_t, 1_{t+\Delta t}=1) \\
&\qquad = \frac{1}{2\nu^i}\exp\left\{\frac{1}{2\nu^i}\left(X^i_t+\kappa^i(\bar{X}^i-X^i_t)\Delta t+\frac{{\sigma^i}^2\Delta t}{\nu^i}-2X^i_{t+\Delta t}\right)\right\} \\
&\qquad\qquad \times \operatorname{erfc}\left(\frac{X^i_t+\kappa^i(\bar{X}^i-X^i_t)\Delta t+\frac{{\sigma^i}^2\Delta t}{\nu^i}-X^i_{t+\Delta t}}{\sigma^i\sqrt{2\Delta t}}\right)
\end{aligned}
$$

where

$$\operatorname{erfc}(x)=\frac{2}{\pi}\int_x^\infty e^{-u^2}du.$$

Since the probability of a jump within this time period is $\ell^i\Delta t$, the transition probability of $X^i_{t+\Delta t}$ is expressed as

$$P(X^i_{t+\Delta t}|X^i_t)=\left(1-\ell^i\Delta t\right)P(X^i_{t+\Delta t}|X^i_t,1_{t+\Delta t}=0)+\left(\ell^i\Delta t\right)P(X^i_{t+\Delta t}|X^i_t,1_{t+\Delta t}=1).$$

A3.2. Occupation time $U_{t,t+T}$

A3.2.1. Conditional moment generating function

Darroch and Morris (1968) derive the conditional moment generating function of occupation time $U_{t,t+T}$. They show that

$$
\begin{aligned}
M_T(c|\mathcal{S}_t=0) &= E_1^\top e^{T(G+cD)}(E_1+E_2) \\
M_T(c|\mathcal{S}_t=1) &= E_2^\top e^{T(G+cD)}(E_1+E_2),
\end{aligned}
$$

where

$$D=\begin{bmatrix}0&0\\0&1\end{bmatrix},\quad E_1=\begin{bmatrix}1\\0\end{bmatrix},\quad \text{and}\quad E_2=\begin{bmatrix}0\\1\end{bmatrix}.$$

Since $T(G+cD)$ is a two-dimensional square matrix, the expression $\left(e^{T(G+cD)}\right)$ represents the matrix exponential. Note that this matrix exponential is different from the matrix

obtained by exponentiating every entry.[6]

A3.2.2. Probability density function

In a two-regime Markov chain model, Pedler (1971) explicitly derives the density function of the occupation time in a certain state . This function is expressed in terms of the Dirac delta function and the modified Bessel function. Applying these results, the density of $U_{t,t+T}$ is derived as

$$\begin{aligned} f_{U_{t,t+T}}(u;\pi_t) &= e^{-\phi_1 u-\phi_0(T-u)}\Big\{\pi_t\delta(T-u)+(1-\pi_t)\delta(u) \\ &\quad +(\pi_t v_1+(1-\pi_t)v_2)I_1(2v_0) \\ &\quad +(\pi_t\phi_1+(1-\pi_t)\phi_0)I_0(2v_0)\Big\} \end{aligned}$$

where

$$v_0=\sqrt{\phi_0\phi_1 u(T-u)},\quad v_1=\sqrt{\frac{\phi_0\phi_1 u}{T-u}},\quad v_2=\sqrt{\frac{\phi_0\phi_1(T-u)}{u}}.$$

Here, $\delta(\cdot)$ represents the Dirac delta function and I_r is the modified Bessel function of the first kind with order r. Namely,

$$I_r(z)=\sum_{j=0}^{\infty}\frac{\left(\frac{z}{2}\right)^{2j+r}}{j!(j+r)!}.$$

[6]Let M be an $n\times n$ matrix. Then, the exponential of M is defined as

$$e^M=\sum_{k=0}^{\infty}\frac{1}{k!}M^K.$$

If M is diagonal, e^M can be calculated by exponentiating every entry on the main diagonal. If M is not diagonal but diagonalizable (that is, $M=P^{-1}DP$ for some diagonal matrix D), then e^M is equal to $P^{-1}e^DP$. Although $T(G+cD)$ is not a diagonal matrix, it is diagonalizable (as long as c is not zero), so one can compute $e^{T(G+cD)}$ by diagonalizing $T(G+cD)$ and applying the above formula.

A3.3. Extraction of two beliefs and its uniqueness

For notational convenience, I define

$$M_{T,L,s} = M_T\left(-\lambda_{Lf}^{\text{rep}}|\mathcal{S}_t = s\right) \quad \text{and} \quad M_{T,H,s} = M_T\left(-\lambda_{Hf}^{\text{rep}}|\mathcal{S}_t = s\right).$$

Using this notation, it follows that

$$\begin{aligned} F^{\text{rep}}(\pi_t, \xi_t; T) &= M_{T,L,0} \\ &+ \left(M_{T,L,1} - M_{T,L,0}\right)\pi_t + \left(M_{T,H,0} - M_{T,L,0}\right)\xi_t \\ &+ \left(M_{T,H,1} + M_{T,L,0} - M_{T,L,1} - M_{T,H,0}\right)\pi_t\xi_t \end{aligned}$$

Therefore, I can re-write the system of equations (3.9) as two quadratic equations:

$$\begin{aligned} a_5\pi_t + b_5\xi_t + \pi_t\xi_t + c_5 &= 0 \\ a_{10}\pi_t + b_{10}\xi_t + \pi_t\xi_t + c_{10} &= 0 \end{aligned}$$

where

$$\begin{aligned} a_T &= \frac{M_{T,L,1} - M_{T,L,0}}{M_{T,H,1} + M_{T,L,0} - M_{T,L,1} - M_{T,H,0}}, \quad b_T = \frac{M_{T,H,0} - M_{T,L,0}}{M_{T,H,1} + M_{T,L,0} - M_{T,L,1} - M_{T,H,0}} \\ c_T &= \frac{M_{T,L,0} - S_{t,t+T}^{\text{rep}} e^{\lambda_n^{\text{rep}} T}}{M_{T,H,1} + M_{T,L,0} - M_{T,L,1} - M_{T,H,0}}. \end{aligned}$$

This system of two quadratic equations in π_t and ξ_t has two mathematical solutions. Specifically, one can show that

$$\pi_t = \frac{-m \pm \sqrt{m^2 - 4nl}}{2l} \tag{A.45}$$

where

$$l = -\frac{a_5 - a_{10}}{b_5 - b_{10}}, \quad m = a_5 - \frac{b_5(a_5 - a_{10})}{b_5 - b_{10}} - \frac{c_5 - c_{10}}{b_5 - b_{10}}, \quad n = c_5 - \frac{b_5(c_5 - c_{10})}{b_5 - b_{10}}.$$

The corresponding value of ξ_t is determined as

$$\xi_t \quad = \quad -\frac{c_5 - c_{10}}{b_5 - b_{10}} - \frac{a_5 - a_{10}}{b_5 - b_{10}}\pi_t.$$

Although mathematically the system of equations generates two solutions, only one solution is economically valid if a limit case of the model is considered. For example, think of the case where $\phi_1 \to 0$, $\lambda_n^{\text{rep}} \to 0$, and $\lambda_{Lf}^{\text{rep}} \to 0$. This economy has two interesting features: (1) once the economy reaches the frailty state, it is impossible to leave it, and (2) the default risk arises only when the economy is in the frailty state and the true severity is high. Thus, if investors believe with certainty that the economy is in the severe frailty state (i.e. $\pi_t = \xi_t = 1$), the economy reduces to a single regime with a constant intensity $\lambda_{Hf}^{\text{rep}}$. This intuition is confirmed when the model survival probability functions are investigated. Note that equation (3.8) implies that

$$S_{t,t+T}^{\text{rep}} = e^{-\lambda_{Hf}^{\text{rep}} T},$$

which suggests that the economy is equivalent to one with constant default intensity $\lambda_{Hf}^{\text{rep}}$.

Under this setup, I check the validity of two mathematical solutions of the extraction equations (3.9). By taking the limits of $\phi_1 \to 0$ and $\lambda_{Lf}^{\text{rep}} \to 0$, I can show that

$$\begin{aligned}
M_{T,H,0} &\to e^{-\phi_0 T} + \frac{\phi_0}{\phi_0 - \lambda_f^{\text{rep}}}\left(e^{-\lambda_f^{\text{rep}}} - e^{-\phi_0 T}\right) \\
M_{T,H,1} &\to e^{-\lambda_f^{\text{rep}}} \\
M_{T,L,0} &\to 0 \\
M_{T,L,0} &\to 0.
\end{aligned}$$

These results together with $\lambda_n^{\text{rep}} \to 0$ imply $l \to 0$, $m \to 1$, and $n \to -1$, which subsequently implies $\sqrt{m^2 - 4nl} \to m$. This suggests that the negative root of solution (A.45) is not valid because π_t diverges in the limit case considered. Note that the denominator of (A.45) approaches zero, but in case of negative root, the numerator approaches a non-zero constant (2). This makes

$$\pi_t \quad \to \quad \infty.$$

The same problem does not arise for the positive root because the numerator also approaches zero. In fact, one can show that with the positive root, both π_t and ξ_t correctly converge to 1. Thus, the extracted state variables are uniquely determined as

$$\begin{aligned} \pi_t &= \frac{-m + \sqrt{m^2 - 4nl}}{2l} \\ \xi_t &= -\frac{c_5 - c_{10}}{b_5 - b_{10}} - \frac{a_5 - a_{10}}{b_5 - b_{10}}\pi_t. \end{aligned}$$

A3.4. Lindeberg's condition

Consider a sequence of random variables $\{W_1, W_2, \cdots, W_I\}$ with finite means ($E[W_i] = \mu_i$) and finite variances ($Var\,(W_i) = \sigma_i^2$). The classical Central limit theorem (CLT) requires this sequence of random variables to be independent and identically distributed. However, the result of the CLT still holds even when the "identical" condition is missing as long as some regularity conditions are met. One such condition is Lindeberg's condition. If it is true that

$$\lim_{I \to \infty} \frac{1}{s_I^2} \sum_{i=1}^{I} E\left[(W_i - \mu_i)^2 1_{\{|W_i - \mu_i| > \epsilon s_I\}}\right] = 0, \quad \text{for every } \epsilon,$$

where

$$s_I^2 = \sum_{i=1}^{I} \sigma_i^2, \quad E[W_i] = \mu_i, \quad Var(W_i) = \sigma_i^2,$$

then, the sum of the random sequence $(\sum_{i=1}^{I} W_i)$ converges to a normal distribution. Namely,

$$\frac{1}{s_I}\sum_{i=1}^{I}(W_i - \mu_i) \xrightarrow{d} N(0,1).$$

Fortunately, the setup in section 3.3.2.1 satisfies Lindeberg's condition. Recall that conditional on the true severity $\left(\lambda_f^i\right)$ and the occupation time $(U_{t,t+T})$, default processes become independent Bernoulli random variables with

$$E[N_{t+T}^i] = P_D^i \quad \text{and} \quad Var\left(N_{t+T}^i\right) = P_D^i(1 - P_D^i).$$

Since I focus on the pool of investment grade firms, and their default probabilities are not too small (zero) nor too large (one), it is somewhat obvious to assume that

$$\lim_{I\to\infty} s_I^2 = \lim_{n\to\infty}\sum_{i=1}^{I} P_D^i(1 - P_D^i) \to \infty.$$

Now, fix any arbitrary $\epsilon > 0$. Then, I can always find $I_0(\epsilon)$ such that

$$\epsilon s_I = \epsilon\sqrt{\sum_{i=1}^{I} P_D^i(1 - P_D^i)} > 1, \qquad \forall I \geq I_0(\epsilon).$$

Note that $|N_{t+T}^i - P_D^i|$ is always smaller than or equal to 1, so it follows that

$$E\left[(N_{t+T}^i - P_D^i)^2 1_{\left\{|N_{t+T}^i - P_D^i| > \epsilon s_I\right\}}\right] = 0, \qquad \forall I \geq I_0(\epsilon).$$

This implies that

$$\sum_{i=1}^{I} E\left[(N_{t+T}^i - P_D^i)^2 1_{\left\{|N_{t+T}^i - P_D^i| > \epsilon s_I\right\}}\right]$$

converges to a finite number as $I \to \infty$. Since s_I^2 diverges to infinity as $I \to \infty$, it can be

shown that Lindeberg's condition holds:

$$\lim_{I\to\infty} \frac{1}{s_I^2} \sum_{i=1}^{I} E\left[(N_{t+T}^i - P_D^i)^2 1_{\{|N_{t+T}^i - P_D^i| > \epsilon s_I\}}\right] = 0.$$

A3.5. CDS pricing

Given a term structure of survival probabilities, a CDS contract can be priced in a model-free way. That is, it is possible to show that the premium leg and the protection leg can be expressed as functions of the term structure of risk-neutral survival probabilities (see, e.g., Duffie and Singleton (2003)). To illustrate, consider a T-year maturity CDS contract with premium C at time 0. Let $t_1, ..., t_M$ denote the insurance premium payment dates and R the recovery rate of the contract. For notational convenience, I define $t_0 = 0$ and $\Delta_m = t_m - t_{m-1}$ as the premium payment interval. It follows that

$$\begin{aligned} \text{Premium Leg} &= C \sum_{m=1}^{M} \mathrm{D}(t_m)\Delta_m S_{0,t_m} - \int_{t_{m-1}}^{t_m} \mathrm{D}(t_m)(t - t_{m-1}) dS_{0,t} \\ \text{Protection Leg} &= -(1-R)\int_0^T D(t) dS_{0,t}, \end{aligned}$$

where $D(\cdot)$ is the risk-free discount function and $S_{0,t}$ is the time-0 risk-neutral survival probability up to time t. Since the (fair or market) CDS spread C^* is defined as the premium that equates these two legs, it follows that

$$C^* = \frac{-(1-R)\int_0^T D(t) dS_{0,t}}{\sum_{m=1}^{M} \mathrm{D}(t_m)\Delta_m S_{0,t_m} - \int_{t_{m-1}}^{t_m} \mathrm{D}(t_m)(t - t_{m-1}) dS_{0,t}}.$$

A3.6. Robustness checks for stock return predictability

There may be two potential concerns regarding the stock return predictability results. I check the robustness of the results by taking these concerns into account.

First, the regression coefficients may be subject to finite sample bias. To check this, I follow the approach by Stambaugh (1999). He shows that if the following predictive system is considered,

$$\begin{aligned} y_t &= \alpha + \beta x_{t-1} + u_t \\ x_t &= \theta + \rho x_{t-1} + v_t \end{aligned}$$

the finite sample bias can be calculated using the following formula:

$$E\left[\hat{\beta} - \beta\right] \simeq -\frac{\sigma_{uv}}{\sigma_v^2}\left(\frac{1+3\rho}{T}\right).$$

According to this formula, the finite sample bias is only 0.026 for the CAT measure. This value is quite small compared to the predictive coefficients, mitigating concerns for bias. The reasons are two-fold. First, the CAT measure is persistent, but it is not as persistent as other classic predictors such as the log P/D. Specifically, the monthly AR(1) coefficient of 10-year horizon CAT 15 is roughly 0.94. Second, the innovations to the CAT measure are not strongly correlated with the residuals of the predictive regression (the correlation between u_t and v_t is -0.69).

Another potential concern arises due to overlapping returns in the predictive regressions. To mitigate this concern, I calculate the VAR-implied R^2 as suggested by Hodrick (1992). Specifically, I impose an assumption that the 1-month non-overlapping stock return and other predictors follow a VAR(1) system and estimate the system using the data time series. Then, I derive the expression for the VAR-implied population R^2 of long-horizon univariate predictive regressions. Furthermore, I calculate the confidence intervals for finite sample R^2 by simulating the estimated system.

Despite the strict assumption that the 1-month stock return and other predictors follow a VAR(1) model, the CAT measure generates sizable R^2, as shown in Table 3.8.

BIBLIOGRAPHY

Aït-Sahalia, Yacine, Julio Cacho-Diaz, and Roger J. A. Laeven, 2013, Modeling financial contagion using mutually exciting jump processes, Working paper.

Aït-Sahalia, Yacine, and Robert L Kimmel, 2010, Estimating affine multifactor term structure models using closed-form likelihood expansions, *Journal of Financial Economics* 98, 113–144.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2013, The Risk Premia Embedded in Index Options, Working paper, Northwestern University.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259–299.

Azizpour, Shahriar, Kay Giesecke, and Gustavo Schwenkler, 2014, Exploring the sources of default clustering, Working paper.

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters Implied by Equity Index Options, *The Journal of Finance* 66, 1969–2012.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bansal, Ravi, and Ivan Shaliastovich, 2011, Learning and asset-price jumps, *Review of Financial Studies* 24, 2738–2780.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long-run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.

Barro, Robert J., and Tao Jin, 2011, On the size distribution of macroeconomic disasters, *Econometrica* 79, 1567–1589.

Barro, Robert J., and Jose F. Ursua, 2008, Macroeconomic crises since 1870, *Brookings Papers on Economic Activity* no. 1, 255–350.

Bates, David S., 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, 181–238.

Bates, David S., 2008, The market for crash risk, *Journal of Economic Dynamics and Control* 32, 2291–2321.

Beeler, Jason, and John Y. Campbell, 2012, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, *Critical Finance Review* 1, 141–182.

Benzoni, Luca, Pierre Collin-Dufresne, Robert Goldstein, and Jean Helwege, 2014, Modeling credit contagion via the updating of fragile beliefs, Working paper.

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

Black, Fischer, and John C Cox, 1976, Valuing corporate securities: Some effects of bond indenture provisions, *The Journal of Finance* 31, 351–367.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–654.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, and Viktor Todorov, 2011a, Estimation of Jump Tails, *Econometrica* 79, 1727–1783.

Bollerslev, Tim, and Viktor Todorov, 2011b, Tails, Fears, and Risk Premia, *The Journal of Finance* 66, 2165–2211.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model specification and risk premia: Evidence from futures options, *Journal of Finance* 62, 1453–1490.

Brunnermeier, Markus, Stefan Nagel, and Lasse H. Pedersen, 2008, Carry trades and currency crashes, *NBER Macroeconomics Annual* 23, 313–347.

Buraschi, Andrea, and Alexei Jiltsov, 2006, Model Uncertainty and Option Markets with Heterogeneous Beliefs, *The Journal of Finance* 61, 2841–2897.

Calvet, Laurent E., and Adlai J. Fisher, 2007, Multifrequency news and stock returns, *Journal of Financial Economics* 86, 178–212.

Campbell, John Y., 2003, Consumption-based asset pricing, in G. Constantinides, M. Harris, and R. Stulz, eds.: *Handbook of the Economics of Finance, vol. 1b* (Elsevier Science, North-Holland).

Campbell, John Y., 2008, Risk and return in stocks and bonds, Lecture 2, Princeton Lectures in Finance.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chacko, George, and Sanjiv Das, 2002, Pricing interest rate derivatives: a general approach, *Review of Financial Studies* 15, 195–241.

Chacko, George, and Luis Viceira, 2005, Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets, *Review of Financial Studies* 18, 1369–1402.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, 2013, Measuring the "Dark Matter" in Asset Pricing Models, Working paper.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2009, The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well, *Management Science* 55, 1914–1932.

Christoffersen, Peter, and Kris Jacobs, 2004, The importance of the loss function in option valuation, *Journal of Financial Economics* 72, 291–318.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Collin-Dufresne, Pierre, Robert S Goldstein, and Fan Yang, 2012, On the Relative Pricing of Long-Maturity Index Options and Collateralized Debt Obligations, *The Journal of Finance* 67, 1983–2014.

Coval, Joshua D, Jakub W Jurek, and Erik Stafford, 2009, Economic catastrophe bonds, *The American Economic Review* pp. 628–666.

Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Dai, Qiang, and Kenneth Singleton, 2002, Expectations puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–442.

Darroch, JN, and KW Morris, 1968, Passage-time generating functions for continuous-time finite Markov chains, *Journal of Applied Probability* pp. 414–426.

Das, Sanjiv R, Darrell Duffie, Nikunj Kapadia, and Leandro Saita, 2007, Common failings: How corporate defaults are correlated, *The Journal of Finance* 62, 93–117.

Das, Sanjiv Ranjan, and Silverio Foresi, 1996, Exact solutions for bond and option prices with systematic jump risk, *Review of derivatives research* 1, 7–24.

Davis, Mark, and Violet Lo, 2001, Infectious defaults, *Quantitative Finance* 1, 382–386.

Drechsler, Itamar, 2012, Uncertainty, Time-Varying Fear, and Asset Prices, forthcoming, *Journal of Finance*.

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

Driessen, Joost, 2005, Is default event risk priced in corporate bonds?, *Review of Financial Studies* 18, 165–195.

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

Duffee, Gregory R, 1999, Estimating the price of default risk, *Review of Financial Studies* 12, 197–226.

Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 369–443.

Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita, 2009, Frailty correlated default, *The Journal of Finance* 64, 2089–2123.

Duffie, Darrell, and Larry G Epstein, 1992, Asset pricing with stochastic differential utility, *Review of Financial Studies* 5, 411–436.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Duffie, D, and KJ Singleton, 2003, Credit Risk: Pricing, Management, and Measurement. Princeton Series in Finance, .

Duffie, Darrell, and Kenneth J Singleton, 1999, Modeling term structures of defaultable bonds, *Review of Financial studies* 12, 687–720.

Duffie, Darrell, and Costis Skiadas, 1994, Continuous-time asset pricing: A utility gradient approach, *Journal of Mathematical Economics* 23, 107–132.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley, 1998, Implied Volatility Functions: Empirical Tests, *The Journal of Finance* 53, 2059–2106.

Epstein, Larry, and Stan Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Eraker, Bjorn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.

Fama, Eugene F, and Kenneth R French, 1988, Dividend yields and expected stock returns, *Journal of financial economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–56.

Feldhütter, Peter, and Mads Stenbo Nielsen, 2012, Systematic and idiosyncratic default risk in synthetic credit markets, *Journal of Financial Econometrics* 10, 292–324.

Fisher, Mark, and Christian Gilles, 1999, Consumption and asset prices with homothetic recursive preferences, Working paper, 99-17 Federal Reserve Bank of Atlanta.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Gabaix, Xavier, 2008, Linearity-generating processes: A modelling tool yielding closed forms for asset prices, Working paper, New York University.

Gabaix, Xavier, 2012, An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

Gallant, A. Ronald, Chien-Te Hsu, and George Tauchen, 1999, Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance, *Review of Economics and Statistics* 81, 617–631.

Gao, George P., and Zhaogang Song, 2013, Rare disaster concerns everywhere, Working paper, Cornell University.

Giesecke, Kay, Francis A Longstaff, Stephen Schaefer, and Ilya Strebulaev, 2011, Corporate bond default risk: A 150-year perspective, *Journal of Financial Economics* 102, 233–250.

Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer, and Ilya A. Strebulaev, 2014, Macroeconomic effects of corporate default crisis: A long-term perspective, *Journal of Financial Economics* 111, 297–310.

Giglio, Stefano, 2014, Credit Default Swap Spreads and Systemic Financial Risk, Working paper, University of Chicago.

Gorton, Gary, 2014, Some Reflections on the Recent Financial Crisis, *Trade, Globalization and Development: Essays in Honour of Kalyan K. Sanyal* pp. 161–.

Gourio, François, 2012, Disaster Risk and Business Cycles, *American Economic Review* 102, 2734–2766.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Jarrow, Robert A, and Fan Yu, 2001, Counterparty risk and the pricing of defaultable securities, *the Journal of Finance* 56, 1765–1799.

Kelly, Bryan, and Hao Jiang, 2014, Tail Risk and Asset Prices, *Review of Financial Studies* pp. –.

Kelly, Bryan T., Hanno N. Lustig, and Stijn Van Nieuwerburgh, 2012, Too-Systemic-To-Fail: What Option Markets Imply About Sector-Wide Government Guarantees, NBER Working Paper # 17149.

Kelly, Bryan T., Lubos Pastor, and Pietro Veronesi, 2014, The Price of Political Uncertainty: Theory and Evidence from the Option Market, Working paper, University of Chicago.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, *Review of Financial Studies* 18, 131–164.

Longstaff, Francis A., and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics* 74, 401–421.

Longstaff, Francis A, and Arvind Rajan, 2008, An empirical analysis of the pricing of collateralized debt obligations, *The Journal of Finance* 63, 529–563.

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium puzzle, *Journal of Monetary Economics* 15, 145–161.

Mortensen, Allan, 2006, Semi-Analytical Valuation of Basket Credit Derivatives in Intensity-Based Models., *Journal of Derivatives* 13, 8–26.

Neuberger, Anthony, 2012, Realized Skewness, *Review of Financial Studies* 25, 3423–3455.

Nowotny, Michael, 2011, Disaster begets crisis: The role of contagion in financial markets, Working paper, Boston University.

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

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy* 111, 642–685.

Pedler, PJ, 1971, Occupation times for two state Markov chains, *Journal of Applied Probability* 8, 381–390.

Rietz, Thomas A., 1988, The equity risk premium: A solution, *Journal of Monetary Economics* 22, 117–131.

Rubinstein, Mark, 1994, Implied Binomial Trees, *The Journal of Finance* 49, 771–818.

Santa-Clara, Pedro, and Shu Yan, 2010, Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options, *Review of Economics and Statistics* 92, 435–451.

Schroder, Mark, and Costis Skiadas, 1999, Optimal consumption and portfolio selection with stochastic differential utility, *Journal of Economic Theory* 89, 68–126.

Shaliastovich, Ivan, 2009, Learning, Confidence and Option Prices, working paper, University of Pennsylvania.

Singleton, Kenneth J, 2001, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* 102, 111–141.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Tauchen, George, 2005, Stochastic volatility in general equilibrium, Working paper, Duke University.

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

Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *Quarterly Journal of Economics* 105, 29–42.

Weitzman, Martin L., 2007, Subjective expectations and asset-return puzzles, *American Economic Review* 97, 1102–1130.

Yan, Shu, 2011, Jump risk, stock returns, and slope of implied volatility smile, *Journal of Financial Economics* 99, 216–233.

Zhou, Guofu, and Yingzi Zhu, 2014, Macroeconomic Volatilities and Long-run Risks of Asset Prices, Working paper, Washington University and Tsinghua University.