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

Sang Byung Seo

The Wharton School, University of Pennsylvania

December 17, 2014

Abstract

I investigate whether the possibility of economic catastrophes, defined as massive correlated defaults, is an important risk factor for asset pricing. I develop a model where default correlations among multiple firms arise through regime and belief shifts. From an estimation using the daily time series of CDS curves for 215 firms, I construct a catastrophic tail risk measure. Using the rich information contained in the term structure and various tail extremities of 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 reveal that seemingly impossible catastrophes are a significant source of risk perceived by investors.

---

*Job market paper, email: sangseo@wharton.upenn.edu. I thank my advisors Nikolai Roussanov, Jessica Wachter, and Amir Yaron. I also thank Yacine Aït-Sahalia, Ian Appel, Rosie Bae, Christine Dobridge, Domenico Cuoco, Vincent Glode, Jakub W. Jurek, Mete Kilic, Michael J. Lee, Doron Levit, Karen Lewis, Krishna Ramaswamy, Krista Schwarz, Ivan Shaliastovich, Robert F. Stambaugh, Colin Ward, and seminar participants at Wharton for valuable comments and advice. All errors are my own.

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 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

[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.

[2] The structure of a CDS contract with its pricing is explained in Section 5.1.

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 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

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

[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).

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-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

[5] In Section 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 3.1.

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

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).

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 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.

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

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

in firms' default probabilities through firms' different sensitivities towards the frailty state: 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 ($\bar{X}^i = 0$).

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

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

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 (1).

[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 6.2, I indeed provide a real example of this case.

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 $P_S^i(t, t+T; \mathcal{G}_t)$ 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:

$$\begin{aligned} E\left[e^{-\int_t^{t+T}\lambda_s^i ds}\,\middle|\,\mathcal{G}_t\right] &= E\left[e^{-\int_t^{t+T}\lambda_n^i+\lambda_f^i 1_{\{\mathcal{S}_s=1\}}+X_s^i ds}\,\middle|\,\mathcal{G}_t\right] \\ &= \underbrace{\left[e^{-\lambda_n^i T}\right]}_{\text{normal default risk}} \underbrace{E\left[e^{-\int_t^{t+T}X_s^i ds}\,\middle|\,\mathcal{G}_t\right]}_{\text{firm-specific risk}} \underbrace{E\left[e^{-\lambda_f^i\int_t^{t+T}1_{\{\mathcal{S}_s=1\}}ds}\,\middle|\,\mathcal{G}_t\right]}_{\text{frailty risk}}. \quad (2) \end{aligned}$$

[13] In contrast, my main model in Section 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).

The last equality holds because the dynamics of idiosyncratic component X_t^i is independent of transitions between regimes. Equation (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 A.1 for details.

Though it appears trickier, the second expectation (frailty risk) of equation (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 B.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 (2)) is calculated as

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

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 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.

$$\begin{aligned} P_S^i(t, t+T) &\equiv E\left[e^{-\int_t^{t+T} \lambda_s^i ds} \,\middle|\, \mathcal{F}_t\right] \\ &= e^{\left\{-\lambda_n^i T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right) X_t^i\right\}} F^i(\pi_t, \xi_t; T), \end{aligned} \tag{3}$$

[14] The rationale for this extension is provided in Section 2.3 and 2.4.

where

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

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 (1) and (3), the default probability is obtained in closed-form as follows:

$$P^i_D(t, t+T) = 1 - e^{\left\{-\lambda^i_n T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right)X^i_t\right\}} F^i(\pi_t, \xi_t; T). \quad (4)$$

The term structure of default probabilities is computed by applying equation (4) with different values of maturity T.

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 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

[15] In Section 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 5.2 provides a detailed explanation about this average firm.

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 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.

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*.

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

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

[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.

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 structure.[22]

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]

[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 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

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.

3.1 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 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

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.

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

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 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.

3.2 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

$$\mathrm{CAT}_t(T;h) = P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i > h\right). \tag{5}$$

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 mea-

[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.

sure. Hence, the CAT measure represents risk-neutral probabilities.

Equation (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 (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 (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.

[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.

3.2.1 Joint default probabilities

Computing the CAT measure using equation (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.

By the law of total probability, it follows that

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

where $U_{t,t+T}$ represents the occupation time in the frailty state between time t and $t+T$ (Section 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 B.2. By applying the law of total probability one more time with respect to the true severity of the frailty state, the equations become

$$\begin{aligned} P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i > h\right) = \int_0^T \Bigg\{ & P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i > h \,\middle|\, \lambda_f^i = \lambda_{Lf}^i, U_{t,t+T} = u\right)(1-\xi_t) \\ & + P_t\left(\frac{1}{I}\sum_{i=1}^{I} N_{t+T}^i > h \,\middle|\, \lambda_f^i = \lambda_{Lf}^i, U_{t,t+T} = u\right)\xi_t \Bigg\} f_{U_{t,t+T}}(u)du. \quad (6) \end{aligned}$$

Now, it is important to observe that conditional on the true severity of the frailty state $\left(\lambda_f^i\right)$ 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_{t+T}^i can be viewed as

an independent Bernoulli random variable with event probability $P_D^i(t, t+T|\lambda_f^i, U_{t,t+T}) = 1 - e^{\left\{-\lambda_n^i T + A\left(T;\Theta^i\right) + B\left(T;\Theta^i\right) X_t^i - \lambda_f^i U_{t,t+T}\right\}}$. If firms are homogeneous, the sum of all these default processes ($\sum_{i=1}^{I} N_{t+T}^i$) 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 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 (6).

[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.

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

4 Data

As described in Section 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 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.

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

[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.

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 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 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 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

[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.

Kimmel (2010).) Specifically, I assume that the individual firms' 5-year survival probabilities are observed without error:

$$\underbrace{S^i_{t,t+5}}_{\text{Data}} \;=\; \underbrace{P^i_S(t,t+5)}_{\text{Model}}. \tag{7}$$

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.

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 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 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 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

[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 E 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.

[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.

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)), 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 2.

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 firm. That is,

$$P_S^{\text{rep}}(t, t+T) \quad = \quad e^{-\lambda_n^{\text{rep}} T} F^{\text{rep}}(\pi_t, \xi_t; T). \tag{8}$$

[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.

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

$$\begin{aligned} S^{\text{rep}}_{t,t+5} &= e^{-5\lambda^{\text{rep}}_n} F^{\text{rep}}(\pi_t, \xi_t; 5) \\ S^{\text{rep}}_{t,t+10} &= e^{-10\lambda^{\text{rep}}_n} F^{\text{rep}}(\pi_t, \xi_t; 10). \end{aligned} \tag{9}$$

In Appendix C, 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 C 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 $(\phi_0, \phi_1, \lambda^{\text{rep}}_{Lf}, \lambda^{\text{rep}}_{Hf}, \lambda^{\text{rep}}_n)$ by minimizing RMSE.[38] Specifically, I find

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

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^i_t'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

[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.

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) and (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

$$X_t^i = \frac{1}{B\left(5;\Theta^i\right)} \left[\log S_{t,t+5}^i + 5\lambda_n^i - A\left(5;\Theta^i\right) - \log\left(F^i(\pi_t, \xi_t; 5) \right) \right]. \tag{10}$$

Not only does equation (10) provide the values of the latent process X_t^i extracted from 5-year survival probabilities but it also facilitates in deriving the expression for the transition density of $S_{t+\Delta t}^i$. Denoting the vector of survival probabilities of maturities other than 5-year

[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.

as $S_t^{i,-5Y}$, 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 (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 A, I derive this transition density of X^i_t in exact (A.3) and approximate (A.4) forms.

6 Interpretation of estimation results

As described in Section 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 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 ($\pi_t(1 - \xi_t)$).

6.1 Extracted time series of investors' beliefs

As can be seen in Figure 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 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 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 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 5, the estimated model fits the data extremely well for all maturities.

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 6 explains the role of these two components using two firms as examples: ConocoPhillips

[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.

and Transocean.

The top left panel of Figure 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.

7 Implications for catastrophic tail risk

Equation (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:

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

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:

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

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 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 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.

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 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 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.

7.2 Term structure of catastrophic tail risk

Figure 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 6.1 and 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 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.

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.

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 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 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. 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.

[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.

Table 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 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, 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 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 5.

8.2 Return predictability

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]

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

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.

From the first panel of Table 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 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.

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

Furthermore, the CAT measure drives out these predictors when they are placed into the same regression. The second and third columns of Table 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 6, in the univariate predictive regression, the coefficient for the default spread was positive, and the coefficient for the log P/D was negative.

The last column of Table 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 F.

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 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.

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 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 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 9). Across all maturities of bonds, the CAT measure performs better than the term spread in terms of R^2.

Table 10 presents multivariate predictive regressions of 5-year bond returns ($n = 5$). The

[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 $h = 12$).

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.

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 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

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

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 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.

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.

Table 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

[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.

[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.

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).

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.

[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.

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.

Appendix

A 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.

A.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

$$\begin{aligned}
B\left(T;\Theta^i\right) &= -\frac{1}{\kappa^i}\left(1-e^{-\kappa^i T}\right) \\
A\left(T;\Theta^i\right) &= \frac{{\sigma^i}^2}{4{\kappa^i}^3}\left(1-e^{-2\kappa^i T}\right)+\left[\frac{\bar{X}^i}{\kappa^i}-\frac{{\sigma^i}^2}{{\kappa^i}^3}\right]\left(1-e^{-\kappa^i T}\right) \\
&\quad +\left[\frac{{\sigma^i}^2}{2{\kappa^i}^2}-\bar{X}^i-\ell^i\right]T+\frac{\ell^i}{\kappa^i+\nu^i}\log\left|\left(1+\frac{\nu^i}{\kappa^i}\right)e^{\kappa^i T}-\frac{\nu^i}{\kappa^i}\right|.
\end{aligned}$$

A.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

$$\begin{aligned}
\tilde{B}\left(\omega;\Delta t,\Theta^i\right) &= i\omega e^{-\kappa^i\Delta t} \\
\tilde{A}\left(\omega;\Delta t,\Theta^i\right) &= i\omega\bar{X}^i\left(1-e^{-\kappa^i\Delta t}\right)-\omega^2{\sigma^i}^2\left(\frac{1-e^{-2\kappa^i\Delta t}}{4\kappa^i}\right) \\
&\quad -\frac{i\ell^i}{\kappa}\left[\arctan\left(\omega\nu^i e^{-\kappa^i\Delta t}\right)-\arctan\left(\omega\nu^i\right)\right] \\
&\quad +\frac{\ell^i}{2\kappa}\log\left(\frac{1+\omega^2{\nu^i}^2e^{-2\kappa^i\Delta t}}{1+\omega^2{\nu^i}^2}\right).
\end{aligned}$$

A.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 A.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.

A.4 Transition density - approximate

The transformation approach in A.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[54]:

$$X_{t+\Delta t}^i = X_t^i + \kappa^i(\bar{X}^i - X_t^i)\Delta t + \sigma^i\sqrt{\Delta t}\ \epsilon_{i,t} + Z_{t+\Delta t}(N_{t+\Delta t}^i - N_t^i)$$

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

$$P(X_{t+\Delta t}^i | X_t^i, 1_{t+\Delta t} = 0) = \frac{1}{\sigma^i\sqrt{2\pi\Delta t}} \exp\left\{-\frac{(X_{t+\Delta t}^i - X_t^i - \kappa^i(\bar{X}^i - X_t^i)\Delta t))}{2{\sigma^i}^2\Delta t}\right\}.$$

[54]The comparison with the result from the transformation approach shows that this approximation is extremely accurate.

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. 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

$$P(X^i_{t+\Delta t}|X^i_t) = \left(1-\ell^i\Delta t\right)P(X^i_{t+\Delta t}|X^i_t, 1_{t+\Delta t}=0) + \left(\ell^i\Delta t\right)P(X^i_{t+\Delta t}|X^i_t, 1_{t+\Delta t}=1).$$

B Occupation time $U_{t,t+T}$

B.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.[55]

B.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

$$v_0 = \sqrt{\phi_0\phi_1 u(T-u)}, \quad v_1 = \sqrt{\frac{\phi_0\phi_1 u}{T-u}}, \quad v_2 = \sqrt{\frac{\phi_0\phi_1(T-u)}{u}}.$$

[55] 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 $P^{-1}e^D P$. 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.

Here, $\delta(\cdot)$ represents the Dirac delta function and I_r is the modified Bessel function of the first kind with order r. Namely,

$$I_r(z) = \sum_{j=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2j+r}}{j!(j+r)!}.$$

C 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 (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

$$\pi_t = \frac{-m \pm \sqrt{m^2 - 4nl}}{2l} \tag{C.1}$$

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

$$\xi_t = -\frac{c_5 - c_{10}}{b_5 - b_{10}} - \frac{a_5 - a_{10}}{b_5 - b_{10}}\pi_t.$$

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 (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 (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 (C.1) is not valid because π_t diverges in the limit case considered. Note that the denominator of (C.1) approaches zero, but in case of negative root, the numerator approaches a non-zero constant (2). This makes

$$\pi_t \to \infty.$$

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

$$\begin{aligned} \pi_t &= \frac{-m + \sqrt{m^2 - 4nl}}{2l} \\ \xi_t &= -\frac{c_5 - c_{10}}{b_5 - b_{10}} - \frac{a_5 - a_{10}}{b_5 - b_{10}} \pi_t. \end{aligned}$$

D 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 3.2.1 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 $|N^i_{t+T} - P^i_D|$ is always smaller than or equal to 1, so it follows that

$$E\left[(N^i_{t+T} - P^i_D)^2 1_{\left\{|N^i_{t+T} - P^i_D| > \epsilon s_I\right\}}\right] = 0, \qquad \forall I \geq I_0(\epsilon).$$

This implies that

$$\sum_{i=1}^{I} E\left[(N^i_{t+T} - P^i_D)^2 1_{\left\{|N^i_{t+T} - P^i_D| > \epsilon s_I\right\}}\right]$$

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:

$$\lim_{I\to\infty} \frac{1}{s_I^2} \sum_{i=1}^{I} E\left[(N^i_{t+T} - P^i_D)^2 1_{\left\{|N^i_{t+T} - P^i_D| > \epsilon s_I\right\}}\right] = 0.$$

E 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

$$C^* = \frac{-(1-R)\int_0^T D(t)dS_{0,t}}{\sum_{m=1}^{M} \mathrm{D}(t_m)\Delta_m S_{0,t_m} - \int_{t_{m-1}}^{t_m} \mathrm{D}(t_m)(t-t_{m-1})dS_{0,t}}.$$

F 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 8.

References

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.

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.

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.

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.

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.

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.

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.

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.

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, 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.

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, 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.

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.

Gabaix, Xavier, 2012, An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

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. –.

Longstaff, Francis A, and Arvind Rajan, 2008, An empirical analysis of the pricing of collateralized debt obligations, *The Journal of Finance* 63, 529–563.

Mortensen, Allan, 2006, Semi-Analytical Valuation of Basket Credit Derivatives in Intensity-Based Models., *Journal of Derivatives* 13, 8–26.

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.

Singleton, Kenneth J, 2001, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* 102, 111–141.

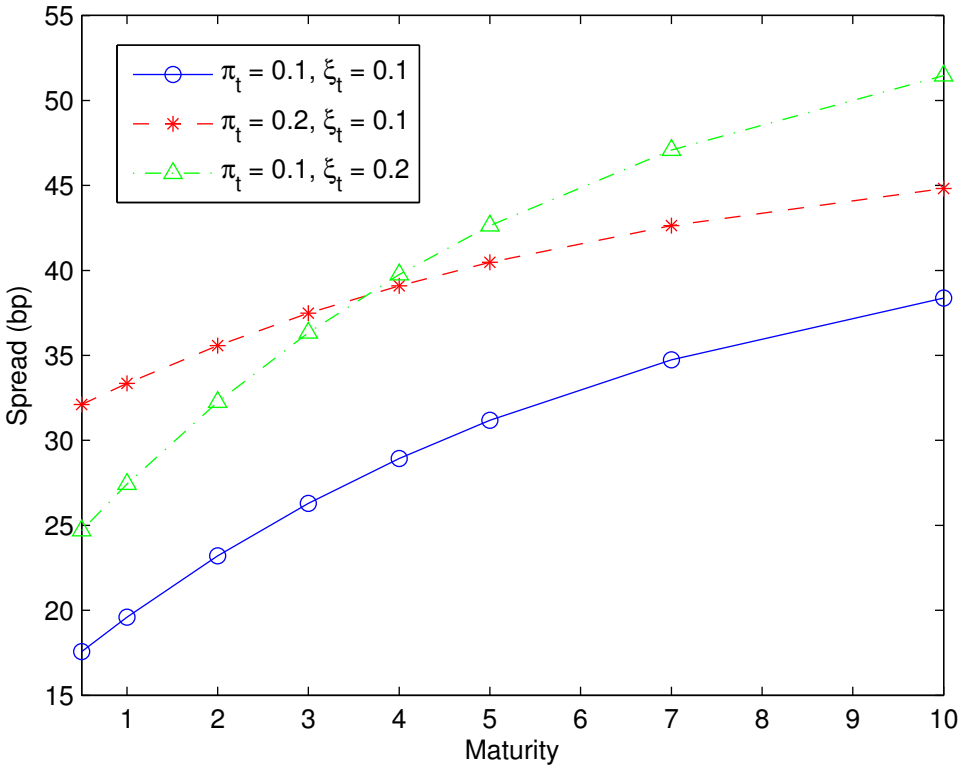
Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

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

Weitzman, Martin L., 2007, Subjective expectations and asset-return puzzles, *American Economic Review* 97, 1102–1130.

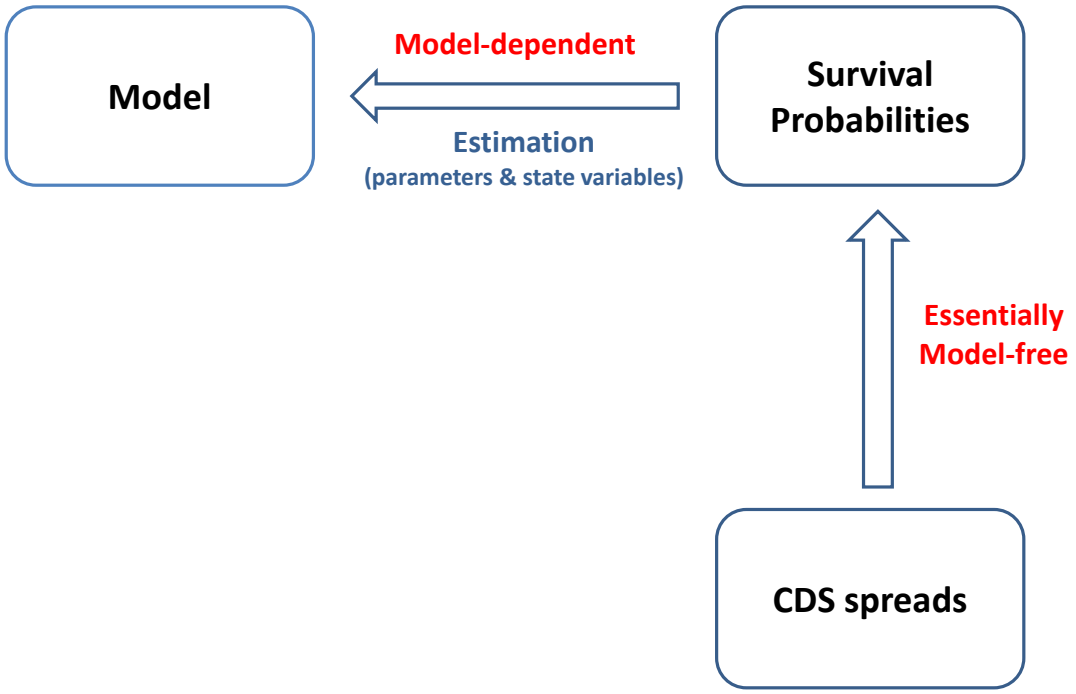
Yan, Shu, 2011, Jump risk, stock returns, and slope of implied volatility smile, *Journal of Financial Economics* 99, 216–233.

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



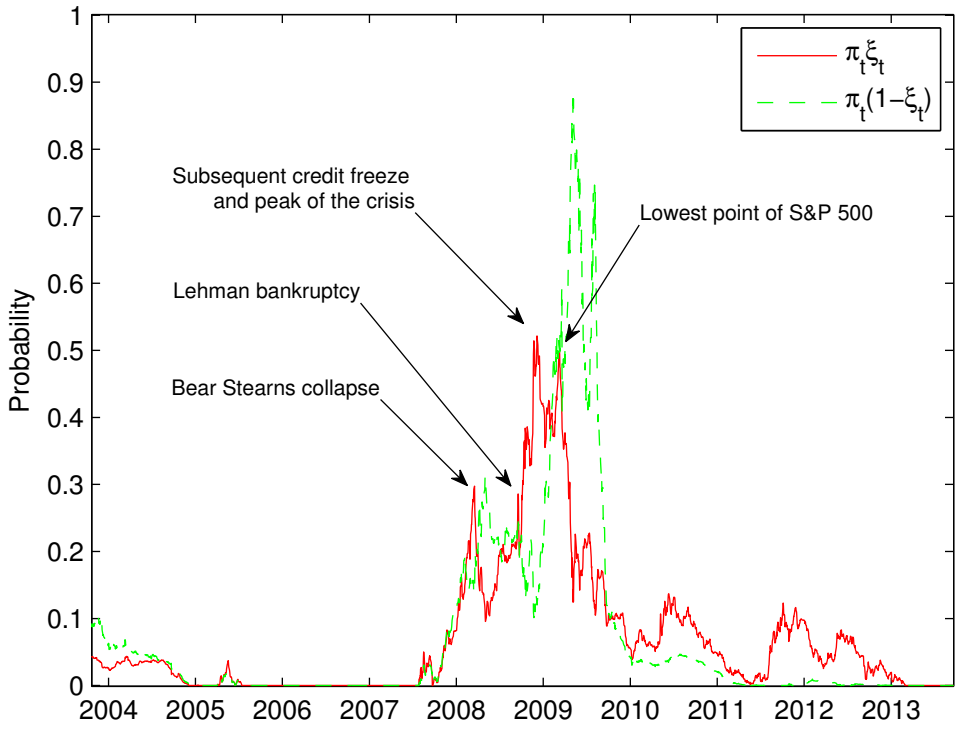
Note: This figure shows the comparative statistics of the model when the two beliefs change. The blue line with circle markers represents the model-implied CDS term structure when the two beliefs (π_t and ξ_t) are all 10%. The red line with asterisk markers fixes ξ_t but increases π_t to 20%. The green 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 2: Estimation of the model



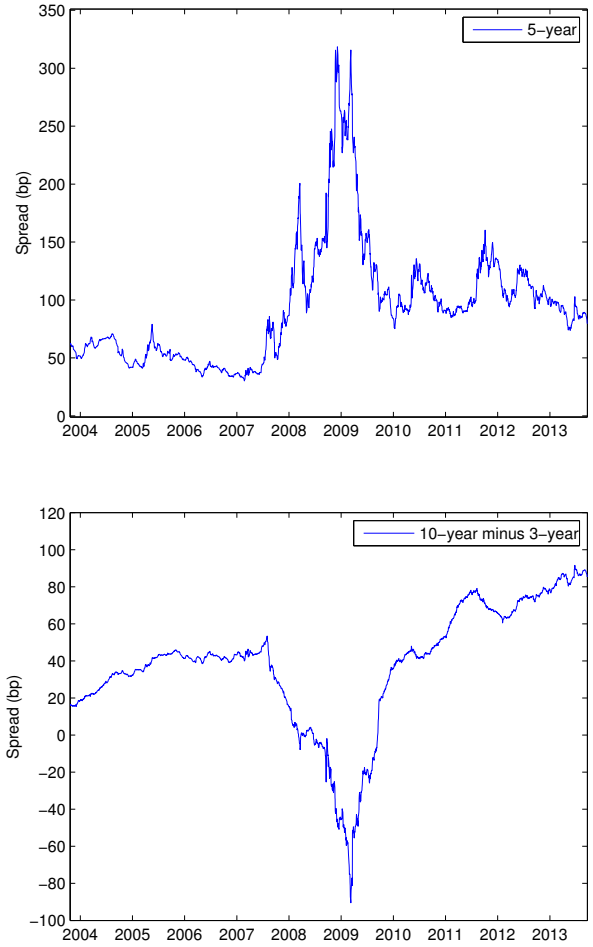
Note: The diagram describes the relationship between the model, survival probabilities, and CDS spreads. Pricing CDS is model-free once the term structure of survival probabilities is obtained (also the reverse direction is also essentially model-free). However, the term structure of survival probabilities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

Figure 3: Implied beliefs of investors



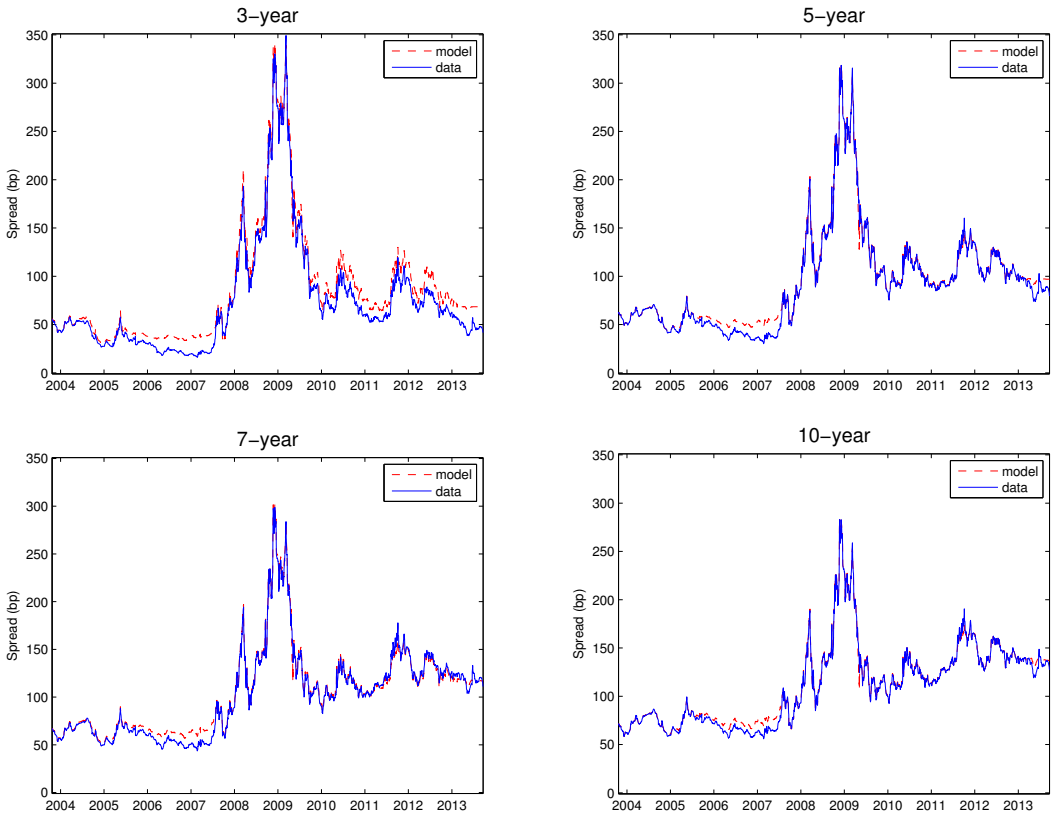
Note: This figure displays the extracted time series of investors' two beliefs. The red solid line represents the probability that the economy is in the severe frailty state ($\pi_t\xi_t$). The green dashed line represents the probability that the economy is in the moderate frailty state ($\pi_t(1-\xi_t)$).

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



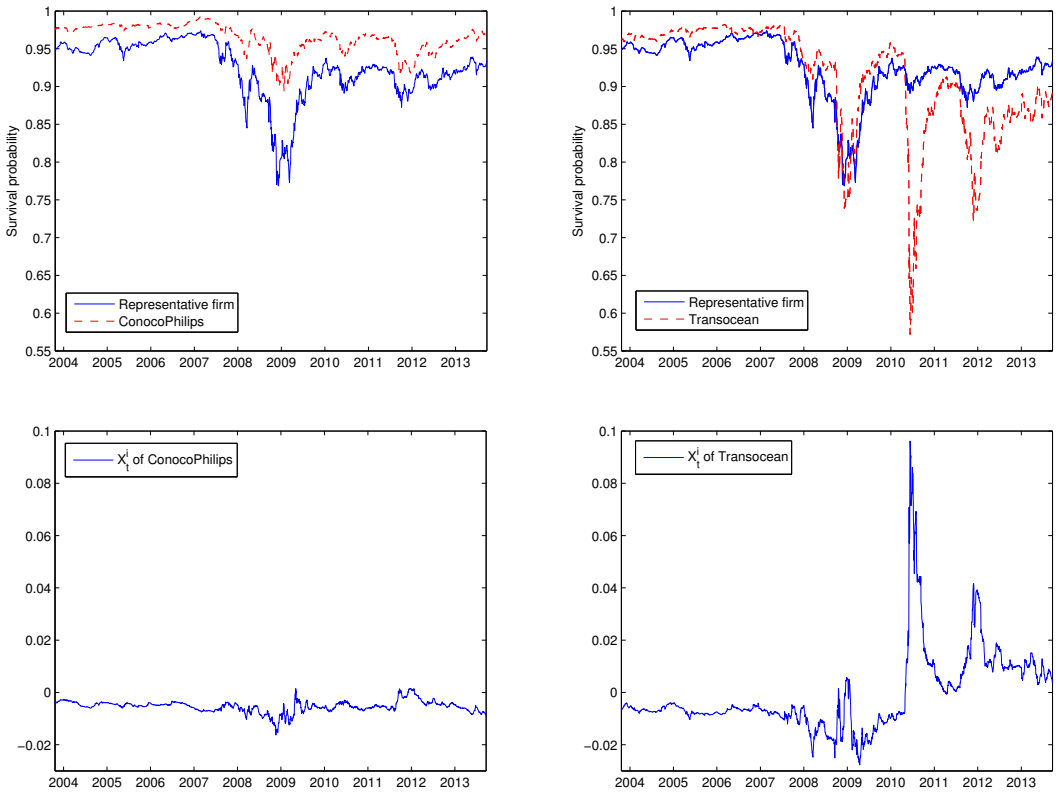
Note: Two panels present the implied CDS curves of the hypothetical firm representing the average firm in the portfolio. The top panel shows the time series of 5-year CDS spreads, which depict the level of the CDS curves. The bottom panel shows the time series of the difference between 10-year spreads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

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



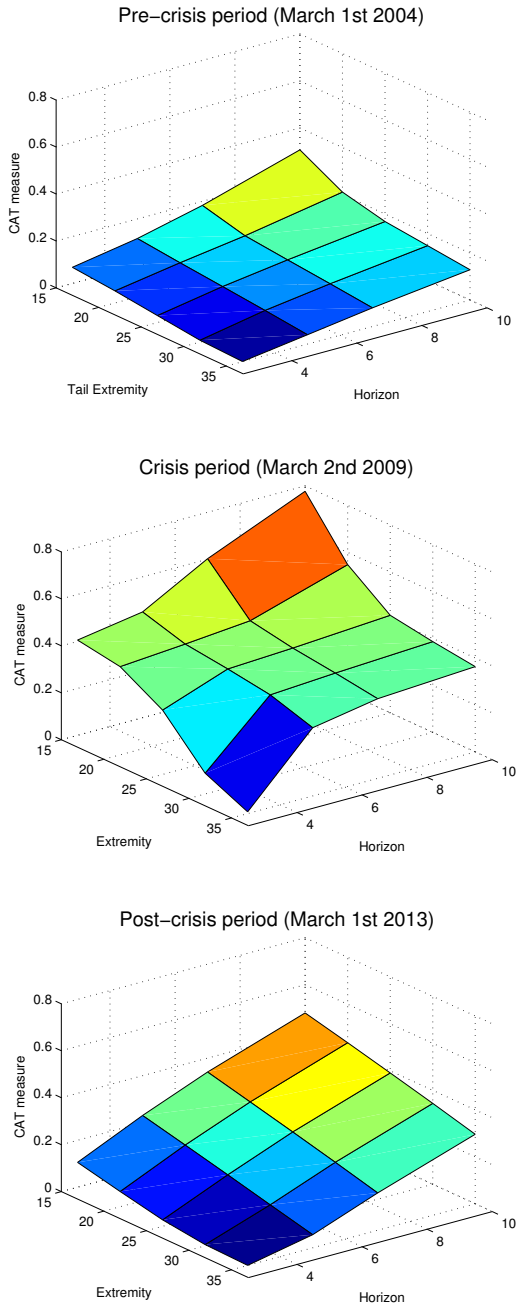
Note: The figure shows the CDS term structure of the average firm in the model and data. The red dashed lines represent the time series of the model implied CDS term structure (3-year, 5-year, 7-year, 10-year maturities). The blue lines represent their data counterparts.

Figure 6: Individual firm dynamics: two examples



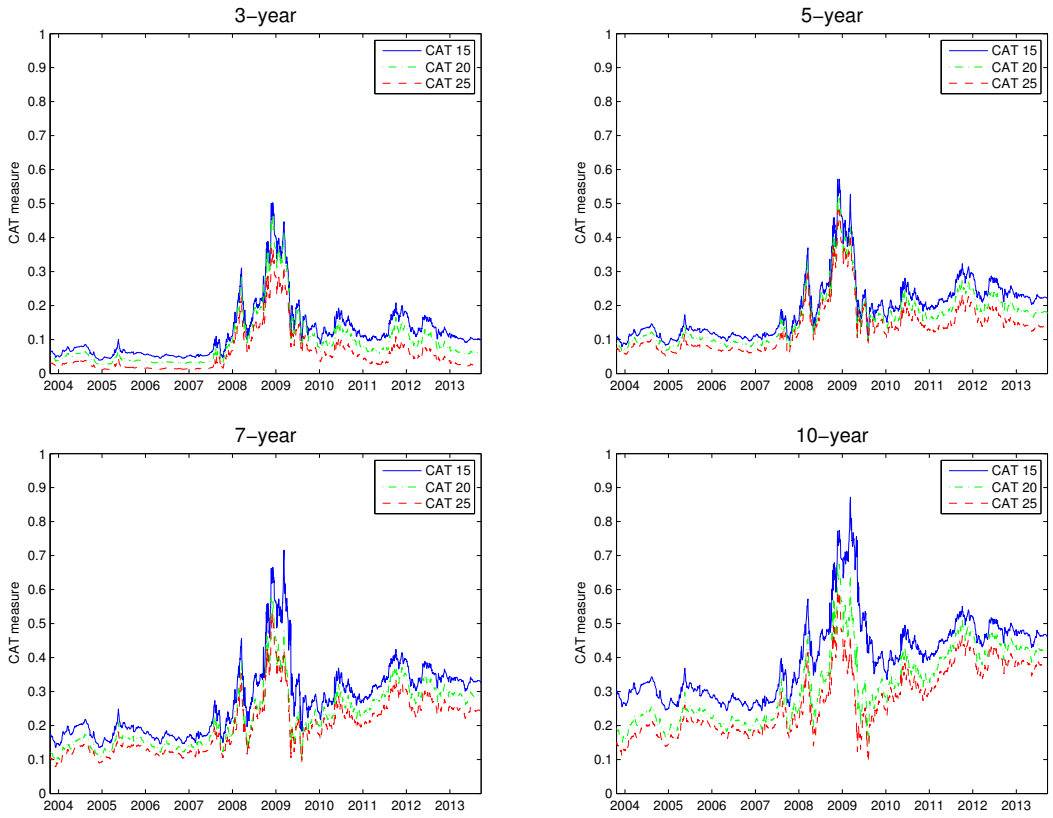
Note: This figure provides two examples of individual firm dynamics. The two left panels are about ConocoPhillips and the two right panels are about Transocean. The top two panels display each firm's time series of 5-year implied survival probabilities (red dashed line) together with the average firm's 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 7: CAT surfaces in the pre-crisis, crisis, and post-crisis period



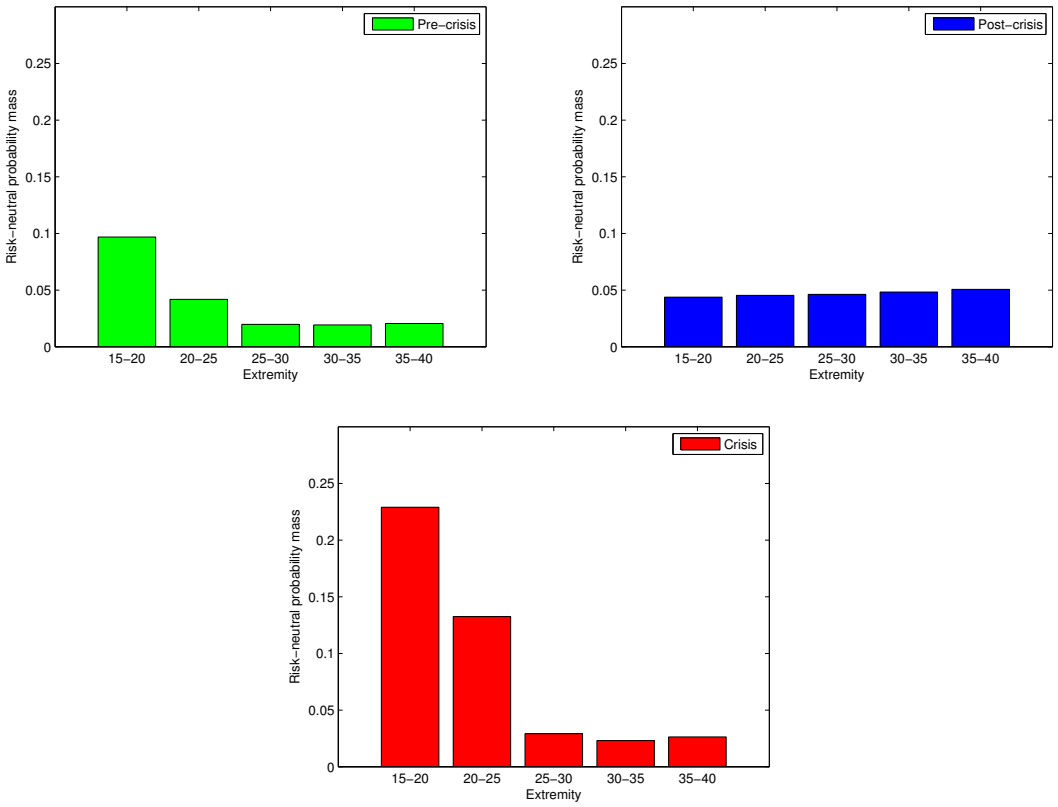
Note: The figure shows three CAT surfaces representing the three periods: the pre-crisis, crisis, and post-crisis periods. Each CAT surface has two dimensions - tail extremity and horizon.

Figure 8: Time series of CAT surfaces



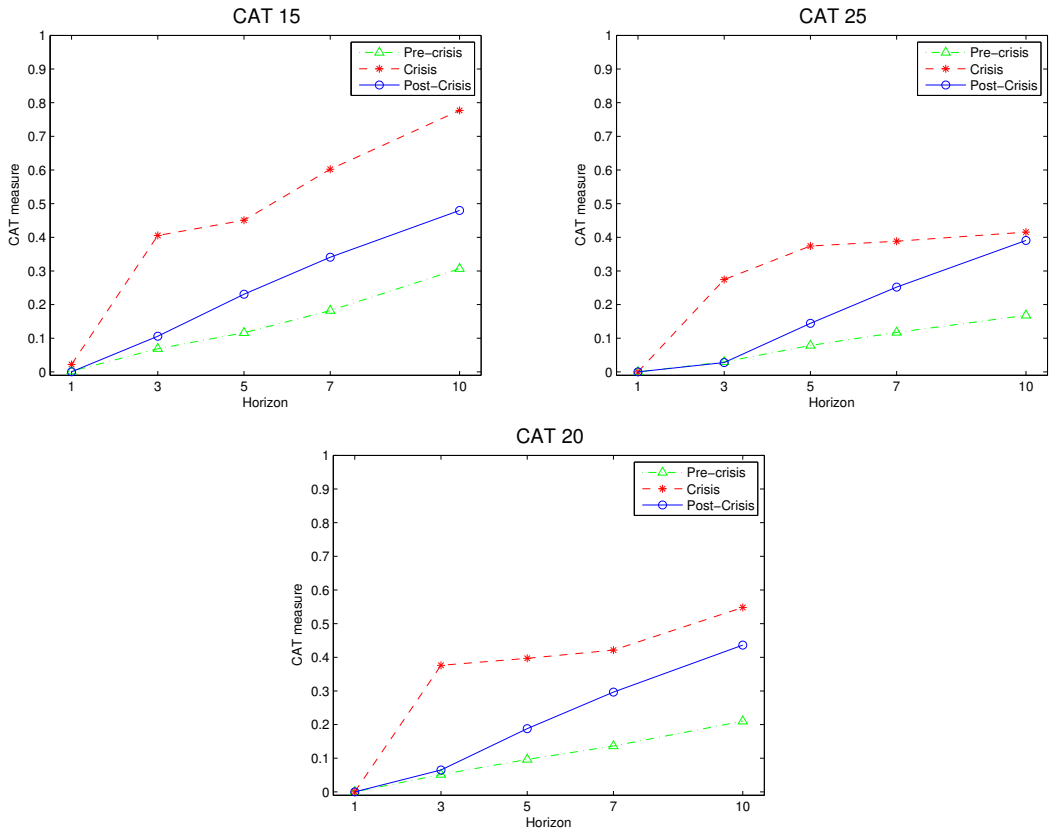
Note: The figure shows the time series of CAT surfaces. Tail extremities 15%, 20%, and 25% are considered at 3-year, 5-year, 7-year, and 10-year horizons. The blue solid lines represent CAT 15, the green dot-dashed lines represent CAT 20, and the red dashed lines represent CAT 25.

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



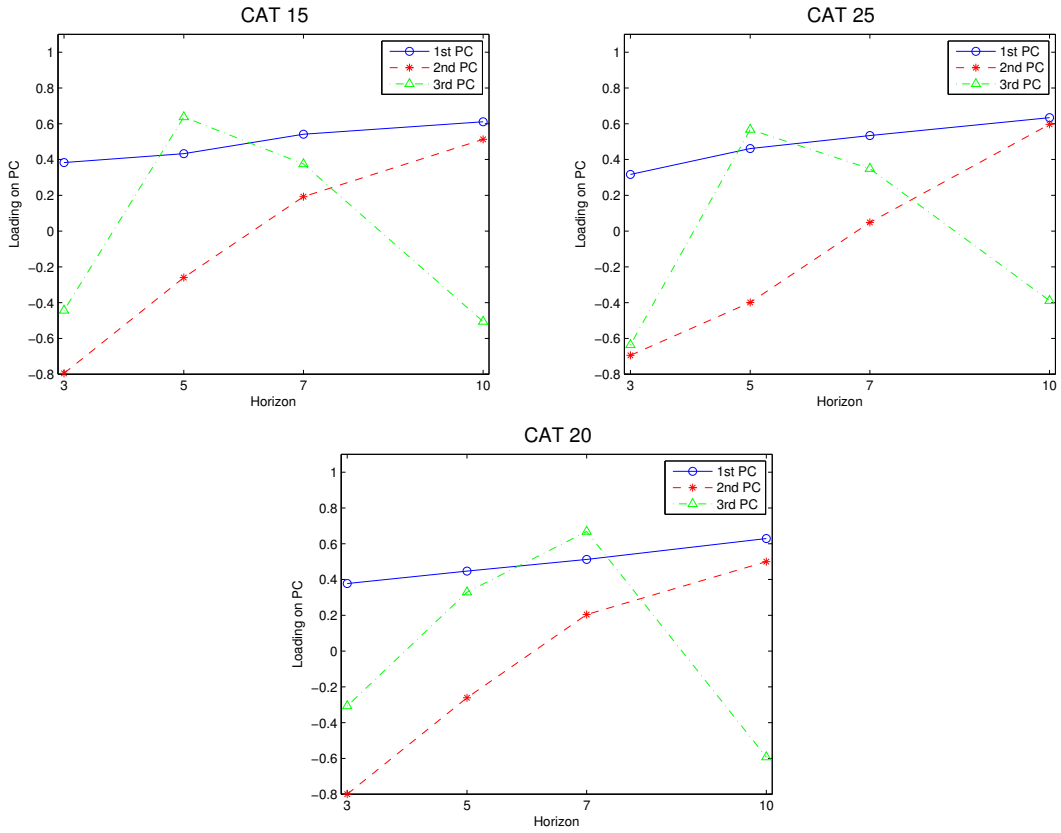
Note: The figure presents three histograms, each of which represents the risk-neutral distribution of catastrophic tail events during one of the three periods: the pre-crisis, crisis, and post-crisis periods. All distributions are under the 10-year horizon. Each histogram has 5 bars representing the 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 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 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 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 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: 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 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 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 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 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 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 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 10: Multivariate predictive regressions of 5-year maturity government bond returns

CAT	0.205 (4.87)	0.133 (3.80)	0.165 (5.64)	0.190 (4.07)	0.122 (2.95)
CP factor (1964-2013)	0.650 (1.45)			0.545 (1.02)	
CP factor (2003-2013)		1.102 (3.12)			1.039 (2.83)
Term spread			0.588 (1.83)	0.275 (0.61)	0.317 (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 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 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.