ESSAYS ON DISASTER RISK AND EQUITY RETURN PREDICTABILITY

A Dissertation

Presented to

The Faculty of the C.T. Bauer College of Business

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Shunlin Liang

May, 2016

ESSAYS ON DISASTER RISK AND EQUITY RETURN PREDICTABILITY

APPROVED:

Kris Jacobs, C.T. Bauer Professor of Finance

Seung Min Yae, Assistant Professor of Finance

Rauli Susmel, Associate Professor of Finance

Tong Lu, Associate Professor of Accounting

Latha Ramchand, Dean C.T. Bauer College of Business

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my advisor, Dr. James (Seung Min) Yae, and my co-chair Dr. Kris Jacobs, for their invaluable guidance and suggestions. Without their care and support, I would not be able to complete this dissertation. Under their guidance, I gain not only valuable knowledge but passion and confidence in finance research. Their insight and enthusiasm always inspire me.

I thank my dissertation committee members, Dr. Rauli Susmel, and Dr. Tong Lu, for their time and valuable advice. Their continuous support also contributes to the success of this dissertation.

I would also like to express my gratitude to Dr. Praveen Kumar, Dr. Ronald Singer, Dr. Stuart Turnbull, Dr. Thomas George, Dr. Nisan Langberg, Dr. Hitesh Doshi, Dr. Sang Byung Seo, and other faculties at the Department of Finance. They taught me fundamental knowledge and skills at the early stage of my PhD study, and were always supportive when I approach them for advice.

Many thanks go to Che-Kuan Chen, Xin Gao, Guanglian Hu, Rui Liu, and other PhD students at the Department of Finance for their friendship, discussion, encouragement, and companion. Grateful thanks to all of my friends in Houston, who have been very supportive of me.

My deepest thanks to my parents for their endless love and timely encouragement, my wife, Wei, for her love and unwavering support, and my daughters , Xinyue and Xindan, for bringing me joy.

ESSAYS ON DISASTER RISK AND EQUITY RETURN PREDICTABILITY

Abstract

Shunlin Liang

May, 2016

This dissertation consists of two essays on disaster risk and equity return predictability. The first essay proposes new measures of firm-level and market-level disaster risk from deviation of put-call symmetry, which is free from being contaminated by the asymmetry between option traders and equity investors. Compared with other known measures of disaster risk, the market-level disaster risk measure robustly predicts aggregate market returns, with out-of-sample ($R^2 = 6.86\%$) for the next twelve months. The cross-sectional analysis shows that firm-level disaster risk also explains variations in expected stock returns. Stocks with high firm-level disaster risk earn an annual four-factor subsequent alpha 8.0% higher than stocks with low firm-level disaster risk. I explore potential mechanisms giving rise to these asset pricing facts.

The second essay finds that the investor's learning of higher moments can account for the time-variation, size, and volatility of equity premium. I estimate the investor's belief on skewness and kurtosis of consumption and dividend growth, and assume investor's Bayesian learning about a skew student's t-distribution with unknown fixed parameters. The predictive regressions show that more negative skewness and higher kurtosis predict higher subsequent market excess returns, which implies the investor's learning generates the time-variation of equity premium although the true distribution is static. The calibrated asset pricing model shows that the investor's learning also explains the size and volatility of the equity premium observed in the data when the investor has a preference for early resolution of uncertainty.

Contents

List of Tables

List of Figures

Chapter 1

The Cross-Section and Time-Series of Disaster Risk Implied by Options

1.1 Introduction

Many of the most influential papers commonly use macroeconomic data (e.g. consumption, dividend price ratio, treasury bills) to predict asset returns, offering links between asset returns and the real economy. Mehra and Prescott (1985) show that fluctuations in the consumption growth over U.S. history can predict an equity premium that is far too small, assuming reasonable levels of risk aversion. Rietz (1988) proposes that the return on equities is high to compensate investors for the rare disaster risk. To test whether the risk is sufficiently high and whether the rare disaster is sufficiently severe, Barro (2006) collects annual data on aggregate GDP growth and asset prices for twenty countries over the 20th century, and shows that the world economy has indeed experienced mayor and rather frequent depressions, which are capable to explain the high equity risk premium and the behavior of the risk-free rate. Rietz-Barro hypothesis, which assumes a constant probability of a disaster over time, led to a fast growing literature on disaster risk models that focus on a number of extensions associated with the stochastic properties of the size and the probability of an economic disaster.

Extending the hypothesis proposed by Rietz (1988) and Barro (2006) to allow for a time-varying disaster probability increases the ability of the model to match the observed behavior of asset prices in many dimensions, Gabaix (2012) proposes a disaster risk model that a disaster reduces the fundamental value of stocks by a time-varying amount, which in return generates a time-varying stock risk premia. Gabaix's model provides a mechanism to explain ten puzzles in asset pricing: (i) the equity premium puzzle of Mehra and Prescott (1985); (ii) the risk-free rate puzzle of Weil (1989); (iii) the excess volatility puzzle of Shiller (1981); (iv) the aggregate stock market returns predictability puzzle of Campbell and Shiller (1988); (v) the cross-section predictability of stocks by stock characteristics vs. covariance with risk factors puzzle of Daniel and Titman (1997); (vi) the yield curve slope puzzle of Campbell (2003); (vii) the long term bond return predictability puzzle of Macaulay (1938), Fama and Bliss (1987) and Campbell and Shiller (1991); (viii) the credit spread puzzle of Almeida and Philippon (2007); (ix) the deep out-of-the money put prices vs. Black-Scholes' implied put prices of Jackwerth and Rubinstein (1996); and (x) the positive relation between high put option prices and their high future returns of Bollerslev, Tauchen, and Zhou (2009).

In this paper, following the disaster risk model in Gabaix (2012), I propose new measures of firm and market level disaster risk from option prices. Firm-level disaster risk measure is constructed as the deviation from put-call symmetry. The expression of deviation from put-call symmetry captures the stochasticity of disaster probability, consumption growth shock, and the stock recovery rate. Market-wide disaster risk measure is extracted from the firm-level disaster risk measure and explains the disaster risk probability induced by consumption shocks. Although firm-level disaster risk measure is a noisy measure as it includes dividend recovery rate and consumption shock, cross-sectionally, it captures only the dividend recovery risk, not the consumption shock risk.

A deviation from put-call symmetry (Carr and Chesney 1996)[1] is a simple relationship between put and call option prices, and provides a new channel to capture the investors' fear about the potential for a disaster in near future. When investors believe a high probability of disaster in the near future, they buy high amounts of put option to protect their equity holdings or to generate profit on option, which in turn increases the out-of-money put prices. Deviation from put-call symmetry is consistent with Bates (1991) that out-of-the-money puts became unusually expensive during the year preceding the crash of October 1987. Therefore, the magnitude of deviation from put-call symmetry represents the investors' belief in the disaster probability. Put-call symmetry is different from put-call parity in that put-call symmetry only depends on option prices, while put-call parity depends both on stock prices and option prices. When option traders have prior information than equity investors, the underlying stock price temperately deviates from the stock price implied from put-call parity. Therefore, deviation from put-call parity captures not only the investors' belief about the disaster probability, but also the asymmetry information between option traders and equity investors. It is difficult to separate these two effects from deviation from put-call parity. My firm-level disaster risk measure is constructed from put-call symmetry, which captures only the investors' belief about the disaster probability, and is free from the information asymmetry concern.

The firm-level disaster risk measure is constructed from the deviation from put-call symmetry. This approach takes advantage of the information contain in the cross-

[1]Under the standard model of frictionless markets and no arbitrage, put-call symmetry is the relationship holding for an American call and an American put when the "moneyness", time to maturity and the volatility structure happen to be the same. Carr and Lee (2009) further relax these assumptions to generalize to unified local/stochastic volatility models and time-changed Lévy processes. In contrast to put-call parity, put-call symmetry allows the underlying asset price, and strike price to differ for the call and the put. This extension can be applied to compare the values of an American call and put at a future date.

section of option. Option prices encode investors' ex-ante assessment of the expected disaster risk, and thus contain forward-looking information that expected to be related to the stock prices. My option-implied firm-level disaster risk is a forward looking measure of stock performance and does not requires the realization of a disaster to occur. Santa-Clara and Yan (2010) argue that the ex-ante disaster risk perceived by investors may be quite different from ex-post realized disaster risk in stock prices since even high probability disaster jumps may fail to materialize in sample. In the direction of option information predicting stock returns, I find that stocks with high firm-level disaster risk over the past month tend to have high returns over the next month. The strength and persistence of this predictability for stock returns from the cross section of firm-level disaster risk are remarkable for several reasons. First, based on Gabaix's model, the firm-level disaster risk reflects the time-varying dynamics of consumption and dividend growths. Second, the predictability is statistically strong and economically large. Decile portfolios formed on firm-level disaster risk have a spread of approximately 0.72% per month in both raw returns and alphas computed using common systematic factor models. From the t-statistics, the predictability of firm-level disaster risk on future returns persists up to three months. After three months, the economic and statistical significance of firm-level disaster risk portfolios disappears, which indicates that the information asymmetry dissipates, on average, within three months. There exists a U-shaped curve across different deciles returns for four-month and up holding periods.

From the expression of the firm-level disaster risk, market-level disaster risk can be extracted by eliminating the idiosyncratic component of firm-level disaster risk. Market-level disaster risk represents the common fluctuation in the disaster risk among individual stocks. This identifying assumption is that the firm-level disaster risk shares similar dynamics. I find that this time-varying market-level disaster risk varies substantially over time with monthly AR(1) coefficient of 0.32. Empirical analysis shows that the market-level disaster risk measure has strong predictive power for aggregate market returns both in-sample and out-of-sample. In the cross-section, stocks with high loadings on past market-level disaster risk earn a monthly three-factor alpha 0.75% higher than stocks with low market-level disaster risk loadings.

The predictability from option-implied firm-level disaster risk to stock returns is consistent with economies where informed traders choose the option market to trade first, such as those developed by Chowdhry and Nanda (1991) and Easley, O'Hara, and Srinivas (1998), which cause the option market to lead the stock market. My findings are also related to a recent literature showing that option prices contain predictive information about stock returns. Cao, Chen, and Griffin (2005) find that merger information hits the call option market prior to the stock market, but focus only on these special corporate events. Ang, Hodrick, Xing, and Zhang (2006) show that idiosyncratic volatility is inversely related to future stock returns. Bali and Hovakimian (2009), Cremers and Weinbaum (2010), Xing, Zhang, and Zhao (2010), Chang, Christoffersen and Jacobs (2013), Conrad, Dittmar, and Ghysels (2013), and An, Ang, Bali, and Cakici (2014) use information in the cross section of options including the difference between implied and realized volatilities, deviations from put-call parity, slope of volatility smirk, innovation in implied market skewness, risk-neutral skewness, and innovations in put and call volatility, respectively. My firm and market level disaster risk measures are also closely related to other tail risk measures in the following literatures. Backus, Chernov, and Martin (2011) use a constant disaster probability version of the Barro-Rietz hypothesis and S&P 500 index options to estimate the parameters of the risk-neutral and physical distribution of equity index returns. Bollerslev and Todorov (2011) use short-dated OTM options on the S&P 500 to estimate a risk-neutral tail measure. Du and Kapadia (2012) create an S&P

500 tail risk index, which capitalizes on the idea that the difference between quadratic variation and integrated variance should isolate the risk-neutral jump intensity in a general class of jump-diffusion models. Kelly and Jiang (2014) apply a power law of extreme value theory on the cross-section of historical equity returns to derive a time-varying measure of tail risk. This paper controls for all of these variables in examining the predictability of stock returns by the firm and market level disaster risk measures.

The remainder of this Chapter 1 is organized as follows. Section 1.2 presents the model to construct firm and market level disaster risk measures. Section 1.3 describes option price data and CRSP stock return data. Section 1.4 shows the empirical analysis for the stock return predictability by firm and market level disaster risk measures. Section 1.5 concludes.

1.2 Model of Rare Disasters

In this section, I first briefly summarize the rare disaster model by Gabaix (2012), and then describe how to construct the firm and market level disaster risk measures from stock option prices.

1.2.1 Gabaix's Rare Disaster Model

In line with Rietz (1988) and Barro (2006), Gabaix (2012) adds a stochastic probability and severity of disasters. Assume a representative agent with utility

$$U = E_0[\Sigma_{t=0}^{\infty} e^{-\rho t} \frac{C_t^{1-\gamma} - 1}{1-\gamma}] \tag{1.1}$$

where $\gamma \geq 0$ is the coefficient of relative risk aversion, $\rho > 0$ is the rate of time preference, and C_t is the consumption endowment at t.

At each period $t+1$, a disaster may happen with a probability p_t. If a disaster does not happen, $\frac{C_{t+1}}{C_t} = e^{g_C}$, where g_C is the constant normal-time growth rate of the economy. If a disaster happens, $\frac{C_{t+1}}{C_t} = e^{g_C} Z_{t+1}$, where $0 < Z_{t+1} < 1$ is a random variable. For example, if $Z_{t+1} = 0.7$, consumption falls by 30%.

$$\frac{C_{t+1}}{C_t} = e^{g_C} \times \begin{cases} 1 & \text{if there is no disaster at } t+1 \text{ with probability } 1-p_t, \\ Z_{t+1} & \text{if there is a disaster at } t+1 \text{ with probability } p_t, \end{cases} \quad (1.2)$$

Given the pricing kernel, the marginal utility of consumption $M_t = e^{\rho t} C_t^{-\gamma}$ follows

$$\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{if there is no disaster at } t+1 \text{ with probability } 1-p_t, \\ Z_{t+1}^{-\gamma} & \text{if there is a disaster at } t+1 \text{ with probability } p_t, \end{cases} \quad (1.3)$$

where $\delta = \rho + \gamma g_C$ is the risk-free rate in an economy that would have a zero probability of disasters. $Z_{t+1}^{-\gamma}$ measures the effect of disaster on the risk-free rate.

Assume a typical stock i with a claim on a stream of dividends $(D_{i,t})_{t\geq 0}$, the dividend process is given by

$$\frac{D_{i,t+1}}{D_{i,t}} = e^{g_{i,D}}(1 + \varepsilon_{i,t+1}^D) \times \begin{cases} 1 & \text{if no disaster at } t+1 \text{ with probability } 1-p_t, \\ F_{i,t+1} & \text{if a disaster at } t+1 \text{ with probability } p_t, \end{cases} \quad (1.4)$$

where $\epsilon_{i,t+1}^D > -1$ is a mean-zero shock that is independent of the disaster event. It matters only for the calibration of dividend volatility. If a disaster does not happen, $D_{i,t}$ grows at an expected rate of $g_{i,D}$. If a disaster happens, the dividend of stock i is partially wiped out by a random variable $0 \leq F_{i,t+1} < 1$, which is the dividend recovery rate in Longstaff and Piazzesi (2004) and Barro (2006). When $F_{i,t+1} = 0$, the dividend is completely destroyed or expropriated, and when $F_{i,t+1} = 1$, there is no dividend loss.

1.2.2 Risk Premium and Disaster Risk Measures

In this subsection, I introduce why disaster risk commands risk premium (Gabaix 2012), and how to construct the disaster risk measures from option prices.

1.2.2.1 Risk Premium

Rietz (1988) proposes that the possibility of rare disasters, such as economic depressions or wars, is a major determinant of asset risk premia. Based on international market, Barro (2006) shows that disasters have been sufficiently frequent and large to make Rietz's proposal viable and account for the high risk premium on equities. Gabaix (2012) further allows the stochasticity to come both from movements in the probability of disasters and from the expected recovery rate of assets to explain puzzles in stock market, bond, and option. To test whether rare disaster risk commands risk premium, we first derive the expected stock returns as following:

From the Euler equation, $1 = E_t[(1 + r_{i,t+1})M_{t+1}/M_t]$, and Eq (1.3), we let

$$1 = e^{-\delta}(1 - p_t)E_t^{ND}(1 + r_{i,t+1}) + e^{-\delta} p_t E_t^D[Z_{t+1}^{-\gamma}(1 + r_{i,t+1})], \tag{1.5}$$

where the first and second terms in the RHS are the no disaster (ND) term and disaster (D) term, respectively. Rearranging the Eq (1.5), we can connect asset i's return $r_{i,t+1}$ at $t + 1$, and expected return $r_{i,t}^e = E_t^{ND}(r_{i,t+1})$ at t conditional on no disaster as follows:

$$r_{i,t}^e = E_t^{ND}(r_{i,t+1}) = \frac{1}{1 - p_t}(e^{\delta} - p_t E_t^D[Z_{t+1}^{-\gamma}(1 + r_{i,t+1})]) - 1 \tag{1.6}$$

In the limit of small time intervals:

$$r_{i,t}^e = \delta - p_t E_t^D[Z_{t+1}^{-\gamma}(1 + r_{i,t+1}) - 1] = r_f - p_t E_t^D[Z_{t+1}^{-\gamma} r_{i,t+1}] \tag{1.7}$$

where $r_f = \delta - p_t E_t^D[Z_{t+1}^{-\gamma} - 1]$ is the real risk free rate in the economy. Eq (1.7) indicates that only the behavior in disasters (the second term on the RHS) creates a risk premium. It is equal to the risk free rate that adjusted by the expected capital loss if there is a disaster.

If a disaster occurs, dividends are multiplied by $F_{i,t}$. Comparing resilience of stock i $H_{i,t} = p_t E_t^D[Z_{t+1}^{-\gamma} F_{i,t+1} - 1]$ (see Appendix A) and Eq (1.7), we have $1 + r_{i,t+1} = F_{i,t+1}$ as $\widehat{H}_{i,t}$ does not change,

$$r_{i,t}^e = \delta - p_t E_t^D[Z_{t+1}^{-\gamma} F_{i,t+1} - 1] = \delta - H_{i,t}, \tag{1.8}$$

and

$$r_{i,t}^e - r_f = p_t E_t^D[Z_{t+1}^{-\gamma} (1 - F_{i,t+1})]. \tag{1.9}$$

Eq (1.9) shows that the expected excess stock returns are high with large consumption drops, and are low with high recovery rates. It indicates that consumption and firm-level disaster risk reduce the fundamental value of a stock by a time-varying amount, and in turn generates time-varying risk premia.

1.2.2.2 Firm-level Disaster Risk Measure

Within the disaster risk model, option prices have a “volatility smirk", where out-of-the-money puts have high prices and implied volatility. Inline with Gabaix's findings, I show that deviation from put-call symmetry represents the risk premium driven by disasters, and thus is a good proxy for firm-level disaster risk.

Deviation from put-call symmetry is a better measure for disaster risk than other option-implied measures (e.g. the slope of volatility smirk, and risk neutral skewness) since deviation from put-call symmetry only requires option prices, while other measures require stock prices as well as options prices. Therefore, a potential information asymmetry between option traders and equity investors can contaminate those measures, while deviation from put-call symmetry is free from such concern. Under the standard model of frictionless markets and no arbitrage, put-call symmetry is the relationship holding for an American call and an American put when the “moneyness", time to maturity and the volatility structure happen to be the same

(Carr and Chesney 1996). For example, it implies that if a forward price M follows geometric Brownian motion under an appropriate pricing measure, $M_0 = 100$, then a 200-strike call on M has time-0 price equal to two times the price of the 50-strike put at the same expiry, which indicates $OP^{put}(k,\sigma) = k\,OP^{call}(k^{-1},\sigma)$, where k is the moneyness.

Following the approach in Gabaix (2012), we derive the put and call option prices with or without disaster (see Appendix A) as follows:

$$\begin{aligned} OP^{put}_{i,t}(K) &= OP^{ND,put}_{i,t}(K) + OP^{D,put}_{i,t}(K) \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,E^{ND}_t[(K\,e^{-\mu_{i,t}} - e^{\sigma_i u_{i,t+1}-\sigma_i^2/2})^+] \\ &\quad +(1-p_t)\,e^{\mu_{i,t}-\delta}\,OP^{BS,put}(K\,e^{-\mu_{i,t}},1,\sigma_i) \end{aligned} \tag{1.10}$$

$$\begin{aligned} OP^{call}_{i,t}(K) &= OP^{ND,call}_{i,t}(K) + OP^{D,call}_{i,t}(K) \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,E^{ND}_t[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2} - K\,e^{-\mu_{i,t}})^+] \\ &\quad +p_t\,e^{\mu_{i,t}-\delta}\,E^{D}_t[Z^{-\gamma}_{t+1}(F_{i,t+1} - K\,e^{-\mu_{i,t}})^+] \end{aligned} \tag{1.11}$$

where $OP^{BS,call(put)}(1,K,\sigma_i)$ is the Black-Scholes value of a call (put) with strike K, volatility σ_i, initial price 1, maturity 1, and interest rate 0.

Next, the firm-level disaster risk is derived as deviation from put-call symmetry as follows:

$$\begin{aligned} FDR_{i,t}(k) &= OP^{put}_{i,t}(k) - k\;\; OP^{call}_{i,t}(k^{-1}) \\ &= e^{\mu_{i,t}-\delta}\,(1-p_t)[OP^{BS,put}(k\,e^{-\mu_{i,t}},\sigma_i) - k\;\; OP^{BS,call}(k^{-1}\,e^{-\mu_{i,t}},\sigma_i)] \\ &\quad +e^{\mu_{i,t}-\delta}\,p_t\,E^D_t[Z^{-\gamma}_{t+1}\,((k\,e^{-\mu_{i,t}} - F_{i,t+1})^+ - k\;\;(F_{i,t+1} - k^{-1}\,e^{-\mu_{i,t}})^+)] \\ &= e^{\mu_{i,t}-\delta}\,p_t\,E^D_t[Z^{-\gamma}_{t+1}\,((k\,e^{-\mu_{i,t}} - F_{i,t+1})^+ - (k\,F_{i,t+1} - e^{-\mu_{i,t}})^+)] \end{aligned} \tag{1.12}$$

where moneyness $k = K/1$, and $OP^{BS,put}(k\,e^{-\mu_{i,t}},\sigma_i) - k\;\; OP^{BS,call}(k^{-1}\,e^{-\mu_{i,t}},\sigma_i) = 0$. Under the assumption that in the limit of small time intervals and $k\,e^{-\mu_{i,t}} > F_{i,t+1}$, the firm-level disaster risk $FDR_{i,t}$ can be deduced as

$$FDR_{i,t}(k) = e^{\mu_{i,t}-\delta}\, p_t\, E_t^D[Z_{t+1}^{-\gamma}\,(k\, e^{-\mu_{i,t}} - F_{i,t+1})] \tag{1.13}$$

where $\delta = \rho + \gamma\, g_c$ is the risk-free rate in an economy that would have a zero probability of disasters, $\mu_{i,t} = g_{i,D} + \ln[\frac{a+b\, e^{-\phi_H}\, \widehat{H}_{i,t}}{a+b\widehat{H}_{i,t}}]$ is the expected dividend growth of firm i.

The $FDR_{i,t}$ depends on economy-wide variables of disaster probability p_t, consumption jump Z_{t+1}, and firm-level recovery rate $F_{i,t+1}$. To derive the effects of above variables on firm-level disaster risk measure, we take the first order derivative of Eq (1.13) for all variables. $(\partial\, FDR/\partial\, \delta) < 0$ represents low firm-level disaster risk at the time of high risk-free rate. $(\partial\, FDR/\partial\, \mu_{i,t}) < 0$ shows that stocks with high expected dividend growth will have low firm-level disaster risk. $(\partial\, FDR/\partial\, F_{i,t+1}) < 0$ shows that stock with high recovery rate will have low firm-level disaster risk. At the time of large consumption drops (low Z_{t+1}), there is high firm-level disaster risk. Therefore, the firm-level disaster risk measure captures not only the market effects from risk-free rate and consumption growth, but also the firm-level effects from dividend growth and recovery rate. Although this firm-level disaster risk measure is a noisy measure since it includes firm and market level disaster risk, but in the cross-section, it represents the firm-level disaster risk, not the market-level disaster risk.

1.2.2.3 Market-wide Disaster Risk Measure

The expression of firm-level disaster risk in Eq (1.13) suggests a way to extract key structural parameters of disasters, such as market-level disaster risk (MDR). Suppose there are two firm-level disaster risks, $FDR_{i,t}(k_1)$ and $FDR_{i,t}(k_2)$ for stock i with moneyness k_1 and k_2, respectively,

$$FDR_{i,t}(k_1) - FDR_{i,t}(k_2) = e^{-\delta}\,(k_1 - k_2)\, E_t^D[p_t\, Z_{t+1}^{-\gamma}], \tag{1.14}$$

then we can have

$$MDR = E_t^D[p_t\, Z_{t+1}^{-\gamma}] = \frac{FDR_{i,t}(k_1) - FDR_{i,t}(k_2)}{e^{-\delta}\,(k_1 - k_2)}, \tag{1.15}$$

where $\delta = \rho + \gamma\, g_c$ is the risk-free rate in an economy that would have a zero probability of disasters. I calibrate time preference $\rho = 6.57\%$, risk aversion $\gamma = 4$ and the normal-time growth rate of the economy g_c as the average consumption growth rate over past 36 months.

After the market-level disaster risk is derived, the firm-level expected recovery rate $E_t^D[p_t\, Z_{t+1}^{-\gamma}\, F_{i,t+1}]$ can be estimated by

$$E_t^D[p_t\, Z_{t+1}^{-\gamma}\, F_{i,t+1}] = e^{\delta - \mu_{i,t}}\, \frac{k_2\, FDR_{i,t}(k_1) - k_1\, FDR_{i,t}(k_2)}{k_1 - k_2} \tag{1.16}$$

where the firm-level expected recovery rate can capture the effects of disaster risk probability, consumption jump, and firm's recovery rate. I will leave this for future research.

1.3 Data

In this section, I describe the sources of the data used in this paper. The data includes option data from OptionMetrics, underlying stock return data from CRSP, and accounting and balance sheet data from COMPUSTAT.

1.3.1 Option Data

The option data originates from OptionMetrics. This database provides end-of-day bid and ask quotes, trading volume, open interest, and option-specific data (e.g., implied volatility, maturity, strike price, delta) for all call and put options on stocks, exchange traded funds (ETFs), and index traded on U.S. exchanges. It also provides the stock price and dividends of the underlying stocks and zero-coupon interest rates. All optionable stocks and ETFs have American-style options. Among the broad-based indices, only limited indices such as the S&P 100 have American-style options. Major broad-based indices, such as the S&P 500, have very actively traded European-style options. Owners of American-style options may exercise at any time before the option expires, while owners of European-style options may exercise only at expiration. American index options cease trading at the close of business on the third Friday of the expiration month. European index options stop trading one day earlier, at the close of business on the Thursday preceding the third Friday.

I use data for options with 10-180 days to expiration on S&P 500 index and stocks from January 1996 to August 2013. I form option pairs that are used to construct the synthetic stock. An option pair consists of an OTM put option with moneyness $k \in (0.80 \sim 0.95)$ on the underlying stock matched with a call option with moneyness $k^{-1} \in (1.05 \sim 1.25)$. I discard option pairs where the quotes for either the call or the put option are locked or crossed. I keep only those option pairs for which the volume for both the call and put is greater than zero and the implied volatility (calculated using the binomial option pricing model) for the call and put is defined. The option prices are taken as the midpoints of the bid and ask quotes, which are the best closing prices across all exchanges on which the option trades.

Table 1.1 contains descriptive statistics for the sample. This table reports the average number of stocks per month for each year from 1996 to 2013. There are 1,455 stocks per month in 1996, and increase to 2,165 stocks per month in 2013. I report the average, standard deviation and the quantiles (e.g. 1%, 25%, 50%, 75% and 99%) of the end-of-day derivation from OTM (moneyness = k) put price and OTM (moneyness = k^{-1}) call price, which is the firm-level disaster risk measure. This firm-level disaster risk measure is high during 2002 and 2003, which coincides with the large decline in stock prices, particularly of technology stocks. During the recent financial crisis in 2008 to 2009, there is a significant increase in average firm-level disaster risk to 8.80, which means the option price of an OTM put is 8.80% more expensive than that of an OTM call.

Figure 1.1 shows the value-weighted firm-level disaster risk across firms. Firm-level disaster risk is expressed in daily terms with 25%, 50%, and 75% quantiles, respectively. The firm-level disaster risk across firms is highly persistent, with a monthly AR(1) coefficient of 0.95. Figure 1.2 shows the market-level disaster risk measure, corresponding to different crisis events in financial markets. This market-level disaster risk can be viewed as the conditional probability that the market has its worst return realizations conditional on previous month market return. Compared with the time series of market returns, higher market-level disaster risk commands with higher subsequent market returns. Figure 1.3 shows the time series of four disaster risk: FDR (average stock) a cross-sectional average of firm-level disaster risk; FDR (market) derived from the S&P 500 index options; MDR (average stock) a cross-sectional average of firm-level market-level disaster risk; and MDR (market) derived from the S&P 500 index options.

1.3.2 Stock Return Data and Other Predictive Variables

I obtain underlying stock return data from CRSP, and accounting and balance sheet data from COMPUSTAT. I also construct the following factor loadings and firm characteristics associated with underlying stock markets that are widely known to forecast the cross section of stock returns.

SIZE: Firm size is measured as the natural logarithm of the market value of equity, which is equal to stock price multiplied by the number of shares outstanding in millions of dollars at the end of the month for each stock.

Book to Market Ratio (BM): Following Fama and French (1992), I compute a firm's book-to-market ratio in month t using the market value of its equity at the end of December of the previous year and the book value of common equity plus balance-sheet deferred taxes for the firm's latest fiscal year ending in the prior calendar year.

Momentum (MOM): Following Jegadeesh and Titman (1993), I compute the momentum variable for each stock in month t as the cumulative return on the stock over the previous 6 months starting two months ago to avoid the short term reversal effect, that is, momentum is the cumulative return from month t-7 to month t-2.

Short-Term Reversal (REV): Following Jegadeesh (1990), I obtain the short-term reversal for each stock in month t as the return on the stock over the previous month from t-1 to t.

Realized Volatility (RVOL): Realized volatility of stock i in month t is defined as the standard deviation of daily returns over the past month t.

Implied Volatility Innovations (ΔPIVOL and ΔCIVOL): Following An, Ang, Bali, and Cakici (2014), I compute the implied put (call) volatility innovations as the change in OTM put (call) implied volatilities from previous month.

Call/Put Volume (C/P-VL): This measure is defined as the ratio of call/put option trading volume at the same month for each stock.

Call/Put Open Interest (C/P-OI): This measure is defined as the ratio of the open interest of call options to that of put options at the same month.

Realized-Implied Volatility Spread (RIVOL): I control for the difference between the monthly realized volatility (RVOL) and the average of the at-the money call and put implied volatilities (IVOL). Bali and Hovakimian (2009) and Goyal and Saretto (2009) show that stocks with high RVOL-IVOL spreads have low future stock returns.

Slope of Volatility Smirk (VOLSKEW): Following Xing, Zhang, and Zhao (2010), I control for slope of volatility smirk, defined as the difference between the out-of-the-money put implied volatility (with moneyness of 0.80-0.95) and the average of the at-the-money call and put implied volatilities (with moneyness of 0.95-1.05), both using maturities of 7-60 days. Xing, Zhang, and Zhao show that stocks with high VOLSKEW tend to have low returns over the following month. In contrast, Conrad, Dittmar, and Ghysels (2013) report the opposite relation using a more general measure of risk-neutral skewness based on Bakshi, Kapadia, and Madan (2003), which is derived using the whole cross section of options.

Risk Neutral Skewness and Kurtosis: Following Bakshi, Kapadia, and Madan (2003), I control for a more general measure of risk-neutral skewness and kurtosis, which are derived from the S&P 500 index options.

Tail Risk (Tail-KJ): Following Kelly and Jiang (2014), I control for common fluctuations in the cross section of stock returns. Kelly and Jiang show that this tail risk has strong predictive power for aggregate market returns. I thank Bryan Kelly for providing the monthly tail risk data.

Investor Fears Index (IFI-BT): Following Bollerslev and Todorov (2011), I control for the Investor Fears Index, which are based on 5-minute S&P 500 futures prices and options. Investor Fears Index $IFI = VRP_t^-(k) - VRP_t^+(k)$, where $VRP_t^-(k)$ ($VRP_t^+(k)$)is the negative (positive) jump variance risk premium. Bollerslev and Todorov show that this investor fears index reveals large time-varying compensation for fears of disasters.

Predictors in Goyal and Welch (2008): Following Goyal and Welch, I control for other predictors as default return spread, default yield spread, dividend payout ratio, dividend price ratio, earnings price ratio, inflation, long-term bond return, long-term yield, net equity expansion, term spread, and treasury-bill rate. The data is from Amit Goyal's Website.

1.4 Empirical Analysis

In this section, I first investigate whether market-level disaster risk can forecast market returns in-sample and out-of-sample, and whether market-level disaster risk can help explain differences in future returns across stocks. Then, I study the firm-level disaster risk across firms, and show how this firm-level disaster risk may help explain differences in expected returns across stocks.

1.4.1 Market Return Predictability by Market-wide Disaster Risk

I test whether market-level disaster risk can forecast aggregate market returns. All regressions are conducted at the monthly frequency and the analyses are for one-, three-, six-, and twelve-month horizons. The dependent variable is the return on the CRSP value-weighted index at the monthly frequency. To illustrate economic magnitudes, all reported predictive coefficients are scaled to be interpreted as the effect of one-standard-deviation increase in the regressor on future annualized returns. Table 1.2 shows that market-level disaster risk has significant forecasting power over one to twelve months. A one-standard-deviation increase in lower market-level disaster risk predicts an increase in future excess returns of 1.73%, with a t-statistic of 4.67 at one-month horizon. Table 2 also compares the forecasting power of market-level disaster risk with a large set of forecasting variables studied in a survey by Goyal and Welch (2008), and other risk measures, as well as the investor fear index (Bollerslev and Tauchen 2011), risk-neutral skewness and kurtosis based on S&P 500 index options (Bakshi, Kapadia, and Madan 2003), and tail risk from cross-sectional stock returns (Kelly and Jiang 2014). The long-term bond return and stock volatility strongly predict up to three-month and nine-month returns, respectively. The effects of tail risk, investor fear index and market-level disaster risk can persist up to twelve months.

Table 1.3 reports results of bivariate predictor performance from monthly predictive regressions of market returns on market-level disaster risk over one, three, six, and twelve-month horizons. The predictive ability of market-level disaster risk is unaffected by including alternative regressors. For one-month forecasts, the market-level disaster risk predictive coefficient remains above 1.3% when combined with each of other predictive variables, with a t-statistic above 3.2 in all cases. At longer horizons, the performance of market-level disaster risk is less strong, but still statistically significant, except the one that at twelve-month horizon and combined with net equity expansion.

Next, I investigate the out-of-sample predictive ability of market-level disaster risk. Using data only through month t (beginning at $t = 80$ to allow for sufficiently large initial estimation period), I run univariate predictive regressions of market returns on market-level disaster risk. This coefficient is used to forecast the $t + 1$ market return. The estimation window is then extended by one month to obtain a new predictive coefficient, and an out-of-sample forecast of the following month's return is constructed. This procedure is repeated until the full sample has been exhausted. Since coefficients are based only on data through t, this procedure mimics the information set that an investor would update with in real time. Using the forecast errors from this approach, I calculate the out-of-sample R^2 as

$$R^2 = 1 - \sum_t (r_{m,t+1} - \hat{r}_{m,t+1|t})^2 / \sum_t (r_{m,t+1} - \bar{r}_{m,t})^2, \tag{1.17}$$

where $\hat{r}_{m,t+1|t}$ is the out-of-sample forecast of the $t + 1$ return based only on data through t, $\bar{r}_{m,t}$ is the historical average market return through t. A negative R^2 indicates that the predictor performs worse than setting forecasts equal to the historical mean. This out-of-sample forecast approach is applied using each of the alternative predictors. Table 1.4 shows the results that market-level disaster risk forecasts demonstrate similar predictive success out-of-sample. At the one-, three-, six-, and twelve-month horizons, the market-level disaster risk out-of-sample R^2 is 0.67%, 1.34%, 2.80%, and 6.86% versus 2.69%, 2.03%, 6.36%, and 4.07% in-sample. A negative R^2 implies that the predictor performs worse than setting forecasts equal to the historical mean. A star next to out-of-sample R^2 is based on significance of MSE-F statistic by McCracken (2004), which tests for equal MSE of the unconditional forecast and the conditional forecast. We can see that market-level disaster risk has significant out-of-sample R^2. Other predictors, such as BM, net equity expansion, IFI-BT, and Tail-KJ, also have positive out-of-sample R^2. Specially, IFI-BT, and Tail-KJ have significant out-of-sample R^2 for twelve-month holding period. In summary, predictive regressions suggest that market-level disaster risk is positively and significantly related to aggregate market returns.

1.4.2 Cross-section Stock Return Predictability by Market-wide Disaster Risk

In the next, I test whether market-level disaster risk (MDR) can help explain differences in future returns across stocks. This is consistent with the hypothesis that investors are averse to market-level disaster risk, stocks with high predictive loadings on market-level disaster risk will be discounted more steeply and thus have higher future returns, and stock with low predictive loadings on market-level disaster risk will have comparatively higher prices and thus have lower future returns.

I then estimate the market-level disaster risk sensitivities of individual stocks with regressions of $E_t[R_{i,t+1}] = \mu_i + \beta_i\, MDR_t$. Stocks with high values of β_i are those that are most sensitive to market-level disaster risk, and thus steeply discounted when market-level disaster risk is high and have higher future return. On the other hand, stocks with low values of β_i are those that are less sensitive to market-level disaster risk, and thus have comparatively higher prices and have lower future returns. In each month, I estimate the market-level disaster risk loading for each stock in regressions that use the most recent 24 months of data. Stocks are sorted into quintile portfolios based on their estimated market-level disaster risk loadings. Then, the average monthly value-weighted (equal-weighted) quintile portfolio returns is derived in a one-month post-formation window.

Panel A of Table 1.5 reports the value- and equal-weighted average out-of-sample one-month holding period portfolio returns. Stocks in the highest market-level disaster risk loading quintile earn value-weighted average monthly returns 0.65% higher than stocks in the lowest quintile, with a Newey-West t-statistic of 2.57, where lag length equal to the number of month in each horizon. Average portfolio return demonstrate a monotonic pattern that is increasing in MDR. To consider alternative priced factors, I report alphas from regressions of portfolio returns on CAPM, Fama-French three factor model, Fama-French-Carhart four factor momentum model, and Fama-French-Carhart plus the Pastor and Stambaugh (2003) traded liquidity five factor model. Alphas of the value-weighted high-minus-low MDR portfolio are large and statistically significant for each of these models. For CAPM model, the alpha is 0.73% per month (t=2.49). For the three-factor model, the alpha is 0.75% per month (t=2.57). For the four-factor model, the alpha is 0.70% per month (t=2.61). For the five-factor model, the alpha is 0.76% per month (t=2.56). Portfolio alphas also demonstrate a monotonic pattern that was observed for average portfolio returns. Panel B report the out-of-sample twelve-month holding period portfolio returns, which show that long horizon portfolio returns have the same qualitative behavior as those of short horizons.

To test the effect of innovation of market-level disaster risk on future returns across stocks, I estimate the $\triangle$MDR sensitivities of individual stocks with $E_t[R_{i,t+1}] = \mu_i + \beta_i \triangle MDR_t$, where $\triangle MDR_t = MDR_t - MDR_{t-1}$. Stocks with high values of β_i are those that are most sensitive to $\triangle$MDR, and thus steeply discounted when $\triangle$MDR is high and have higher future return. On the other hand, stocks with low values of β_i are those that are less sensitive to $\triangle$MDR, and thus have comparatively higher prices and have lower future returns.

1.4.3 Cross-Section Stock Return Predictability by Firm-level Disaster Risk

In this subsection, I study the cross-section stock return predictability by firm-level disaster risk measure. I estimate the differences in returns between deciles sorted by firm-level disaster risk, study the longer-term predictive power of firm-level disaster risk, and examine whether firm-level disaster risk can predict the next month's returns while controlling for firm characteristics and other known disaster risk measures.

1.4.3.1 Portfolios Sorted by Firm-level Disaster Risk

To investigate whether firm-level disaster risk (FDR) may help explain differences in expected returns across stocks, I sort firm-level disaster risk into 10 portfolios. Portfolio 1 (Low FDR) contains stocks with the lowest individual FDR in the previous month and Portfolio 10 (High FDR) includes stocks with the highest individual firm-level disaster risk in the previous month. The decile portfolios are rebalanced every month, and stocks in each decile portfolio are value-weighted.

Table 1.6 shows that the average raw return of stocks in decile 1 with the lowest firm-level disaster risk is 0.31% per month and this monotonically increases to 1.03% per month for stocks in decile 10. The difference in average raw returns between deciles 1 and 10 is 0.72% per month, with a highly significant Newey-West t-statistic of 3.46. This result translates to a monthly Sharpe ratio of 0.17 and an annualized

Sharpe ratio of 0.61 for an investment strategy taking a long position in high firm-level disaster risk stocks and a short position in low firm-level disaster risk stocks. The differences in returns between deciles 1 and 10 are very similar if I risk-adjust using the CAPM, at 0.68% per month (t-statistic of 3.32), and the Fama-French (1993) three-factor model including market, size, and book-to-market factors, at 0.72% per month (t-statistic of 3.19). I also report alphas from regressions of portfolio returns on the Fama-French-Carhart four-factor momentum model and the Fama-French-Carhart model plus the Pastor and Stambaugh (2003) traded liquidity factor as a fifth control. Alphas of the equal-weighted high minus low firm-level disaster risk portfolio are large and statistically significant for each of these models. Portfolio alphas demonstrate a monotonic pattern that was observed for average portfolio returns. The results indicates that stocks that have high firm-level disaster risk over the past month tend to have high returns in the next month.

1.4.3.2 Long-Term Predictability of Portfolios Sorted by Firm-level Disaster Risk

I further investigate the longer-term predictive power of firm-level disaster risk up to twelve months by constructing portfolios with overlapping holding periods following Jegadeesh and Titman (1993). In a given month t, the strategy holds portfolios that are selected in the current month as well as in the previous N-1 months, where N is the holding period (N=1 to 12 months). At the beginning of each month t, I perform dependent sorts on firm-level disaster risk over the past month. Based on these rankings, 10 portfolios are formed for firm-level disaster risk. In each month t, the strategy buys stocks in the high firm-level disaster risk decile and sells stocks in the low firm-level disaster risk decile, holding this position for N months. In addition, the strategy closes out the position initiated in month t-N. Hence, under this trading strategy, I revise the weights on $1/N$ of the stocks in the entire portfolio in any given month and carry over the rest from the previous month. The profits of the above strategies are calculated for a series of portfolios that are rebalanced monthly to maintain value weights.

I report the long-term predictability results in Table 1.7. The average raw and risk-adjusted return differences between high firm-level disaster risk and low firm-level disaster risk portfolios are statistically significant for one to three month holding periods. There is a pronounced drop in the magnitude of the average holding return, which is reduced by 40% between holding period one-month and two-month from 0.72% per to 0.44%. There is a further reduction to 0.18% for four-month holding period. From the t-statistics, the predictability of firm-level disaster risk on future returns persists up to three months. After three months, the economic and statistical significance of firm-level disaster risk portfolios disappears, which indicates that the information asymmetry dissipates, on average, after three months. Holding the firm-level disaster risk portfolios longer than 4 months, there exists a U-shaped returns across low to high firm-level disaster risk portfolios deciles.

1.4.3.3 Firm Characteristics of Portfolios Sorted by Firm-level Disaster Risk

To highlight the firm characteristics, risk, volatility spread, and skewness attributes of stocks in the portfolios of firm-level disaster risk, Table 1.8 presents descriptive statistics for the stocks in the various deciles. The decile portfolios in Table 1.8 are formed by sorting stocks based on firm-level disaster risk. In each month, I record the median values of various characteristics within each portfolio. These characteristics are all observable at the time the portfolios are formed. Table 1.8 reports the average of the median characteristic values across months of market beta (BETA), log market capitalization (SIZE), book-to-market (BM), the cumulative return over the 6 months prior to portfolio formation (MOM), the return in the portfolio formation month (REV), the slope of volatility smirk (VOLSKEW), the realized volatility (RVOL), the realized minus implied volatility (RIVOL), the implied volatility spread (IVOL-S), and innovation of out-of-the-money put (call) volatility ΔPIVOL (ΔCIVOL).

From Table 1.8, the low firm-level disaster risk to the high firm-level disaster risk decile, the average return on firm-level disaster risk portfolios increases from 0.31% to 1.03%. The return spread between the extreme decile portfolios is 0.72% per month, with a t-statistic of 3.46. There are no discernible patterns of market beta, size, book-to-market and momentum across the portfolios. I investigate whether the skewness or volatility spread attributes of stocks provide an explanation for the high returns of stocks with large firm-level disaster risk. There is a pronounced pattern of decreasing the slope of volatility smirk (VOLSKEW) moving from 6.21 for the Low1 FDR decile to 3.56 for the High10 FDR decile. While there is a significant pattern of increasing implied volatility spread (IVOL-S) moving from 2.15 for the Low1 FDR decile to 8.46 for the High10 FDR decile. Note that VOLSKEW is computed as the spread between the implied volatilities of out-of-the-money puts and the average of at-the-money put and calls, and IVOL-S is computed as the spread between implied volatilities of out-of-the-money puts and out-of-the-money calls. Decreasing VOLSKEW across the FDR deciles is equivalent to these stocks experiencing simultaneous increases in put volatilities as firm-level disaster risk increases. This is consistent with informed trading whereby informed bearish investors with a high degree of confidence in future price depreciation buy puts and sell calls. In cross-sectional regressions, I control for VOLSKEW and IVOL-S along with other regressors in examining firm-level disaster risk predictability.

1.4.3.4 Cross-Sectional Regressions with Firm-level Disaster Risk and Other Predictors

Finally, I illustrate that the firm-level disaster risk predicts underlying equity returns. I argue that firm-level disaster risk reflects investors' expectation of a downward price jump. If informed traders choose the options market to trade in first and the stock market is slow to incorporate the information embedded in the options market, then I should see the information from the options market predicting future stock returns. While the analysis in Table 1.8 shows that most firm characteristics measures are unlikely to play a role in predictability of the cross section of stock returns sorted by firm-level disaster risk, it does not control simultaneously for multiple sources of risk. I conduct a Fama-MacBeth (1973) regression to examine whether firm-level disaster risk can predict the next month's returns, while controlling for different firm characteristics. I run the following cross-sectional regression:

$$R^e_{i,t+1} = \beta_{0,t} + \beta_{1,t}\, FDR_{i,t} + \beta'_{2,t}\, CONTROLS_{i,t} + \epsilon_{i,t+1} \tag{1.18}$$

where $R^e_{i,t+1}$ is the realized excess return on stock i in month $t+1$ and $CONTROLS_{i,t}$ is a collection of stock-specific control variables observable at time t for stock i, which includes information from the cross section of stocks and the cross section of options. I estimate the regression in this equation across stocks i at time t and then report the cross-sectional coefficients averaged across the sample. The cross-sectional regressions are run at the monthly frequency from January 1996 to August 2013. To compute standard errors, I take into account potential autocorrelation and heteroscedasticity in the cross-sectional coefficients, and then compute Newey-West (1987) t-statistics on the time series of slope coefficients. The Newey-West standard errors are computed with twelves lags.

Table 1.9 reports the average firm-level cross-correlations of stocks' firm-level disaster risk (FDR), innovation of firm-level disaster risk (ΔFDR), market capitalization (SIZE), book-to-market (BM), the cumulative return over the 6 months prior to portfolio formation (MOM), the return in the portfolio formation month (REV), implied volatility spread (IVOL-S), realized volatility (RVOL), realized minus implied volatility (RIVOL), slope of volatility smirk (VOLSKEW), innovation of implied put (call) volatility ΔPIVOL (ΔCIVOL), ratio of put call open interest (P/C-OI), ratio of put call volumn (P/C-VL), tail risk (Tail-KJ) by Kelly and Jiang (2014), and investor fear index (IFI-BT) by Bollerslev and Todorov (2011). Table 1.9(continues) reports the average firm-level cross-correlations of stocks' FDR measure, and the other predictors in Goyal and Welch (2008).

Table 1.10 presents firm-level cross-sectional regressions with firm-level disaster risk, together with controls for firm characteristics and risk factors. In the presence of risk loadings and firm characteristics, regression (1) in Table 1.11 shows that the average slope coefficient on FDR is 0.158, which is highly significant with a t-statistic of 3.26. In regression (2), the average slope coefficient on ΔFDR is 0.027 with a t-statistic of 2.88. In regression (3), the average slope coefficient on IVOL-S is 0.043 with a t-statistic of 3.12. In regression (4), the average slope coefficient on VOLSKEW is 0.096 with a t-statistic of 3.50. In regression (5), the average slope coefficients on ΔCIVOL (ΔPIVOL) is -0.133 (0.084) with a t-statistic of -2.42 (0.92). In regression (6), the average slope coefficient on Tail-KJ is 0.074 with a t-statistic of 1.98. In regression (7), the average slope coefficient on IFI-BT is -0.051 with a t-statistic of -2.30. Regression (8,9,10) shows that the average slope coefficient on FDR is positive and significant. Table 1.10 provides no evidence for a significant link between trading volume P/C-VL (or open interest P/C-OI) and the cross-section of expected returns, which is consistent with Pan and Poteshman (2006), who show that publicly available option volume information contains little predictive power. In regression (1), implied volatility spread (RIVOL) carries a negative and statistically significant coefficient, consistent with Bali and Hovakimian (2009).

1.5 Conclusion

Following the hypothesis that a disaster reduces the fundamental value of stock by a time-varying amount (Gabaix 2012), I construct new measures of consumption and firm-level disaster risk by using the cross-section of deviation from put-call symmetry. Stocks with high firm-level disaster risk over the previous month tend to have high subsequent returns. From the t-statistics, the predictability of firm-level disaster risk on future returns persists up to three months. After three months, the economic and statistical significance of firm-level disaster risk portfolios disappears, which indicates that the information asymmetry dissipates, on average, after three months. Holding the firm-level disaster risk portfolios longer than 4 months, there exists a U-shaped returns across different deciles.

I further show that market-level disaster risk has strong predictive power for aggregate market returns with out-of-sample. Compared with known measures of disaster risk, the market-level disaster risk measure robustly predicts aggregate market returns. In the cross-section, stocks with high loadings on past market-level disaster risk earn a monthly three-factor alpha 0.75% higher than stocks with low market-level disaster risk loadings. These finding suggests that market-level disaster risk is priced in predicting aggregate market returns and cross-sectional stock returns.

Appendices

A. Option Prices

In the following, I briefly summarize the the approach used in Gabaix (2012) to derive the option prices. During the process, we need to derive variables of stock resilience, stock prices, and then option prices.

A1. Resilience

The notion of resilience $H_{i,t}$ of stock i is introduced with the time variation in the stock's recovery rate $F_{i,t+1}$,

$$H_{i,t} = p_t E_t^D [Z_{t+1}^{-\gamma} F_{i,t+1} - 1], \tag{1.19}$$

where E_t^D (E_t^{ND}) is the expected value conditionally on a disaster (no disaster) happening at $t+1$. The stock that with high recover rate (as of high resilience) is expected to do well in a disaster. In the cross-section, stocks with high resilience are safer, and then command low risk premia.

To specify the dynamics of $H_{i,t}$, resilience $H_{i,t}$ is split into a constant part H_{i*} and a variable part $\widehat{H}_{i,t}$ as follows

$$H_{i,t} = H_{i*} + \widehat{H}_{i,t}, \tag{1.20}$$

and the variable part $\widehat{H}_{i,t}$ follows a linearity-generating (Gabaix 2009) process,

$$\widehat{H}_{i,t+1} = \frac{1 + H_{i*}}{1 + H_{i,t}} e^{-\phi_H} \widehat{H}_{i,t} + \varepsilon_{i,t+1}^H, \tag{1.21}$$

where $E_t \varepsilon_{i,t+1}^H = 0$, and $\varepsilon_{i,t+1}^H$ is uncorrelated with ε_{t+1}^D. As $H_{i,t}$ hovers around H_{i*}, $\frac{1+H_{i*}}{1+H_{i,t}}$ is close to 1. It implies that $\widehat{H}_{i,t+1} \simeq e^{-\phi_H} \widehat{H}_{i,t} + \varepsilon_{i,t+1}^H$, which means $\widehat{H}_{i,t}$ mean-reverts to 0 at a speed of ϕ_H, and innovates at each period.

A2. Stock Prices

A linear generating (LG) process is introduced by Gabaix (2009), and keeps all expressions for stocks and bonds in closed form, which is a linear function of an arbitrary number of factors. The LG class and the affine class yield have the same expression to a first order approximation.

LG processes are given by $M_t D_{i,t}$, a pricing kernel M_t times a dividend $D_{i,t}$, and $\widehat{H}_{i,t}$, an n-dimensional vector of factors. For instance, for bonds, the dividend is $D_{i,t} = 1$. By definition, a process $M_t D_{i,t}(1, \widehat{H}_{i,t})$ is LG if and only if there are constant $\alpha \in \mathbb{R}$, $\gamma, \delta \in \mathbb{R}^n$ and $\Gamma \in \mathbb{R}^{n \times n}$, such that the following relations hold at all $t \in N$,

$$E_t[\frac{M_{t+1} D_{i,t+1}}{M_t D_{i,t}}] = \alpha + \delta' \widehat{H}_{i,t}, \tag{1.22}$$

$$E_t[\frac{M_{t+1} D_{i,t+1}}{M_t D_{i,t}} \widehat{H}_{i,t+1}] = \gamma + \Gamma \widehat{H}_{i,t}. \tag{1.23}$$

where higher moments need not be specified. $M_t\, D_{i,t}(1, \widehat{H}_{i,t})$ is an LG process with generator $\Omega = [\alpha, \delta'; \gamma, \Gamma]$. Take a natural logarithm, $\omega = \ln(\Omega) = [\delta_i, -1; 0, \delta_i + \phi_H]$, where $\delta_i = \delta - h_{i*} - g_{i,D}$ as the stock's effective discount rate and $h_{i*} = \ln(1 + H_{i,*})$.

$$E_t[\frac{M_{t+1}D_{i,t+1}}{M_t D_{i,t}}\widehat{H}_{i,t+1}] = (1,0)\Omega^T(1, \widehat{H}_{i,t})'. \tag{1.24}$$

In the limit of small time intervals:

$$E_t[\frac{M_{t+1}D_{i,t+1}}{M_t D_{i,t}}] \approx \frac{1}{1-e^{-\delta_i}}(1 + \frac{e^{-\delta_i - h_{i*}}\widehat{H}_{i,t}}{1 - e^{-\delta_i - \phi_H}}) \tag{1.25}$$

To derive the stock prices $P_{i,t} = E_t[\frac{M_{t+1}D_{i,t+1}}{M_t}]$ with its resilience $H_{i,t}$,

$$P_{i,t} = E_t[\frac{M_{t+1}D_{i,t+1}}{M_t}] = D_{i,t}\,\frac{1}{1-e^{-\delta_i}}(1 + \frac{e^{-\delta_i - h_{i*}}\widehat{H}_{i,t}}{1 - e^{-\delta_i - \phi_H}}) \tag{1.26}$$

In the limit of short time period, $P_{i,t} = \frac{D_{i,t}}{\delta_i}(1 + \frac{\widehat{H}_{i,t}}{\delta_i + \phi_H})$. We can have the price to dividend ratio as $\frac{P_{i,t}}{D_{i,t}} = \frac{1}{\delta_i} + \frac{1}{\delta_i(\delta_i+\phi_H)}\widehat{H}_{i,t} = a + b\widehat{H}_{i,t}$.

The return $P_{i,t+1}/P_{i,t}$ at $t+1$ is given by

$$\begin{aligned} \frac{P_{i,t+1}}{P_{i,t}} &= \frac{P_{i,t+1}/D_{i,t+1}}{P_{i,t}/D_{i,t}}\frac{D_{i,t+1}}{D_{i,t}} \\ &= \frac{a + b\widehat{H}_{i,t+1}}{a + b\widehat{H}_{i,t}} e^{g_{i,D}}(1 + \varepsilon^D_{i,t+1}) \times \begin{cases} 1 & \text{if no disaster at } t+1 \\ F_{i,t+1} & \text{if a disaster at } t+1 \end{cases} \\ &= e^{\mu_{i,t}} \times \begin{cases} e^{\sigma_i u_{i,t+1} - \sigma_i^2/2} & \text{if no disaster at } t+1 \\ F_{i,t+1} & \text{if a disaster at } t+1 \end{cases} \end{aligned} \tag{1.27}$$

where $\mu_{i,t} = g_{i,D} + \ln[\frac{a + b\,e^{-\phi_H}\,\widehat{H}_{i,t}}{a + b\widehat{H}_{i,t}}]$, is the expected dividend growth rate of firm i, and $u_{i,t+1}$ is a standard Gaussian variable.

A3. Option Prices

We consider the price of a European one-period put on a stock i with strike K expressed as a ratio to the initial price: $OP_{i,t} = E_t[\frac{M_{t+1}}{M_t}\, max\,(0, K - \frac{P_{i,t+1}}{P_{i,t}})]$. In the following, I derive the put and call option price with or without disaster. The value of a put with strike K and a one-period maturity is $OP^{put}_{i,t}(K) = OP^{ND,put}_{i,t}(K) + OP^{D,put}_{i,t}(K)$ with $OP^{ND,put}_{i,t}$ and $OP^{D,put}_{i,t}$ corresponding to the events with no disaster

(ND) and with disaster (D), respectively,

$$\begin{aligned} OP_{i,t}^{ND,put}(K) &= (1-p_t)\,e^{-\delta}\,E_t^{ND}[(K - e^{\mu_{i,t}+\sigma_i u_{i,t+1}-\sigma_i^2/2})^+] \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,E_t^{ND}[(K\,e^{-\mu_{i,t}} - e^{\sigma_i u_{i,t+1}-\sigma_i^2/2})^+] \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,OP^{BS,put}(K\,e^{-\mu_{i,t}}, 1, \sigma_i) \end{aligned} \tag{1.28}$$

$$\begin{aligned} OP_{i,t}^{D,put}(K) &= p_t\,e^{-\delta}\,E_t^{D}[Z_{t+1}^{-\gamma}(K - e^{\mu_{i,t}}F_{i,t+1})^+] \\ &= p_t\,e^{\mu_{i,t}-\delta}\,E_t^{D}[Z_{t+1}^{-\gamma}(K\,e^{-\mu_{i,t}} - F_{i,t+1})^+] \end{aligned} \tag{1.29}$$

where $OP^{BS,put}(K,1,\sigma_i)$ is the Black-Scholes value of a put with strike K, volatility σ_i, initial price 1, maturity 1, and interest rate 0 (See the proof in Appendix B).

The value of call option $OP_{i,t}^{call}(K) = OP_{i,t}^{ND,call}(K) + OP_{i,t}^{D,call}(K)$ can be derived in the same way:

$$\begin{aligned} OP_{i,t}^{ND,call}(K) &= (1-p_t)\,e^{-\delta}\,E_t^{ND}[(e^{\mu_{i,t}+\sigma_i u_{i,t+1}-\sigma_i^2/2} - K)^+] \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,E_t^{ND}[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2} - K\,e^{-\mu_{i,t}})^+] \\ &= (1-p_t)\,e^{\mu_{i,t}-\delta}\,OP^{BS,call}(1, K\,e^{-\mu_{i,t}}, \sigma_i) \end{aligned} \tag{1.30}$$

$$\begin{aligned} OP_{i,t}^{D,call}(K) &= p_t\,e^{-\delta}\,E_t^{D}[Z_{t+1}^{-\gamma}(e^{\mu_{i,t}}F_{i,t+1} - K)^+] \\ &= p_t\,e^{\mu_{i,t}-\delta}\,E_t^{D}[Z_{t+1}^{-\gamma}(F_{i,t+1} - K\,e^{-\mu_{i,t}})^+] \end{aligned} \tag{1.31}$$

where $OP^{BS,call}(1,K,\sigma_i)$ is the Black-Scholes value of a call with strike K, volatility σ_i, initial price 1, maturity 1, and interest rate 0 (See the proof in Appendix B).

B. Proof of Eq (1.28 and 1.30)

We follow the derivations in Farhi and Gabaix (2014) and derive the discrete-time Girsanov theorm to prove Eq (1.28, 1.30). Suppose that (x, y) are jointly Gaussian distributed under the physical probability measure $\mathbb{P}$. Consider the measure $\mathbb{Q}$ defined by $d\,\mathbb{Q}/d\,\mathbb{P} = \exp(x - E[x] - Var(x)/2)$. Then, under $\mathbb{Q}$, y is Gaussian, with distribution

$$y \sim^{\mathbb{Q}} N(E^{\mathbb{P}}[y] + Cov^{\mathbb{P}}(x,y), Var^{\mathbb{P}}(y)), \tag{1.32}$$

where $E[y]$, $Cov(x,y)$, and $Var(y)$ are calculated under $\mathbb{P}$.

Proof: Calculate the moment generating function (MGF) of y under $\mathbb{Q}$ using the Radon-Nikodym derivative,

$$\begin{aligned} E^{\mathbb{Q}}[e^{m\,y}] &= E^{\mathbb{P}}[\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\,e^{my}] \\ &= E[e^{my+x}]\,e^{-E[x]-Var(x)/2} \\ &= e^{m\,E[y]+m\,Cov(y,x)+m^2\,Var(y)/2} \end{aligned} \tag{1.33}$$

We start with the equation of

$$E_t^{ND}[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2} - K\, e^{-\mu_{i,t}})^+] = E_t^{\mathbb{Q}}[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2} - K\, e^{-\mu_{i,t}})^+] \tag{1.34}$$

where $u \sim^{\mathbb{Q}} N(0,1)$, and $e^{\sigma_i u_{i,t+1}-\sigma_i^2/2}$ is log-normal with mean $-\sigma_i^2/2$ and variance σ_i^2 under measure $\mathbb{Q}$.

Define a new measure $\mathbb{Q}$ such that $x = \sigma_i u_{i,t+1} \sim^{\mathbb{Q}} N(0, \sigma_i^2)$, and the Radon-Nikodym derivative is $d\,\mathbb{Q}/d\,\mathbb{P} = \exp(x - E[x] - Var(x)/2) = \exp(\sigma_i u_{i,t+1} - \sigma_i^2/2)$. Then, under the new measure $\mathbb{Q}$, $y = \sigma_i^2/2 - \sigma_i u_{i,t+1}$ is Gaussian with the following distribution:

$$y \sim^{\mathbb{Q}} N(E^{\mathbb{P}}[y] + Cov^{\mathbb{P}}(x,y), Var^{\mathbb{P}}(y)) \tag{1.35}$$

where $E^{\mathbb{P}}[y] = E[\sigma_i^2/2 - \sigma_i u_{i,t+1}] = \sigma_i^2/2$, $Cov^{\mathbb{P}}(x,y) = Cov(\sigma_i u_{i,t+1}, \sigma_i^2/2 - \sigma_i u_{i,t+1}) = -\sigma_i^2$, and $Var^{\mathbb{P}}(y) = Var(\sigma_i^2/2 - \sigma_i u_{i,t+1}) = \sigma_i^2$. Thus, we have $y \sim^{\mathbb{Q}} N(-\sigma_i^2/2, \sigma_i^2)$.

The MGF of y under $\mathbb{Q}$ using the Radon-Nikodym derivative

$$E^{\mathbb{Q}}[e^y] = E^{\mathbb{P}}[\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\, e^y] = E[e^{\sigma_i u_{i,t+1}-\sigma_i^2/2} e^{\sigma_i^2/2-\sigma_i u_{i,t+1}}] = 1 \tag{1.36}$$

Hence,

$$\begin{aligned}
E_t^{ND}[(e^y - K\, e^{-\mu_{i,t}})^+] &= E^{\mathbb{Q}}[(e^y - K\, e^{-\mu_{i,t}})^+] \\
&= E_t^{\mathbb{P}}[\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\,(e^y - K\, e^{-\mu_{i,t}})^+] \\
&= E_t^{\mathbb{P}}[(\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\, e^y - \frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\, K\, e^{-\mu_{i,t}})^+] \\
&= E_t^{\mathbb{P}}[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2}\, e^{\sigma_i^2/2-\sigma_i u_{i,t+1}} - e^{\sigma_i u_{i,t+1}-\sigma_i^2/2}\, K\, e^{-\mu_{i,t}})^+] \\
&= E_t^{\mathbb{P}}[(1 - e^{\sigma_i u_{i,t+1}-\sigma_i^2/2}\, K\, e^{-\mu_{i,t}})^+] \\
&= OP^{BS,call}(1, K\, e^{-\mu_{i,t}}, \sigma_i)
\end{aligned} \tag{1.37}$$

where $OP^{BS,call}(1, K\, e^{-\mu_{i,t}}, \sigma_i)$ is the Black-Scholes value of a call with strike $K\, e^{-\mu_{i,t}}$, volatility σ_i, initial price 1, maturity 1, and interest rate 0.

Applying the same approach, we can derive the proof for Eq (1.28) as

$$E_t^{ND}[(K\, e^{-\mu_{i,t}} - e^y)^+] = E_t^{\mathbb{P}}[(e^{\sigma_i u_{i,t+1}-\sigma_i^2/2}\, K\, e^{-\mu_{i,t}} - 1)^+] = OP^{BS,put}(K\, e^{-\mu_{i,t}}, 1, \sigma_i) \tag{1.38}$$

where $OP^{BS,put}(K\, e^{-\mu_{i,t}}, 1, \sigma_i)$ is the Black-Scholes value of a put with strike $K\, e^{-\mu_{i,t}}$, volatility σ_i, initial price 1, maturity 1, and interest rate 0.

Bibliography

An, Byeong-Je, Andrew Ang, Turan G. Bali, and Nusret Cakici, 2014, The joint cross section of stocks and options, *Journal of Finance* 69, 2279-2337.

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1969-2012.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101-143.

Bali, Turan G., and Armen Hovakimian, 2009, Volatility spreads and expected stock returns, *Management Science* 55, 1797-1812.

Bates, D.S., 1991, The crash of '87 - was it expected ? The evidence from options markets, *Journal of Finance* 46, 1009-1044.

Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears and risk premia, *Journal of Finance* 66, 2165-2211.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481-1509.

Barro, Robert, 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823-866.

Campbell, John Y., and John Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stockmarket behavior, *Journal of Political Economy* 107, 205-251.

Cao, Charles, Zhiwu Chen, and John M. Griffin, 2005, The informational content of option volume prior to takeovers, *Journal of Business* 78, 1073-1109.

Carr, Peter, and Marc Chesney, 1996, American put call symmetry, Working paper, New York University, New York.

Carr, Peter, and Roger Lee, 2009, Put-call symmetry: extensions and applications, *Mathematical Finance* 19, 523-560.

Chang, Bo Young, Peter Christoffersen, and Kris Jacobs, 2013, Market skewness risk and the cross-section of stock returns, *Journal of Financial Economics* 107, 46-68

Chowdhry, Bhagwan, and Vikram Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies* 4, 483-511.

Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels, 2013, Ex ante skewness and expected stock returns, *Journal of Finance* 68, 85-124.

Cremers, Martijn, and David Weinbaum, 2010, Deviations from put-call parity and stock return predictability, *Journal of Financial and Quantitative Analysis* 45, 335-367.

Du, Jian, and Nikunj Kapadia, 2012, Tail and volatility indices from option prices, Working paper, University of Massachusetts, Amherst.

Easley, David, Maureen O'Hara, and P. S. Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade, *Journal of Finance* 53, 431-465.

Epstein, G. Larry, and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937-969.

Fama, Eugene F., and Kenneth R. French, 1992, Cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk and return: Some empirical tests, *Journal of Political Economy* 81, 607-636.

Farhi, Emmanuel, and Xavier Gabaix, 2014, Rare disasters and exchange rates, Working Paper.

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

Gabaix, Xavier, 2009, Linearity-generating processes: A modelling tool yielding closed forms for asset prices, Working Paper, New York Uuniversity.

Gabaix, Xavier, 2011, Disasterization: A simple way to fix the asset pricing properties of macro-economic models, *American Economic Review, Papers and Proceedings* 101, 406-409.

Goyal, Amit, and Alessio Saretto, 2009, Cross-section of option returns and volatility, *Journal of Financial Economics* 94, 310-326.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455-1508.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Kelly, Bryan, and Hao Jiang, 2014, Tail risk and asset prices, *Review of Financial Studies* 27, 2841-2871.

Longstaff, Francis, and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics* 74, 401-421.

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145-161.

McCracken, McCracken W., 2007, Asymptotics for Out of Sample Tests of Granger Causality, *Journal of Econometrics* 140, 719-752.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Rietz, Thomas, 1988, The equity risk premium: A solution, *Journal of Monetary Economics* 22, 117-131.

Santa-Clara, Pedro, and Shu Yan, 2010, Crashes, volatility, and the equity premium: Lessons from S&P 500 options, *Review of Economics and Statistics* 92, 435-451.

Schneider Paul, and Fabio Trojani, 2015, Fear trading, Working Paper.

Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao, 2010, What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis* 45, 641-662.

Figure 1.1 **Firm-level disaster risk measure**
This figure shows the quantiles of firm-level disaster risk across firms. Firm-level firm-level disaster risk is defined as deviation from OTM put and OTM call option symmetry. Firm-level disaster risk are expressed in daily terms with 25%, 50%, and 75% quantiles. The sample period is from January 1996 to August 2013.

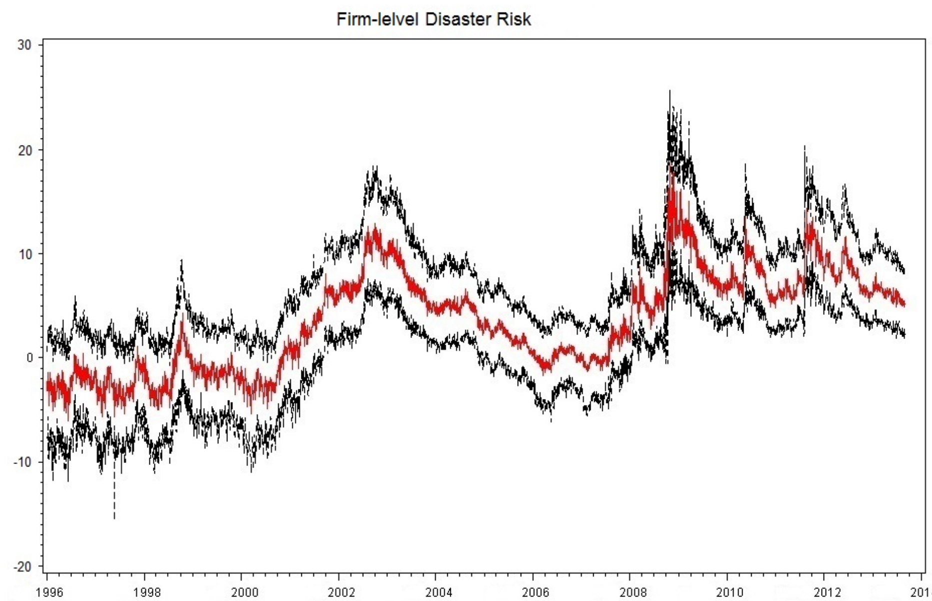


Figure 1.2 **Market-wide disaster risk and subsequent market returns**
This figure shows the market-level disaster risk (MDR). Assume there are two firm-level disaster risks: $FDR_{i,t}(k_1)$ and $FDR_{i,t}(k_2)$ for stock i at moneyness k_1 and k_2 respectively, $MDR = E_t^D[p_t\, Z_{t+1}^{-\gamma}] = \frac{FDR_{i,t}(k_1)-FDR_{i,t}(k_2)}{e^{-\delta}\,(k_1-k_2)}$. Then, I plot the average value (value weighted) of market-level disaster risk across stocks, and the realized subsequent one-month market returns. To emphasize comparison, the market-level disaster risk time series have been scaled to have same mean and variance as those of subsequent market returns time series. The sample period is from January 1996 to August 2013.

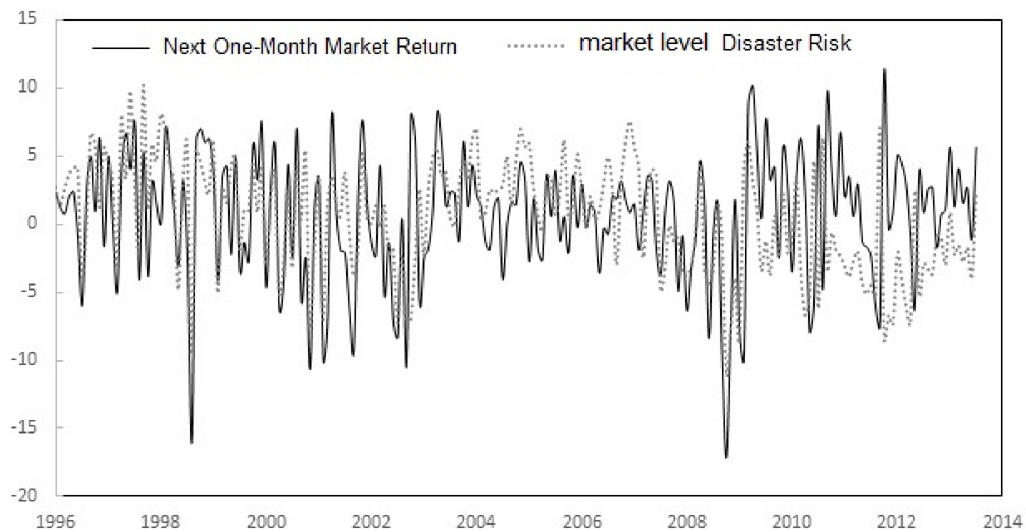


Figure 1.3 **Average firm and market level disaster risk**

This figure shows the average firm-level disaster risk (FDR) and market-level disaster risk (MDR). FDR (average stock) is a cross-sectional average of firm-level FDR, FDR (market) is derived from the S&P 500 index options, MDR (average stock) is a cross-sectional average of firm-level MDR, and MDR (market) is derived from the S&P 500 index options. The sample period is from January 1996 to August 2013.

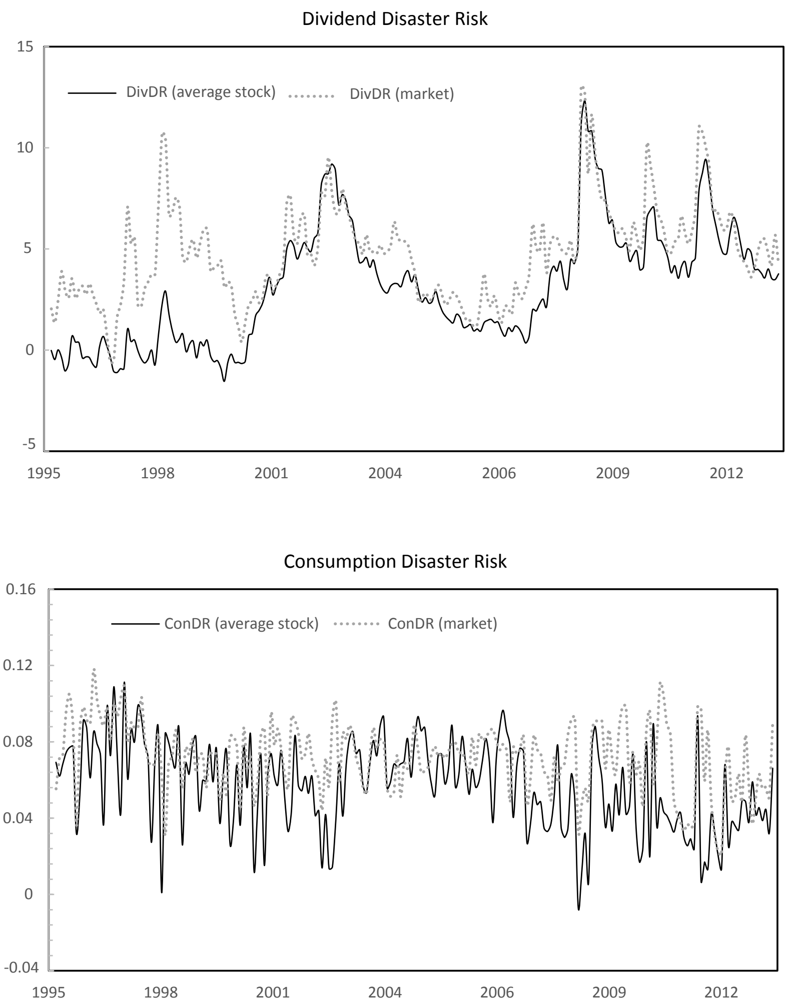


Table 1.1 **Descriptive Statistics for Firm-level Disaster Risk Measure**

This table reports the average number of stocks per month for each year from 1996 to 2013. I report the average, standard deviation and the quantiles (e.g. 1%, 25%, 50%, 75% and 99%) of the end-of-day cross-firm firm-level disaster risk.

date	# of stocks	1%	25%	50%	75%	99%	Average	Stdev
1996	1455	-8.70	-2.43	-0.87	0.12	7.48	-1.20	2.93
1997	1728	-8.70	-2.46	-0.87	0.13	6.79	-1.23	2.85
1998	1920	-8.74	-2.08	-0.56	0.50	10.64	-0.67	3.38
1999	1959	-8.53	-2.18	-0.60	0.46	8.04	-0.87	3.01
2000	1941	-9.03	-2.78	-0.83	0.36	7.59	-1.22	3.21
2001	1813	-5.80	-0.27	0.86	3.23	20.16	2.20	4.71
2002	1761	-1.51	0.88	2.81	8.30	34.10	5.99	7.78
2003	1747	-1.38	0.92	2.85	7.76	31.63	5.55	7.00
2004	1958	-2.51	0.44	1.82	5.38	24.28	3.74	5.30
2005	1990	-6.41	-0.46	0.60	2.59	24.08	1.78	4.94
2006	2165	-8.07	-1.62	-0.17	1.06	18.37	0.17	4.28
2007	2383	-8.46	-1.42	0.07	1.84	23.29	0.90	5.26
2008	2350	-2.82	1.00	3.63	9.22	34.72	6.46	7.98
2009	2261	-0.82	2.03	5.81	13.55	35.97	8.86	8.77
2010	2357	-1.08	1.71	5.22	12.78	35.36	8.30	8.56
2011	2351	-1.29	1.87	6.02	14.22	37.04	9.27	9.39
2012	2161	-1.02	1.99	6.30	14.55	36.99	9.38	9.18
2013	2165	-1.16	1.59	5.36	11.65	32.98	7.76	7.87
Average	2026	-4.78	-0.18	2.08	5.98	23.86	3.62	5.91

Table 1.2 **Market Return Predictability: Univariate Predictor Performance**

This table reports results of univariate predictor performance from monthly predictive regressions of CRSP value-weighted market index return over one-, three-, six-, and twelve-month horizons. The first row reports forecasting results based on market-level disaster risk (MDR (average stock)) time series. Next are the results from MDR (market) from S&P 500 index options, predictors studies in Goyal and Welch (2008), the investor fear index (Bollerslev and Tauchen 2011), risk-neutral skewness and kurtosis based on S&P 500 index options (Bakshi, Kapadia, and Madan 2003), and tail risk from cross-section stock returns (Kelly and Jiang 2014). For comparison, reported predictive coefficients are scaled as the percentage change in annualized expected returns resulting from a one-standard-deviation increase in each predictor variable. Newey-West t-statistics are given in parentheses, where lag length equal to the number of month in each horizon.

	One-Month			Three-Month			Six-Month			Twelve-Month		
	Coeff.	t-stat.	R^2	Coeff.	t-stat.	R^2	Coeff.	t-stat.	R^2	Coeff.	t-stat.	R^2
MDR (average stock)	1.73	4.67	2.69	1.08	4.28	2.03	1.41	3.75	6.36	1.18	2.07	4.07
MDR (market)	1.50	4.81	2.04	0.89	3.46	2.47	1.23	3.25	4.84	1.00	1.75	2.64
Book to Market	-0.35	-1.05	0.53	0.19	0.71	0.24	1.16	3.05	4.31	3.06	5.73	6.70
Default return spread	0.90	2.80	1.22	0.36	1.36	0.88	0.01	0.02	0.00	-0.25	-0.44	0.09
Default yield spread	-0.45	-1.39	0.93	-0.32	-1.20	0.69	0.14	0.37	0.07	1.16	2.05	1.98
Dividend payout ratio	-0.01	-0.01	0.00	0.17	0.64	0.20	0.71	1.83	1.59	1.39	2.46	2.83
Dividend price ratio	-0.38	-1.15	0.63	0.39	1.48	1.05	1.38	3.65	2.05	3.42	6.55	7.15
Earnings price ratio	-0.20	-0.60	0.17	0.01	0.05	0.00	-0.07	-0.19	0.02	0.23	0.41	0.08
Inflation	0.13	0.40	0.08	0.26	1.00	0.48	-0.50	-1.29	0.80	-1.23	-2.17	2.23
Long-term bond return	-1.03	-3.21	1.74	-0.51	-1.96	1.42	-0.34	-0.88	0.37	-0.35	-0.61	0.18
Long-term yield	-0.07	-0.23	0.02	-0.30	-1.15	0.64	-0.86	-2.23	2.35	-1.62	-2.88	3.85
Net equity expansion	0.51	1.56	0.58	0.68	2.60	1.57	1.58	4.24	5.99	2.68	4.93	6.51
Stock volatility	-1.74	-5.70	2.72	-1.12	-4.45	3.72	-1.03	-2.69	3.37	-0.13	-0.23	0.03
Term spread	-0.06	-0.19	0.02	-0.30	-1.13	0.61	-0.84	-2.19	2.25	-1.48	-2.62	3.21
Treasury-bill rate	-0.14	-0.42	0.08	-0.26	-1.00	0.48	-0.75	-1.95	1.81	-2.10	-3.78	4.45
RN-skewness	-0.13	-0.37	0.09	-0.38	-1.35	1.28	-0.52	-1.36	1.30	0.06	0.10	0.01
RN-kurtosis	0.50	1.36	1.29	0.42	1.48	1.53	0.55	1.42	1.42	0.19	0.29	0.06
IFI-BT	0.37	1.96	1.68	0.17	2.57	0.80	0.43	2.30	0.66	1.03	2.61	4.64
Tail-KJ	0.88	2.53	1.59	0.47	1.68	1.48	0.81	1.96	1.99	2.29	3.94	5.58

Table 1.3 **Market Return Predictability: Bivariate Predictor Performance**

This table reports results of bivariate predictor performance from monthly predictive regressions of CRSP value-weighted market index return over one-, three-, six-, and twelve-month horizons. For each horizon, the first two columns are the coefficient and t-statistic for market-level disaster risk (MDR) time series, whereas the third and fourth columns are the coefficient and t-statistic for the alternative predictors in Goyal and Welch (2008), the investor fear index (Bollerslev and Tauchen 2011), risk-neutral skewness and kurtosis based on S&P 500 index options (Bakshi, Kapadia, and Madan 2003), and tail risk from cross-section stock returns (Kelly and Jiang 2014). For comparison, reported predictive coefficients are scaled as the percentage change in annualized expected returns resulting from a one-standard-deviation increase in each predictor variable. Newey-West t-statistics are given in parentheses, where lag length equal to the number of month in each horizon.

	One-Month					Three-Month				
	MDR Coeff.	MDR t-stat.	Coeff.	t-stat.	R^2	MDR Coeff.	MDR t-stat.	Coeff.	t-stat.	R^2
Book to Market	1.80	4.59	0.22	0.68	1.14	1.27	3.80	0.58	2.21	2.57
Default return spread	1.66	4.46	0.74	2.43	1.32	1.06	3.16	0.25	0.98	2.14
Default yield spread	1.76	4.47	0.11	0.33	1.12	1.09	3.09	0.03	0.12	2.03
Dividend payout ratio	1.74	4.68	0.13	0.44	1.13	1.10	3.35	0.26	1.01	2.15
Dividend price ratio	1.74	4.52	0.02	0.08	1.12	1.24	3.83	0.67	2.63	2.78
Earnings price ratio	1.73	4.64	-0.14	-0.46	1.13	1.08	3.27	0.05	0.19	2.04
Inflation	1.73	4.64	-0.01	-0.02	1.12	1.07	3.21	0.18	0.70	2.09
Long-term bond return	1.77	4.97	-1.09	-3.68	1.56	1.10	3.39	-0.55	-2.20	2.56
Long-term yield	2.05	5.31	-0.85	-2.60	1.35	1.40	4.21	-0.83	-3.10	3.05
Net equity expansion	1.71	4.42	0.11	0.35	1.12	0.98	2.78	0.45	1.73	2.36
Stock volatility	1.30	3.21	-1.32	-4.25	1.70	0.80	2.08	-0.86	-3.30	3.18
Term spread	2.02	5.24	-0.80	-2.46	1.33	1.37	4.15	-0.80	-2.99	2.99
Treasury-bill rate	2.05	5.35	-0.87	-2.71	1.37	1.35	4.07	-0.75	-2.81	2.88
RN-skewness	2.83	6.37	0.01	0.05	2.78	1.28	3.30	-0.31	-1.18	3.20
RN-kurtosis	2.81	6.23	0.11	0.35	2.79	1.26	3.17	0.24	0.90	3.11
IFI-BT	3.36	5.25	0.58	1.79	3.43	1.85	4.84	0.37	1.63	4.62
Tail-KJ	1.92	5.00	0.62	1.93	1.57	1.21	3.47	0.31	1.15	2.74

Table 1.3. (Continues)

	Six-Month					Twelve-Month				
	MDR Coeff.	MDR t-stat.	Coeff.	t-stat.	R^2	MDR Coeff.	MDR t-stat.	Coeff.	t-stat.	R^2
Book to Market	1.97	4.21	1.78	4.71	7.73	2.37	4.40	3.80	7.06	12.11
Default return spread	1.43	2.76	-0.14	-0.36	3.21	1.22	2.13	-0.38	-0.66	2.24
Default yield spread	1.62	3.10	0.66	1.66	3.80	1.72	2.92	1.71	2.90	5.88
Dividend payout ratio	1.48	2.94	0.82	2.20	4.25	1.30	2.31	1.49	2.66	5.28
Dividend price ratio	1.83	3.97	1.80	4.90	8.06	2.08	4.01	3.90	7.52	13.13
Earnings price ratio	1.41	2.73	-0.03	-0.08	3.18	1.19	2.08	0.27	0.48	2.14
Inflation	1.46	2.88	-0.62	-1.64	3.78	1.28	2.28	-1.33	-2.37	4.63
Long-term bond return	1.43	2.78	-0.39	-1.04	3.42	1.19	2.09	-0.39	-0.69	2.26
Long-term yield	2.02	4.16	-1.62	-4.14	6.77	2.08	3.52	-2.41	-4.07	9.32
Net equity expansion	1.10	1.92	1.33	3.51	5.82	0.58	1.04	2.54	4.55	10.98
Stock volatility	1.21	2.04	-0.64	-1.61	3.76	1.27	2.11	0.28	0.46	2.14
Term spread	1.98	4.07	-1.56	-4.00	6.55	1.98	3.34	-2.20	-3.72	8.19
Treasury-bill rate	1.94	3.94	-1.45	-3.70	6.10	2.22	3.86	-2.90	-5.03	12.76
RN-skewness	1.90	3.73	-0.42	-1.18	7.45	2.78	4.07	0.21	0.34	10.59
RN-kurtosis	1.88	3.61	0.29	0.79	7.21	2.80	4.06	-0.20	-0.33	10.59
IFI-BT	2.98	5.55	1.02	1.48	11.04	3.63	5.06	1.46	2.33	14.60
Tail-KJ	1.66	3.13	0.59	1.46	5.08	1.48	2.56	2.09	3.61	10.69

Table 1.4 **Market Return Predictability: out-of-sample R^2**

This table reports the out-of-sample forecasting R^2 in percent from univariate predictive regressions of market returns on market-level disaster risk MDR (average stock), MDR (market), the alternative predictors in Goyal and Welch (2008), the investor fear index (Bollerslev and Tauchen 2011), risk-neutral skewness and kurtosis based on S&P 500 index options (Bakshi, Kapadia, and Madan 2003), and tail risk from cross-section stock returns (Kelly and Jiang 2014). Using data only through month t (beginning at t=80 to allow for sufficiently large initial estimation period), coefficient is used to forecast the $t+1$ return. The estimation window is then extended by one month to obtain a new predictive coefficient, and an out-of-sample forecast of the following months' return is constructed. This procedure is repeated until the full sample has been exhausted. I calculate the out-of-sample R^2 as $R^2 = 1 - \sum_t (r_{m,t+1} - \hat{r}_{m,t+1|t})^2 / \sum_t (r_{m,t+1} - \bar{r}_{m,t})^2$, where $\hat{r}_{m,t+1|t}$ is the out-of-sample forecast of the $t+1$ return based only on data through t, and $\bar{r}_{m,t}$ is the historical average market return through t. A negative R^2 implies that the predictor performs worse than setting forecasts equal to the historical mean. A star next to out-of-sample R^2 is based on significance of MSE-F statistic by McCracken (2004), which tests for equal MSE of the unconditional forecast and the conditional forecast. Significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively.

	One-Month	Three-Month	Six-Month	Twelve-Month
MDR (average stock)	0.67	1.34	2.80	6.86*
MDR (market)	0.76	1.53	2.26	4.03
Book to Market	-0.23	1.12	1.53	3.05
Default return spread	-0.15	-0.93	-1.44	-0.90
Default yield spread	-0.60	-2.88	-5.36	-9.87
Dividend payout ratio	-0.67	-1.96	-1.42	-2.85
Dividend price ratio	-0.98	-4.59	-10.10	-17.20
Earnings price ratio	-0.59	-1.12	-0.80	-1.60
Inflation	-0.14	-0.48	-2.25	-4.22
Long-term bond return	-0.20	-0.68	-4.79	-7.84
Long-term yield	-0.77	-2.97	-4.89	-9.78
Net equity expansion	-0.32	-0.46	0.75	3.76
Stock volatility	-0.68	-1.24	-7.04	-14.08
Term spread	-0.63	-2.24	2.45	-10.10
Treasury-bill rate	-0.20	-0.83	-0.74	-1.47
RN-skewness	-0.84	-3.31	-4.17	-7.60
RN-kurtosis	-0.37	-1.57	-3.56	-7.13
IFI-BT	0.31	0.59	1.77	4.10*
Tail-KJ	0.42	0.84	2.11	4.86*

Table 1.5 **MDR-sorted Portfolio Returns**

This table reports monthly stock return (value-weighted) statistics for portfolios formed on the basis of market-level disaster risk (MDR) beta β_i, $(E_t[R_{i,t+1}] = \mu_i + \beta_i\, MDR_t.)$. Each month, stocks are sorted into quintile portfolios based on market-level disaster risk loadings that are estimated from monthly data over the most current 36 months. Panel A reports the value- and equal-weighted average out-of-sample one-month holding period portfolio returns. Panel B report the out-of-sample twelve-month holding period portfolio returns. To consider alternative priced factors, I report alphas from regressions of portfolio returns on Fama-French three factor model, Fama-French-Carhart four factor momentum model, and Fama-French-Carhart plus the Pastor and Stambaugh (2003) traded liquidity five factor model. Newey-West t-statistics are given in parentheses, where lag length equal to the number of month in each horizon.

	Low	2	3	4	High	High-Low	t-stat.
Panel A: One-month returns							
			Value-weighted				
Average return	0.43	0.65	0.81	0.94	1.08	0.65	2.57
CAPM alpha	-0.07	0.25	0.46	0.58	0.66	0.73	2.49
FF alpha	-0.10	0.28	0.35	0.56	0.65	0.75	2.57
FF+Carhart alpha	-0.09	0.23	0.39	0.47	0.61	0.70	2.61
FF+Carhart+Liq alpha	-0.12	0.21	0.33	0.53	0.64	0.76	2.56
			Equal-weighted				
Average return	0.52	0.74	0.87	1.05	1.20	0.68	2.65
CAPM alpha	0.13	0.37	0.55	0.67	0.80	0.67	2.55
FF alpha	0.16	0.35	0.52	0.69	0.85	0.69	2.54
FF+Carhart alpha	0.12	0.33	0.48	0.72	0.79	0.67	2.42
FF+Carhart+Liq alpha	-0.04	0.29	0.45	0.74	0.81	0.85	2.46
Panel B: Twelve-month returns							
			Value-weighted				
Average return	0.46	0.65	0.83	0.94	1.11	0.65	2.58
CAPM alpha	-0.08	0.28	0.43	0.59	0.64	0.72	2.49
FF alpha	-0.11	0.26	0.37	0.54	0.68	0.79	2.61
FF+Carhart alpha	-0.02	0.29	0.36	0.51	0.65	0.67	2.65
FF+Carhart+Liq alpha	-0.17	0.23	0.35	0.57	0.73	0.90	2.61
			Equal-weighted				
Average return	0.52	0.74	0.85	1.03	1.26	0.74	2.63
CAPM alpha	0.13	0.36	0.52	0.63	0.78	0.65	2.60
FF alpha	0.10	0.38	0.57	0.66	0.83	0.73	2.61
FF+Carhart alpha	0.09	0.35	0.51	0.70	0.76	0.67	2.46
FF+Carhart+Liq alpha	-0.07	0.27	0.47	0.73	0.80	0.87	2.45

Table 1.6 **Decile Returns for Portfolios Sorted by Firm-level Disaster Risk**

Portfolio 1 (Low FDR) contains stocks with the lowest monthly firm-level disaster risk measure in the previous month and Portfolio 10 (High FDR) includes stocks with the highest monthly firm-level disaster risk measure in the previous month. I value-weight stocks in each decile portfolio and rebalance monthly. For each decile of firm-level disaster risk, the columns report the average raw returns, the CAPM, three-factor Fama-French (FF) alphas, extended four-factor (FF+Carhart) alphas, and five-factor (FF+Carhart+Liquidity) alphas (Pastor and Stambaugh 2003). The row 10-1 Diff. reports the difference in average raw and risk-adjusted returns between the High FDR and Low FDR deciles. Newey-West t-statistics are given in parentheses with twelve lags.

	Return	CAPM Alpha	FF Alpha	FF+Carhart Alpha	FF+Carhart+Liq Alpha
Low FDR	0.31	0.08	0.08	0.11	0.10
2	0.56	0.22	0.21	0.25	0.26
3	0.68	0.37	0.27	0.29	0.33
4	0.72	0.40	0.33	0.33	0.39
5	0.80	0.44	0.40	0.36	0.47
6	0.82	0.49	0.46	0.48	0.51
7	0.86	0.56	0.53	0.53	0.57
8	0.89	0.60	0.62	0.64	0.65
9	0.95	0.67	0.69	0.71	0.76
High FDR	1.03	0.77	0.80	0.86	0.90
10-1 Diff.	0.72	0.68	0.72	0.75	0.80
t-stat	(3.46)	(3.32)	(3.19)	(3.24)	(3.35)

Table 1.7 **Decile Returns for Holding FDR Portfolios for Long Term**

Portfolio 1 (Low FDR) contains stocks with the lowest monthly firm-level disaster risk measure in the previous month and Portfolio 10 (High FDR) includes stocks with the highest monthly firm-level disaster risk measure in the previous month. I value-weight stocks in each decile portfolio and rebalance monthly. For each decile of firm-level disaster risk, the columns report the average raw returns for holding these portfolios for one to twelve months. The row 10-1 Diff. reports the difference in average raw and risk-adjusted returns between the High FDR and Low FDR deciles. Newey-West t-statistics are given in parentheses, where lag length equal to the number of month in each horizon.

	1-Month	2-Month	3-Month	4-Month	5-Month	6-Month	9-Month	12-Month
Low FDR	0.31	0.52	0.57	0.75	0.78	0.79	0.83	0.81
2	0.56	0.60	0.60	0.70	0.72	0.72	0.77	0.74
3	0.68	0.60	0.62	0.64	0.64	0.67	0.69	0.71
4	0.72	0.60	0.60	0.60	0.59	0.60	0.58	0.60
5	0.80	0.66	0.64	0.65	0.63	0.62	0.66	0.64
6	0.82	0.73	0.72	0.74	0.74	0.74	0.72	0.73
7	0.86	0.83	0.81	0.79	0.79	0.79	0.80	0.79
8	0.89	0.88	0.91	0.89	0.89	0.89	0.87	0.87
9	0.95	0.95	0.94	0.91	0.91	0.92	0.90	0.87
High FDR	1.03	0.97	0.95	0.93	0.95	0.94	0.95	0.95
10-1 Diff	0.72	0.44	0.38	0.18	0.17	0.15	0.12	0.14
t-stat	(3.46)	(2.73)	(2.51)	(1.54)	(1.33)	(1.22)	(1.02)	(1.32)

Table 1.8 **Descriptive Statistics for Decile Portfolios Sorted by FDR**

This table shows that decile portfolios are formed by sorting the stocks based on firm-level disaster risk (FDR). Then, within each FDR decile, stocks are sorted into decile portfolios ranked based on the monthly FDR measure, so that $FDR1$ ($FDR10$) contains stocks with the lowest (highest) FDR. This table reports the average across months in the sample of the median values within each month of various characteristics for the stocksâĂŤone-month-ahead return (Return), innovation of FDR measure (ΔFDR), market capitalization (SIZE), book-to-market (BM), the cumulative return over the 6 months prior to portfolio formation (MOM), the return in the portfolio formation month (REV), implied volatility spread (IVOL-S), realized volatility (RVOL), realized minus implied volatility (RIVOL), innovation of implied put (call) volatility ΔPIVOL (ΔCIVOL), and the slope of volatility smirk (VOLSKEW). VOLSKEW is defined as the difference between out-of-the-money put implied volatility and the average of at-the-money call and put implied volatilities. Newey-West t-statistics are given in parentheses with twelve lags.

	FDR	ΔFDR	Return	IVOL-S	RVOL	RIVOL	VOLSKEW	ΔPIVOL	ΔCIVOL	BETA	SIZE	BM	MOM	REV
Low FDR	-3.53	-7.55	0.31	2.15	43.32	-1.60	6.21	1.31	1.06	1.24	7.71	0.67	10.47	1.85
2	-1.12	-5.49	0.56	3.57	43.56	-1.53	5.80	-0.45	0.98	1.07	7.62	0.65	10.51	1.80
3	-0.04	-4.56	0.68	4.04	43.77	-1.71	5.38	-0.62	0.33	1.13	7.57	0.67	10.60	1.89
4	0.92	-3.69	0.72	4.43	43.69	-1.53	5.16	-0.59	0.04	1.16	7.61	0.68	10.37	1.70
5	2.00	-2.68	0.80	4.71	43.30	-1.72	5.00	-0.37	-0.01	1.18	7.70	0.69	10.50	1.80
6	3.34	-1.27	0.82	5.02	42.58	-1.77	4.89	-0.76	-0.76	1.11	7.88	0.68	10.50	1.74
7	5.02	0.35	0.86	5.42	42.70	-1.59	4.69	-0.50	-0.85	1.35	8.07	0.67	10.80	1.80
8	7.25	2.61	0.89	5.98	42.14	-1.74	4.50	-0.36	-1.23	1.20	7.93	0.67	10.68	1.76
9	10.54	5.93	0.95	6.70	41.86	-1.76	4.22	0.13	-1.26	1.22	7.66	0.69	9.92	1.59
High FDR	13.62	7.10	1.03	8.46	41.96	-1.80	3.56	2.22	-1.10	1.20	7.49	0.69	10.27	1.62
10-1 Diff	17.15	14.65	0.72	6.31	-1.36	-0.21	-2.65	0.91	-2.15	-0.04	-0.22	0.02	-0.20	-0.23
t-stat	(3.19)	(3.00)	(3.46)	(3.61)	(-1.85)	(-1.99)	(-3.46)	(0.93)	(-2.44)	(-0.51)	(-1.20)	(1.64)	(-0.85)	(-2.35)

Table 1.9 **Average Firm-level Correlations**

This table reports the average firm-level cross-correlations of stocks' firm-level disaster risk (FDR), innovation of FDR (ΔFDR), market capitalization (SIZE), book-to-market (BM), the cumulative return over the 6 months prior to portfolio formation (MOM), the return in the portfolio formation month (REV), implied volatility spread (IVOL-S), realized volatility (RVOL), realized minus implied volatility (RIVOL), slope of volatility smirk (VOLSKEW), innovation of implied put (call) volatility ΔPIVOL (ΔCIVOL), ratio of put call open interest (P/C-OI), ratio of put call volumn (P/C-VL), tail risk (Tail-KJ) by Kelly and Jiang (2014), and investor fear index (IFI-BT) by Bollerslev and Todorov (2011).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
FDR (1)	1.000															
ΔFDR (2)	0.570	1.000														
IVOL-S (3)	0.424	0.200	1.000													
RVOL (4)	-0.064	-0.016	0.154	1.000												
RIVOL (5)	-0.005	-0.003	-0.001	0.609	1.000											
VOLSKEW (6)	-0.126	-0.139	-0.282	0.144	0.006	1.000										
ΔPIVOL (7)	0.004	0.009	0.188	0.365	-0.030	0.125	1.000									
ΔCIVOL (8)	-0.046	-0.077	0.030	0.370	-0.024	0.041	0.967	1.000								
SIZE (9)	-0.059	-0.074	-0.008	0.199	0.026	-0.113	-0.299	-0.314	1.000							
BM (10)	0.062	0.017	0.033	0.005	0.018	0.023	-0.010	-0.011	-0.058	1.000						
MOM (11)	-0.069	-0.012	-0.097	-0.075	-0.053	-0.090	0.003	0.003	0.015	-0.092	1.000					
REV (12)	-0.028	-0.005	-0.034	-0.059	-0.056	-0.028	-0.005	-0.007	0.001	-0.040	0.305	1.000				
P/C-OI (13)	0.019	0.005	0.014	-0.025	-0.003	0.004	-0.010	-0.012	0.025	0.005	0.018	0.018	1.000			
P/C-VL (14)	0.039	0.016	0.034	-0.022	-0.008	0.021	-0.009	-0.015	0.038	0.007	-0.010	-0.003	0.055	1.000		
Tail-KJ (15)	0.075	0.000	0.021	-0.221	-0.041	0.027	0.000	0.000	0.009	-0.011	-0.057	0.005	0.011	0.011	1.000	
IFI-BT (16)	-0.223	0.000	-0.274	-0.361	-0.198	-0.222	0.000	0.000	0.017	-0.042	0.192	0.147	0.009	-0.013	0.102	1.000

Table 1.9 Average Firm-level Correlations (Continue)

This table reports the average firm-level cross-correlations of stocks' firm-level disaster risk (FDR), and the other predictors in Goyal and Welch (2008).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
FDR (1)	1.000												
Treasury bills (2)	-0.505	1.000											
Long term yield (3)	-0.439	0.771	1.000										
Net equity expansion (4)	-0.141	0.171	0.412	1.000									
Inflation (5)	-0.090	0.120	0.114	0.031	1.000								
Long term rate of return (6)	0.024	0.012	-0.090	0.028	-0.241	1.000							
Dividend price ratio (7)	0.345	-0.577	-0.563	-0.493	-0.165	0.025	1.000						
Dividend yield (8)	0.344	-0.584	-0.571	-0.467	-0.163	-0.024	0.978	1.000					
Earning price ratio (9)	-0.023	0.027	-0.247	0.093	0.061	0.058	-0.101	-0.110	1.000				
Dividend payout ratio (10)	0.166	-0.268	-0.026	-0.290	-0.123	-0.039	0.513	0.511	-0.906	1.000			
Default yield spread (11)	0.328	-0.497	-0.416	-0.545	-0.287	0.006	0.689	0.666	-0.512	0.735	1.000		
Term spread (12)	0.383	-0.843	-0.308	0.092	-0.082	-0.094	0.387	0.389	-0.248	0.379	0.391	1.000	
Default return spread (13)	-0.173	0.254	0.426	0.116	0.259	-0.940	-0.217	-0.174	-0.137	0.026	-0.148	-0.020	1.000

Table 1.10 **Cross-sectional Equity Returns by FDR, and other Predictors**

This table shows the Fama-MacBeth (1973) regression ($R^e_{i,t+1} = \beta_{0,t} + \beta_{1,t}\, FDR_{i,t} + \beta'_{2,t}\, CONTROLS_{i,t} + \epsilon_{i,t+1}$) to examine whether firm-level disaster risk (FDR) can predict the next month's realized excess returns ($R^e_{i,t+1}$), while controlling for a collection of stock-specific control variables observable at time t for stock i. Control variables include innovation of firm-level disaster risk (ΔFDR), market beta (BETA), market capitalization (SIZE), book-to-market (BM), the cumulative return over the 6 months prior to portfolio formation (MOM), the return in the portfolio formation month (REV), implied volatility spread (IVOL-S), realized volatility (RVOL), realized minus implied volatility (RIVOL), innovation of implied put (call) volatility ΔPIVOL (ΔCIVOL), the slope of volatility smirk (VOLSKEW), Tail risk (Tail-KJ) by Kelly and Jiang (2014), and Investor Fear Index (IFI-BT) by Bollerslev and Todorov (2011). Newey-West t-statistics are given in parentheses with twelve lags.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FDR	0.158							0.186	0.212	0.190
	(3.26)							(2.97)	(3.10)	(2.74)
ΔFDR		0.027								0.032
		(2.88)								(2.63)
IVOL-S			0.043					0.039		0.031
			(3.12)					(3.00)		(2.64)
VOLSKEW				-0.096				-0.075		-0.072
				(-3.50)				(-2.69)		(-2.44)
ΔCIVOL					-0.133			-0.16		-0.152
					(-2.42)			(-2.26)		(-2.05)
ΔPIVOL					0.084			0.066		0.073
					(0.92)			(0.85)		(0.62)
Tail-KJ						0.074			0.069	0.08
						(1.98)			(1.95)	(1.75)
IFI-BT							-0.051		-0.0634	-0.056
							(-2.30)		(-2.55)	(-2.17)
BETA	-0.0076	-0.0064	-0.007	-0.0082	-0.0058	-0.0073	-0.0045	-0.0052	-0.008	-0.0038
	(-0.48)	(-0.45)	(-0.54)	(-0.23)	(-0.60)	(-0.37)	(-0.33)	(-0.40)	(-0.74)	(-0.32)
SIZE	-0.043	-0.047	-0.039	-0.05	-0.046	-0.053	-0.047	-0.034	-0.038	-0.035
	(-1.27)	(-1.20)	(-1.10)	(-1.24)	(-1.31)	(-1.36)	(-1.25)	(-1.07)	(-1.12)	(-1.05)
BM	0.064	0.08	0.077	0.072	0.076	0.065	0.083	0.057	0.064	0.066
	(1.85)	(2.01)	(1.91)	(1.73)	(1.96)	(1.88)	(1.99)	(1.71)	(1.65)	(1.70)
MOM	-0.0021	-0.0018	-0.0024	-0.0016	-0.0023	-0.002	-0.0024	-0.0026	-0.0023	-0.0024
	(-0.80)	(-0.64)	(-0.75)	(-0.54)	(-0.48)	(-0.50)	(-0.68)	(-0.73)	(-0.36)	(-0.33)
REV	-0.026	-0.023	-0.024	-0.02	-0.028	-0.025	-0.024	-0.021	-0.018	-0.019
	(-2.30)	(-2.42)	(-2.68)	(-2.55)	(-2.64)	(-2.37)	(-2.26)	(-2.31)	(-2.43)	(-2.35)
RVOL		-0.0075	-0.0114	-0.0096	-0.0103	-0.0085	-0.00132	-0.0099	-0.0086	-0.007
		(-1.90)	(-1.72)	(-1.96)	(-1.66)	(-1.74)	(-1.87)	(-2.05)	(-1.97)	(-1.71)
RIVOL	-0.338									
	(-1.96)									
P/C-OI	0.072	0.074	0.077	0.068	0.073	0.081	0.069	0.075	0.072	0.061
	(1.36)	(1.48)	(1.13)	(1.20)	(1.25)	(1.42)	(1.29)	(1.25)	(1.10)	(0.97)
P/C-VL	0.0065	0.0058	0.0074	0.007	0.0068	0.0062	0.0085	0.0072	0.0061	0.0063
	(0.34)	(0.31)	(0.35)	(0.39)	(0.36)	(0.45)	(0.41)	(0.38)	(0.27)	(0.29)
R^2	8.86%	7.72%	7.47%	7.69%	8.25%	7.21%	7.37%	9.43%	8.98%	9.60%
	(9.29)	(10.04)	(9.85)	(8.63)	(9.12)	(9.43)	(10.21)	(9.42)	(9.75)	(9.49)

Chapter 2

Higher Moments, Parameter Uncertainty, and Equity Return Predictability

2.1 Introduction

Many empirical research indicate the evidence of predictability in asset returns. However, three major asset pricing puzzles still capture the attention of macroeconomic finance: the equity premium, risk free rate and equity volatility puzzles. The equity premium puzzle refers to the failure of rational expectations equilibrium (REE) model to explain a historical difference of $\sim 6\%$ between the average return from a representative stock market portfolio and the average return from a representative portfolio of relatively safe bond of value (Mehra and Prescott 1985). The risk free rate puzzle refers to the $\sim 4\%$ discrepancy between the risk free rate that predicted by the REE Ramsey formula and actually observed (Weil 1989). The equity volatility puzzle refers to the empirical fact that actual returns on a representative stock market index have a variance some two orders of magnitude larger than the variance of consumption and dividend (LeRoy and Porter 1981, Shiller 1981, Campbell 1996). A common explanation of these three asset pricing puzzles is that markets are behaving as if the investors fear some unknown hidden randomness that isn't obviously obtained from the data. This unknown hidden randomness is linked to rare risk.

To incorporate the risk into the model, Bansal and Yaron (2004) propose a long run risk model which adds a small persistent expected growth rate component and a conditional volatility component. They find that changes in these fundamentals drive the risks and volatility in asset prices, and any adverse movements in the long-run growth components lower asset prices and concomitantly the wealth and consumption, making investors a high equity risk compensation for holding risky equity. Barro (2006), Gabaix (2012) and Wachter (2013) consider the discrete or continuous time rare disasters by introducing a downward jump to create negative skewness on normal distribution, and find low probability ($\sim 2\%$ per year) disasters explain the equity premium puzzle along with other asset market puzzles. Weitzman (2007), Bakshi and Skoulakis (2010), and Gvozdeva and Kumar (2012) study the subjective expectations on structural distribution with parameter uncertainty. Boguth and Kuehn (2014) use consumption volatility to predict future returns, generating a spread across quintile portfolios in excess of 7% annually.

During a macroeconomic disaster, aggregate consumption growth falls by a time varying amount and generates a rare risk. The investor requires a high compensation for bearing time varying risk due to the shocks and the likelihood of rare disasters. We can see that there are three important features of dividend and consumption growths: stochastic, skewness and fat-tail. The exact distribution of consumption and dividend growths is still subject to considerable debate. Rietz (1988), Longstaff and Piazzesi (2004), Barro (2006), Weitzman (2007), and Colacito, Ghysels and Meng (2012) agree the effects of macroeconomic disasters on consumption growth, GDP growth and asset returns. However, they use different distributions about these fundamental variables. Rietz simply chooses an arbitrary distribution and illustrates its impact. Longstaff and Piazzesi argue that a distribution based on U.S. experience cannot match the equity premium with modest degrees of risk aversion. Barro studies broader collections of countries, which in principle can tell us about alternative histories the United States might have experienced. In the panel of international macroeconomic data, the frequency and magnitude of disasters are significantly larger than those we have seen in U.S. history. Weitzman proposes a thickened posterior predictive left tail to represent the structural uncertainty about bad events. Colacito, Ghysels and Meng use an unconditional skew normal distribution of GDP growth to predict equity excess returns. From those studies, the distribution of consumption and dividend growths is still under debate. This paper uses a skew student's t-distribution, which can capture the tail risk and asymmetric growth prospects in asset pricing models.

Earlier studies by Barsky and DeLong (1993), Timmerman (1993), Bossaerts (1995), Cecchetti, Lam and Mark (2000), Veronesi (2000), Brennan and Xia (2001), Abel (2002), Brav and Heaton (2002), Lewellen and Shanken (2002), Weitzman (2007), Bakshi and Skoulakis (2010) and Gvozdeva and Kumar (2012) indicate that the need for Bayesian learning about structural parameters, which reduces the degree of one or another equity anomaly. However, Geweke (2001) applies a Bayesian framework to the most standard model prototypically used to analyze behavior towards risk and then demonstrates the extraordinary fragility of the existence of finite expected utility itself. Therefore, potential problems of using student's t-distribution are the representative investor's expected utility is negative infinity and then the moment generating function (MGF) is undefined (Weitzman 2007). To solve this problem, we apply higher (up to 4th) moments method to approx the exact distribution.

Asset pricing theories commonly assume a particularly strong form of knowledge by the investors: they know the true model and true parameter values. However, parameter uncertainty is intuitively important, especially in asset pricing models with numerous parameters and increasingly complex dynamics. Statistical learning and its implications for asset pricing have attracted an enormous amount of attention. A recent survey has been provided by Pastor and Veronesi (2009). One of the key implications is that Bayesian learning generates persistent and long-term changes to the agents's beliefs, which have important influence on stock valuation, risk measures, and time series predictability. Among others, Timmerman (1993, 1996) and Lewellen and Shanken (2002) show that learning may generate excess volatility and predictability in stock returns. Johannes, Korteweg, and Polson (2014) investigate sequential learning and return predictability. Johannes, Lochstoer, and Mou (2014) focus on learning about consumption dynamics. Collin-Dufresne, Johannes, and Lochstoer (2013) study parameter learning in a general equilibrium setup and its implications for asset pricing. Fulop, Li, and Yu (2014) concurrently learn about parameters and state variables.

In this paper, investors learn about consumption and dividend simultaneously (Gvozdeva and Kumar 2012), not like previous literatures focus on the learning about either consumption (e.g. Bakshi and Skoulakis 2010) or dividend (e.g. Timmermann 1993), or forces dividends and consumption to be equal or as two separate processes (Campbell 1996, Bansal and Yaron 2004). Simultaneously modeling the parameter uncertainty with respect to consumption and dividend growths can reinforce each other, and generate a sizeable risk premium. Similar evidence is shown by Bansal, Dittmar, and Lundblad (2005) that aggregate consumption risks embodied in cash flows (e.g dividends) can account for the puzzling differences in risk premia across book-to-market, momentum, and size-sorted portfolios.

We further embed the subjective beliefs into a formal equilibrium model assuming Epstein-Zin preferences. The estimated subjective beliefs reflect a high amount of uncertainty over consumption and dividend dynamics and, over time, the beliefs ÂĞfluctuate and drift substantially. Shocks to the beliefs over the long-run properties of consumption dynamics are highly volatile and strongly counter-cyclical, due to parameter uncertainty. For example, at each time t, we price a levered claim to future consumption and dividend given beliefs over parameters and states, computing quantities such as ex-ante expected returns. Then, at time $t+1$, the representative investor updates beliefs using new macroeconomic realizations at time $t+1$, and recomputes expected returns and risk free rate. From this time series of prices, we compute realized asset returns, risk free rate, volatilities, skewness, kurtosis, etc. In the learning channel, the timing of belief revisions is important. If the investors subjective beliefs change, then asset prices should change at the same time. We also find that these realistic and difficult learning problems generate subjective beliefs about consumption and dividend dynamics that differ substantially and in important ways from beliefs generated using the standard implementations of the same models, which ÂĔfix parameters at the most likely full-sample values and assumes the model is known.

The main findings of this paper can be summarized as follows: (a) Learning the skewness and fat-tail determines an increase in the average equity risk premium, which is around 50% higher than the skew-normal and student's t-distribution, and even 90% higher than the normal distribution. (b) the skewness and kurtosis have predictive power for the conditional equity premium. (c) By comparing 3 cases: (1) learning with parameter uncertainty, (2) learning without parameter uncertainty, and (3) no learning, we determine the effects of learning and parameter uncertainty on determining the expected return.

The rest of Chapter 2 is structured as follows. Section 2.2 presents the annual aggregate data in long term and its' descriptive statistics. Section 2.3 describes the consumption and dividend dynamics and the general equilibrium model (e.g. equity premium and risk free rate). Section 2.4 provides the estimation by Bayesian learning. Section 2.5 concludes.

2.2 Data

To obtain the longest possible data series on annually aggregate consumption, dividend, stock index prices and risk free rates, we take annually data on consumption, risk free interest rates, Standard and Poor's Composite Stock Price Index values, and dividend from Robert Shiller's website, starting from 1890 to 2010. We update the data using the personal consumption series from the national economic accounts of the Bureau of Economic Analysis (BEA), population estimates from the U.S. Census Bureau, and interest rate series from the Federal Reserve Board. We deflate all nominal quantities using the Consumer Price Index (CPI). Our data on consumption and dividend growth rates covers 1890-2010 while our market returns data covers 1891-2010, which will therefore be the period we simulate in the model.

Table 2.1 shows the summary statistics for these data. For the entire period 1891-2010, the average annual equity premium is 5.53%, the average risk-free rate is 1.91%, the volatility of the market return is 18.67%, the risk-free rate volatility is 1.33%, and the Sharpe ratio is 0.30. Consumption growths, dividend growths and equity returns exhibit negative skewness ($\approx$ -0.35, -0.65, and -0.065 respectively), which indicates investor's fear of occasional disasters.

2.3 Preferences

A representative investor in the economy exhibits recursive preferences as in Epstein and Zin (1989) and Weil (1989). The key feature of these preferences is that they allow agents to be risk averse in future utility in addition to future consumption. The single period utility separates risk aversion and intertemporal elasticity of substitution (IES) in the following form:

$$U_t = [(1-\beta)C_t^{\frac{1-\gamma}{\theta}} + \beta(E_t[U_{t+1}^{1-\gamma}])^{\frac{1}{\theta}}]^{\frac{\theta}{1-\gamma}} \tag{2.1}$$

where the parameter $0 < \beta < 1$ is the time discount factor, $\gamma \geq 0$ is the risk aversion coefficient, $\psi \geq 0$ is the IES, and θ is defined by $\frac{1-\gamma}{1-\frac{1}{\psi}}$. The sign of θ is determined by the magnitudes of the risk aversion and IES.

We assume the assets are traded in a frictionless market. Conditional on the information set at date t, Φ_t, the representative investor faces the following first-order condition, or the Euler's equation:

$$E[\beta^\theta G_{c,t+1}^{-\frac{\theta}{\psi}} R_{a,t+1}^{(\theta-1)} R_{i,t+1} | \Phi_t] = 1 \tag{2.2}$$

where $G_{c,t+1} = C_{t+1}/C_t$ is the aggregate gross growth rate of per-capita consumption, $R_{a,t+1}$ is the gross return on an asset that delivers aggregate consumption as its dividends each period, and $R_{i,t+1}$ is the gross returns on any asset i. Φ_t is the representative consumers' information set, which includes the observed history of aggregate consumption and dividend growth rates up to t: G_c^t and G_d^t.

As in Campbell (1996), the return to the aggregate consumption claim, $R_{a,t+1}$, is not observed in the data while the return on the dividend claim corresponds to the observed return on the market portfolio $R_{m,t+1}$. The levels of market dividends and consumption are not equal, aggregate consumption is much larger than aggregate dividends. The difference is financed by labor income. In the model, aggregate consumption and aggregate dividends are treated as two separate processes and the difference between them implicitly defines the investor's labor income process. In order to price any individual asset, we alternatively replace $R_{i,t+1}$ in the above equation with either the aggregate consumption portfolio returns $R_{a,t+1}$, or with the market portfolio returns $R_{m,t+1}$ that pay the aggregate market dividend, or with the risk free asset returns $R_{f,t+1}$ that pay one unit of consumption good as dividends every period. We characterize the average market risk premium $E(R_{m,t} - R_{f,t})$, the average risk free rate $E(R_{f,t} - 1)$, the market volatility σ_m, the volatility of the risk-free rate σ_f, and the Sharpe ratio of the equity premium $S = E(R_{m,t} - R_{f,t})/\sigma(R_{m,t} - R_{f,t})$. We will henceforth use lowercase letters to denote the logarithm of associated variables.

We derive the asset prices using the logarithm form

$$E_t[\exp(\theta \ln \beta - \frac{\theta}{\psi} g_{c,t+1} + (\theta - 1) r_{a,t+1} + r_{i,t+1})|\Phi_t] = 1 \tag{2.3}$$

where $g_{c,t+1} = \log(G_{c,t+1})$, $r_{a,t+1} = \log(R_{a,t+1})$, and $r_{i,t+1} = \log(R_{i,t+1})$. We first start by solving the special case where $r_{i,t+1} = r_{a,t+1}$ and then solve for the market return $r_{m,t+1}$ and the risk-free rate r_f.

$$E_t[\exp(\theta \ln \beta - \frac{\theta}{\psi} g_{c,t+1} + \theta r_{a,t+1})|\Phi_t] = 1 \tag{2.4}$$

To derive these solutions for the model, we use the standard approximations utilized in Campbell and Shiller (1989),

$$r_{a,t+1} = k_{c,0} + k_{c,1} z_{c,t+1} - z_{c,t} + g_{c,t+1} \tag{2.5}$$

where $z_{c,t} = \log(P_t/C_t)$ is the log price-consumption ratio, and $k_{c,0}$ and $k_{c,1}$ are approximating constants that both depend only on the average level of z_c.

$$E_t[\exp(\theta \ln\beta + \theta\, k_{c,0} + \theta(k_{c,1}z_{c,t+1} - z_{c,t}) + \theta(1 - \frac{1}{\psi})g_{c,t+1})|\Phi_t] = 1 \tag{2.6}$$

As in Equation (2.5), we can express $r_{m,t+1}$ in terms of the price-dividend ratio, $z_{m,t} = \log(P_{m,t}/D_t)$ and dividend growth rate, $g_{d,t+1} = \log(D_{t+1}/D_t)$; i.e.,

$$r_{m,t+1} = k_{m,0} + k_{m,1}z_{m,t+1} - z_{m,t} + g_{d,t+1} \tag{2.7}$$

We apply the the projection method to relate the logarithm of price to consumption ratio $z_{c,t}$ and the logarithm of price to dividend ratio $z_{m,t}$ to the growth rate $g_{c,d,t}$ in the next subsection. Then, given the distribution of $g_{c,d,t}$, we can solve $g_{c,d,t+1}$, $r_{a,t+1}$ and $r_{m,t+1}$.

$$E_t[\exp(\theta \ln\,\beta + (\theta-1)(k_{c,0} + k_{c,1}z_{c,t+1} - z_{c,t}) + k_{m,0} + k_{m,1}z_{m,t+1} - z_{m,t} + (\theta - 1 - \tfrac{\theta}{\psi})\, g_{c,t+1} + g_{d,t+1})|\Phi_t] = 1 \tag{2.8}$$

In a similar fashion, we solve for the risk-free rate:

$$r_{f,t+1} = \ln\left[(E_t[\exp(\theta \ln\beta + (\theta - 1 - \frac{\theta}{\psi})g_{c,t+1} + (\theta-1)(k_{c,0} + k_{c,1}z_{c,t+1} - z_{c,t}))|\Phi_t])^{-1}\right] \tag{2.9}$$

Projection Method

As Cochrane (2008) notes, if both returns and dividend (or consumption) growth are unforecastable, then the price to dividend (or consumption) ratio is constant. However, the price/dividend and price/consumption ratio from 1891 to 2010, demonstrates that this is not the case. Other studies (e.g. Bansal and Yaron 2004) conjecture a solution for the price/dividend and price/consumption ratio as a linear function of a state variable, which is the expected growth rate of consumption x_t. However, Beeler and Campbell (2012) argue that U.S. data do not show as much univariate persistence in consumption or dividend growth as implied by the model. Therefore, we apply the projection method (Judd 1992) to conjecture the price-to-consumption (or dividend) ratio. Projection method is widespread used in the sciences to approximate the solution function by a member of a class of parameterized functions. This makes projection methods equivalent to parameterized expectations for the asset pricing model where the integral and the solution function are the same object.

$$z_{c,t} = \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{1,kl} T_{k-1}(2\frac{g_{c,t} - g_{c,max}}{g_{c,max} - g_{c,min}} + 1) T_{l-1}(2\frac{g_{d,t} - g_{d,max}}{g_{d,max} - g_{d,min}} + 1) \tag{2.10}$$

$$z_{m,t} = \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{2,kl} T_{k-1}(2\frac{g_{c,t} - g_{c,max}}{g_{c,max} - g_{c,min}} + 1) T_{l-1}(2\frac{g_{d,t} - g_{d,max}}{g_{d,max} - g_{d,min}} + 1) \tag{2.11}$$

where $\varphi_{j,kl}$ are vectors of coefficients, and T_{k-1} is the $(k-1)^{th}$ Chebyshev polynomial (e.g. $T_0(x) = 1$, $T_1(x) = x$, and $T_k(x) = 2x\,T_{k-1}(x) - T_{k-2}(x)$ for $k \geq 2$). We use polynomials of order $N = 4$ in our computations. The order of the polynomial approximation and the convergence criteria are important. We then use this polynomial approximation in Equations (2.5,2.7,2.8,2.9) to solve the $r_{a,t}$, $r_{m,t}$ and $r_{f,t}$

2.4 Dynamics of Consumption and Dividend

2.4.1 Skew Student's t-distribution

For the growing interest in the literature on parametric families of multivariate distributions which represent some degree of departures from the multivariate normal family, we introduce skewness into a student's t-distribution. Hansen (1994) was the first to consider a skew student's t-distribution to model skewness in conditional distributions of financial returns. Since then, several skew extensions of the student's t-distribution have been proposed for financial and other applications (Fernandez and Steel 1998, Theodossiou 1998, Branco and Dey 2001, Jones and Faddy 2003, Sahu et al. 2003, Azzalini and Capitanio 2003, Bauwens and Laurent 2005, and Aas and Haff 2006). Following the extension model by Sahu et al. (2003), and letting C_t and D_t to be the aggregate consumption and dividend at time t respectively, the logarithms of consumption growth ($g_{c,t} = \log(C_t/C_{t-1})$) and dividend growth ($g_{d,t} = \log(D_t/D_{t-1})$) follow a bivariate process

$$\begin{bmatrix} g_{c,t} \\ g_{d,t} \end{bmatrix} = \begin{bmatrix} \mu_{c,t} \\ \mu_{d,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{c,t} \\ \varepsilon_{d,t} \end{bmatrix} \tag{2.12}$$

where the first part $[\mu_{c,t}, \mu_{d,t}]'$ is constant, the second part $[\varepsilon_{c,t}, \varepsilon_{d,t}]'$ are shocks, which have $i.i.d$ bivariate skew student's t-distribution $ST_2(0, \Sigma_t, \Lambda_t, v_t)$ (see more details in Sahu et al. (2003)), with scale covariance matrix $\Sigma_t = [\sigma^2_{c,t}, \sigma_{cd,t}\,;\, \sigma_{cd,t}, \sigma^2_{d,t}]$, skewness matrix $\Lambda_t = \text{Diag}\{\lambda_{1,\text{t}}, \lambda_{2,\text{t}}\}$, and degree of freedom vector $v_t = [v_{c,t}, v_{d,t}]'$.

For simplicity, we assume $\sigma_{cd,t} = 0$. The density of the multivariate skew-t distribution is given by

$$f(\mathbf{g_t}|\mu_\mathbf{t}, \Sigma_t, \Lambda_t, v_t) = 2^2 t_{2,v_t}(\mathbf{g_t}|\mu_\mathbf{t}, \Omega_t, v_t)\, T_{2,v_t+2}((\frac{v_t + Q(\mathbf{g_t})}{v_t+2})^{-1/2}\, \frac{\Lambda_t}{\Sigma_t^{1/2}}\, \frac{(\mathbf{g_t} - \mu_\mathbf{t})}{|\Omega_t|^{1/2}}) \tag{2.13}$$

where $\Omega_t = \Sigma_t + \Lambda_t^2$, $Q(g_t) = (\mathbf{g_t} - \mu_\mathbf{t})'\Omega_t^{-1}(\mathbf{g_t} - \mu_\mathbf{t})$ and

$$t_2(\mathbf{g_t}|\mu_\mathbf{t}, \Omega_t, v_t) = \frac{\Gamma((v_t+2)/2)}{|\Omega_t|(\pi\, v_t)\Gamma(v_t/2)}(1 + \frac{Q(\mathbf{g_t})}{v_t})^{-(v_t+2)/2} \tag{2.14}$$

is the density function of a 2-dimensional t-distribution with degree of freedom vector v_t, and $T_{2,v_t+2}(\cdot)$ denotes the cumulative density function (CDF) of $t_{2,v_t+2}(0, I)$.

Three interesting particular cases of the multivariate skew student's t-distribution are (i)the multivariate student's t-distribution $t(\mu_t, \Sigma_t, v_t)$ obtained when $\Lambda = 0$; (ii)multivariate skew normal distribution $SN(\mu_t, \Sigma_t, \Lambda_t)$ obtained when $v_t \to \infty$; and (iii) the multivariate normal distribution $N(\mu_t, \Omega_t)$ obtained when $\Lambda_t = 0$ and $v_t \to \infty$.

2.4.2 Moments

The recent meltdown in financial markets were hit by catastrophic events whose ex-ante probabilities were considered negligible. Only considering low distribution moments can hardly account for rare and disaster events, since this rare effect is multiplied by a very small probability. Considering high distribution moments, the rare and extremely negative effect can be raised to a higher power, making its effect substantial regardless of the small probability associated with it. Therefore, if relying on the first two distribution moments, the performance evaluation will underestimate the effects of rare disasters. There is a large body of work in asset pricing suggests that investors favor right skewness (e.g., Kraus and Litzenberger 1976, Kane 1982, and Harvey and Siddique 2000), but are averse to tail-risk and rare disasters (e.g., Barro 2006), and Chen, Joslin, and Tran 2012).

In the literature of market volatility risk, Ang et al. (2006) use the option implied volatility index (VIX) to show that innovations in aggregate market volatility carry a negative price of risk in the cross-section. Adrian and Rosenberg (2008) use a GARCH-inspired model to decompose market volatility into short and long run components and show how each of the two components affects the cross-section of asset prices. These papers use measures of stock market volatility.

There are also research of finding structural asset pricing interpretations of skewness. For example Damodaran (1985) suggests skewed distributions of asset returns are caused by investors reacting asymmetrically to good and bad company news. Chen, Hong and Stein (2001) argue that differences of opinion among investors combined with short-sale constraints generate skewed returns. Chabi-Yo, Ghysels, and Renault (2010) show that allowing for heterogeneity in investors' preferences and beliefs can give rise to additional factors related to skewness and kurtosis in the pricing of nonlinear risks, whereas Mitton and Vorkink (2007) show that allowing for heterogeneity in investors' preferences for skewness can also lead to right skewed securities having higher prices.

Preference for the fourth moment, kurtosis, has both a utility-based and an intuitive rationale. Kurtosis can be described as the degree to which, for a given variance, a distribution is weighted toward its tails (Darlington 1970). That is, kurtosis measures the probability mass in the tails of the distribution. Thus, kurtosis is distinguished from the variance, which measures the dispersion of observations from the mean, in that it captures the probability of outcomes that are highly divergent from the mean. In a multivariate distribution, random variables may also exhibit co-kurtosis. This measure captures the sensitivity to extreme states. Moments beyond the fourth are difficult to interpret intuitively and are not explicitly restricted by standard preference theory.

We derive the first four moments for the multivariate distribution $ST(\mu_t, \Sigma_t, \Lambda_t, v_t)$:

$$M_{\bar{g}}(t) = 2^m \Phi_{m,t}(\Lambda_t\, t)\, \texttt{exp}[t\mu_t \mp t'\,(\Sigma_t \mp \Lambda_t^2)t/2] \tag{2.15}$$

where $\Phi_{m,t}$ is the cumulative density function of the dimensional t-distribution with mean 0, covariance matrix identity, and degree freedom v_t. Let $\Lambda_t = \lambda_t\, I$ and $c_t = (v_t/\pi)^{1/2}\frac{\Gamma((v_t-1)/2)}{\Gamma(v_t/2)}$, the order m moment can be written as

$$\mu_{g_t} = E[g_t] = \mu_t + c_t\lambda_t,\ \ \texttt{if}\ (v_t > 1) \tag{2.16}$$

$$\sigma^2_{g_t} = E[(g_t - E[g_t])^2] = [v_t/(v_t-2)]\Omega_t - c_t^2\Lambda_t^2,\ \ \texttt{if}\ (v_t > 2) \tag{2.17}$$

$$\begin{aligned} s^3_{g_t} = {} & E[(g_t - E[g_t])^3] \\ = {} & [v_t/(v_t-2)][\Omega_t \otimes \mu_t + \mu_t \otimes \Omega_t + \mathrm{vec}(\Omega_t) \otimes \mu_t'] + \mu_t \otimes \mu_t' \otimes \mu_t \\ & + [c_t v_t/(v_t-3)][\lambda_t \otimes \Omega_t + \mathrm{vec}(\Omega_t)\lambda_t' + (\mathrm{I}_2 \otimes \lambda_t)\Omega_t - \lambda_t \otimes \lambda_t' \otimes \lambda_t] \\ & + c_t[\lambda_t \otimes \mu_t' \otimes \mu_t + \mu_t \otimes \lambda_t' \otimes \mu_t + \mu_t \otimes \mu_t' \otimes \lambda_t],\ \ \texttt{if}\ (v_t > 3) \end{aligned} \tag{2.18}$$

$$k^4_{g_t} = E[(g_t - E[g_t])^4],\ \ \texttt{if}\ (v_t > 4)\,(\texttt{see details in Appendix A}) \tag{2.19}$$

where μ_{g_t}, $\sigma^2_{g_t}$, $s^3_{g_t}$, and $k^4_{g_t}$ are the mean, volatility, skewness and kurtosis, respectively.

2.4.3 Likelihood Inference

For numerical computation of the maximum likelihood estimation (MLE), it is advantageous to make use of the expressions of the derivatives of the log-likelihood. Let $\theta_q = (\mu, \Sigma, \Lambda, v)$, and $g = (g_c, g_d)$ of size 2, the log-likelihood function for a regression model of type (Equation 2.13) and ST error terms is given

$$
\begin{aligned}
l_T(\theta_{\mathbf{q}}|\mathbf{g}) = & \quad \sum_{t=1}^{T} ln\, f(\mathbf{g_t}|\theta_{\mathbf{q}}) \\
= & \quad 2T\ln 2 + \sum_{t=1}^{T}[\ln\, t_2(\mathbf{g_t}; \mu_t, \Omega_t, v_t) \\
& \quad + \ln\, T_{2,v+2}((\frac{v_t + Q(\mathbf{g_t})}{v+2})^{-1/2}\frac{\Lambda}{\Sigma^{1/2}}\frac{(\mathbf{g_t} - \mu_{\mathbf{t}})}{|\Omega|^{1/2}})]
\end{aligned} \tag{2.20}
$$

The likelihood function conditions on the predictor variable. We can derive the derivatives of the log-likelihood respect to θ_q. Maximization of the log-likelihood function must be accomplished numerically. To improve efficiency, the derivatives of the log-likelihood can be supplied to an optimization algorithm.

2.4.4 Taylor Series Approximation

Base on the above approximation that z_c and z_m depend both on g_c and g_d, the equilibrium functions in Equations (2.6 and 2.8) can be written as $E_t[f(g_{t+1})|\Phi_t] = 1$, where $f(g_{t+1})$ is a complicated nonlinear function (see Appendix B), which depends on the posterior probability of $p(g_{t+1}|g^t)$. Using the Taylor series approximation with time-varying coefficients, $f(g_{t+1})$ can be written as

$$
f(g_{t+1}) \approx f(\bar{g^t}) + f'(\bar{g^t})(g_{t+1} - \bar{g^t}) + \tfrac{f''(\bar{g^t})}{2}(g_{t+1} - \bar{g^t})^2 + \tfrac{f'''(\bar{g^t})}{3!}(g_{t+1} - \bar{g^t})^3 + \tfrac{f''''(\bar{g^t})}{4!}(g_{t+1} - \bar{g^t})^4 + O(g_{t+1}^4) \tag{2.21}
$$

where $\bar{g^t} \approx \frac{1}{t}\sum_{\omega=1}^{t} g_\omega$ and $O(g^4)$ is the Taylor remainder. The Euler equation $E_t[f(g_{t+1})|g^t] = 1$ can be expressed as

$$
f(\bar{g^t}) + f'(\bar{g^t})E_t[(g_{t+1} - \bar{g^t})] + \tfrac{f''(\bar{g^t})}{2}E_t[(g_{t+1} - \bar{g^t})^2] + \tfrac{f'''(\bar{g^t})}{3!}E_t[(g_{t+1} - \bar{g^t})^3] + \tfrac{f''''(\bar{g^t})}{4!}E_t[(g_{t+1} - \bar{g^t})^4] = 1 \tag{2.22}
$$

where $E_t[(g_{t+1} - \bar{g^t})^m]$, $(m = 1, 2, 3, 4)$ are the first to fourth moments (e.g. μ_g, σ_g^2, s_g^3 and k_g^4) in Equations (16,17,18 and 19), respectively. For simplicity, we assume there is no cross sections between moments in the expectation.

2.5 Parameter Learning

2.5.1 Prior Distribution

Learning begins with initial beliefs on the prior distribution. We consider the problem of an investor who takes into account the predictability of returns but is uncertain about the parameters of the return model given by Equation (2.22), which is based on the investor's subjective posterior joint distribution reflecting the information contained in the historical data and the investor's prior beliefs about the parameters. The Bayesian learning approach can employ useful prior information about quantities of interest, account for estimation risk and uncertainty, and facilitate the use of fast, intuitive, and easily implementable numerical algorithms to simulate the complex economic quantities. Three main building blocks are in underlying Bayesian learning approach. The first block is to form prior beliefs, which are typically represented by a probability density function on the stochastic parameters. The prior density can reflect information about events, historical data, and asset pricing theories. The second block is to formulate the law of motion governing the evolution of asset returns. The third block is to recover the predictive distribution of future asset returns, analytically or numerically, incorporating prior information, as well as risk and uncertainty. The predictive distribution, which integrates out the parameter space, characterizes the entire uncertainty about future asset returns. The Bayesian optimal rule is obtained by maximizing the expected utility with respect to the predictive distribution.

One of the key implications of learning is that the representative investor's beliefs are non-stationary. For example, the investor may gradually learn that one model provides a better fit for the data than an other model or that a parameter value is higher or lower than the previously thought, both of which generate nonstationarity in beliefs. One easy way to understand this is to note that the posterior mean of a parameter, $E[\theta_q|g^t]$, where $g^t = (g_1, ..., g_t)$ denote the observed history of growth rate at time t, is trivially a martingale. Then, revisions in beliefs represent permanent and nonstationary shocks, which is an important implication in asset pricing theories. Nonstationary dynamics can generate a quantitatively gap between ex-post outcomes and ex-ante beliefs, which provides an alternative explanation for asset pricing quantities such as the observed equity premium or excess return predictability. Thus, learning is difficult, and it is harder to learn the parameters governing the state dynamics when states (e.g. recessions or depressions) are unobserved.

We assume that the representative investor does not know the location vector μ, scale covariance matrix Σ, skewness matrix Λ and degree of freedom vector v. However, the investor can learn about these parameters by observing the realized values of the growth rates over time. Based on a three-stage hierarchical specification, the skew student's t-distribution in Equation (2.13) can be derived as a Gaussian mixture model.

$$g|(Z,\tau) \sim N_2(\mu + \Lambda Z, \Sigma/\tau) \tag{2.23}$$

$$Z|\tau \sim TN_2(0, I_2/\tau) \tag{2.24}$$

$$\tau \sim \Gamma(v/2, v/2) \tag{2.25}$$

It's probability density function is given by:

$$p(g; Z, \tau) \sim \int\int p(g|Z,\tau)p(Z|\tau)p(\tau|v)d\tau\, dZ \qquad (2.26)$$

When the prior information is not available, a convenient strategy of avoiding improper posterior distribution is to use diffuse proper priors. The prior distributions are adopted as follows:

$$\mu \sim N_2(\mu_0, \kappa^{-1}), \;\; B \sim W_2(2B_0, (2H)^{-1}), \;\; \log(1/v) \sim U_2(-10, 10) \qquad (2.27)$$

$$\Sigma^{-1}|B \sim W_2(2a_0, (2B)^{-1}), \;\; \lambda \sim N_2(0, \Gamma) \qquad (2.28)$$

where $(\mu_0, \kappa, a_0, B_0, H, \Gamma)$ are fixed as appropriate quantities to yield the proper posterior distributions. We choose the initial values: $\mu_0 = (0.02, 0.01)$, $\kappa = (0.001, 0; 0, 0.01)$, $a_0 = (0.05, 0.03; 0.03, 0.08)$, $B_0 = (0.2, 0; 0, 0.5)$, $H = (0.07, 0.08; 0.08, 0.04)$ and $\Gamma = 4$. Thus, the joint prior density function of $\theta_q = (\mu, \Sigma, \Lambda, v)$ and B is

$$\begin{aligned}\pi(\theta_q, B) \propto\; & |B|^{\mu_0+(2B_0-2-1)/2}|\Sigma|^{(2\mu_0-2-1)/2}\exp\{-\frac{1}{2}(\mu-\mu_0)'\kappa(\mu-\mu_0)\} \\ & \times \exp\{-tr\,((\Sigma^{-1}+H)B) - \frac{1}{2}\lambda'\Gamma^{-1}\lambda\}J_v \qquad (2.29)\end{aligned}$$

where $J_v = v^{-1}$ $(0 < v < \infty)$ is the Jacobian of transformation $\log(1/v)$ to v.

2.5.2 Posterior Distribution

We apply the learning process of a general equilibrium model by embedding the subjective beliefs. The investor's expected utility is a function of the economic state as summarized by the posterior distribution. The investor updates his prior beliefs to form the posterior distribution upon seeing the data. The posterior of the parameters conditional on the return predictability, and the posterior probability of that returns are predictable. The joint posterior density can be obtained by

$$p(\theta_q, B, \tau, Z, g_{t+1}|g^t) \propto \pi(\theta_q, B)p(g_{t+1}|g^t, \tau, Z, \theta_q)p(\tau|g^t, Z, \theta_q)p(Z|g^t, \theta_q)p(\theta_q|g^t) \quad (2.30)$$

This joint distribution is the posterior in our model. The joint learning problem is a difficult and high-dimensional problem, as posterior beliefs depend in a complicated and non-analytic manner on past data and vary substantially over time. Characterizing this type of learning is difficult for a number of reasons, most notably the presence of confounding effects. Confounding occurs when uncertainty about one quantity makes learning about other quantities more difficult. We can use the Gibbs sampler to sample from the joint distribution if we knew the full conditional distributions for each parameter, which is conditional on the known information and all the other parameters. The full conditional posterior densities are described as follows:

$$\begin{aligned}
P(g_{t+1}|g^t, \tau_t, Z_t, \theta_{q,t}) &\sim \quad N_2(\mu_t, \tau_t^{-1}\Sigma_t) \\
p(\tau_t|g^t, Z_t, \theta_{q,t}) &\sim \quad \Gamma(\frac{2 \times 2 + v_t + 1}{2}, \frac{(Z_t - \Pi_t)'\Delta_t(Z_t - \Pi_t) + Q_t + v_t}{2}) \\
p(Z_t|g^t, \theta_{q,t}) &\sim \quad Tt_2(\Pi_t, \frac{v_t + Q_t}{v_t + 1}\Delta_t, v_t + 1) \\
p(\mu_t|\cdots) &\sim \quad N_2(\mu^*, \kappa^*) \\
p(B|\cdots) &\sim \quad W_2(2Z^*, (2H^*)^{-1}) \\
p(\Sigma_t^{-1}|\cdots) &\sim \quad W_2(2a_0 + t, B^{*-1}) \\
p(\lambda_t|\cdots) &\sim \quad N_2(\delta^*, \Gamma^*) \\
p(v_t|\cdots) &\propto \quad [\frac{(v_t/2)^{v_t/2}}{\Gamma(v_t/2)}]^t \tau_t^{v_t t/2} \exp(-\frac{v_t t}{2}\tau_t) J_{v_t}
\end{aligned} \quad (2.31)$$

where $\Delta_t = (I_2 + \Lambda_t'\Sigma_t^{-1}\Lambda_t)^{-1}$, $\Pi_t = \Lambda_t\Omega_t^{-1}(g^t - \mu_t)$, $\kappa^* = (\tau\Sigma^{-1} + \kappa)^{-1}$, $\mu^* = \kappa^*(\Sigma\tau(g^t - \Lambda Z) + \kappa\mu_0)$, $Z^* = a_0 + Z_t$, $H^* = H + \Sigma^{-1}$, $B^* = 2B + \tau(g^t - \mu_t - \Lambda\, Z_t)(g^t - \mu_t - \Lambda\, Z_t)'$, $\Gamma^* = (\Gamma^{-1} + \Sigma^{-1}\tau\, ZZ')^{-1}$, and $\delta^* = \Gamma^*(\Sigma^{-1}\tau\, Z(g^t - \mu_t)')$.

E-M steps by Gibbs sampler simulation are presented as follows: (1)Generate τ from $\Gamma(\frac{2\times 2+v_t+1}{2}, \frac{(Z_t-\Pi_t)'\Delta_t(Z_t-\Pi_t)+Q_t+v_t}{2})$; (2)Generate Z_t from $Tt_2(\Pi_t, \frac{v_t+Q_t}{v_t+1}\Delta_t, v_t + 1)$, (3) Generate μ_t from $N_2(\mu^*, \kappa^*)$, (4) Generate B_t from $W_2(2Z^*, (2H^*)^{-1})$, (5) Generate Σ_t from $W_2(2a_0 + t, B^{*-1})$, (6) Generate λ_t from $N_2(\delta^*, \Gamma^*)$, (7) Generate v_t from $[\frac{(v_t/2)^{v_t/2}}{\Gamma(v_t/2)}]^t \tau_t^{v_t t/2} \exp(-\frac{v_t t}{2}\tau_t) J_{v_t}$.

Finally, we obtain the investor's posterior distribution of the data on next period by integrating the conditional (e.g. on the consumption and dividend growth and unknown parameters) distribution as Equation (2.31). The likelihoods conditional on parameters are integrated over the prior distribution of the parameters.

2.5.3 Calibration and Estimation

In Equation (2.22), there are 48 parameters (e.g. β, γ, ψ, $k_{c,0}$ $k_{c,1}$, $k_{m,0}$, $k_{m,1}$, $\varphi_{2,4*4}$, μ, Σ, Λ, and v). We start from $t = (48/2) + 1 = 25th$ observations and forecast one period ahead. Based on the first 25 data (from 1891 to 1915), we have the baseline calibration for these 48 parameters and present the main parameters in Table 2.2, where risk aversion $\gamma = 8$, intertemporal elasticity of substitution (IES) $\psi = 1.2$, discount factor $\beta = 0.988$, consumption growth volatility $\Sigma_{1,1} = 0.0015$, dividend growth volatility $\Sigma_{2,2} = 0.013$, correlation between consumption growth and dividend growth $\rho = 0.35$, skewness for consumption growth $\lambda_c = -0.32$, skewness for dividend growth $\lambda_d = -0.68$, the degree freedom for consumption growth $v_c = 9$ and the degree freedom for dividend growth $v_d = 11$. The value of discount factor indicates a reasonably patient representative investor, and is consistent with the business cycle and asset pricing literatures. The value of risk aversion is below 10, which is consistent with Mehra and Prescott (1985), Kocherlakota (1996), and Bansal and Yaron (2004). However, Brandt et al (2004) and Ljungqvist and Sargent (2004) argue for lower values than 4. The IES parameter in our paper is slightly larger than 1, which is consistent with Bansal and Yaron (2004).

Table 2.3 shows the estimated moments of consumption and dividend growths using skew student's t-distribution. This table reports the unconditional mean, volatility, skewness, and kurtosis for real U.S log-consumption growth (in Panel A) and log-dividend growth (in Panel B) computed using annual data from 1915 to 2010, and using the model of skew student's t-distribution (simulating the model 1000 times with sample size 96 years). Our estimated moments are consistent with the data. After solving Equation (2.22), we obtain the estimated values for the parameters in the skew student's t-distribution, and the expected returns $E_t(r_{a,t+1})$, $E_t(r_{m,t+1})$, and $E_t(r_{f,t+1})$ in Equations (2.5, 2.7 and 2.9).

2.5.4 Effects by Skewness and Fat-Tail

We investigate the effects by skewness and fat tail by comparing the results from the benchmark model (skew student's t-distribution) with those from skew normal distribution ($v \to \infty$), student's t-distribution ($\lambda = 0$), and normal distribution ($v \to \infty$ and $\lambda = 0$). Several results ought to be noticed in Table 2.4. First of all, the introduction of skewness and fat-tail determine an increase in the average equity risk premium, which is around 50% higher than the skew normal or student's t-distribution, and even 90% higher than that in the absence of skewness and fat-tail dynamics. Second, the average risk free rate seems to be almost unaffected by the introduction of skewness and fat-tail dynamics. We also find that the result derived from the skew student's t is more consistent to the data.

2.5.5 Model Comparison and Parameter Learning

We compare four models with different distributions of the consumption and dividend dynamics: (1)the multivariate normal distribution $N(\mu_t, \Omega_t)$ obtained when $\Lambda_t = 0$ and $v_t \to \infty$; (2) the multivariate student's t-distribution $t(\mu_t, \Sigma_t, v_t)$ obtained when $\Lambda = 0$; (3) the multivariate skew normal distribution $SN(\mu_t, \Sigma_t, \Lambda_t)$ obtained when $v_t \to \infty$; and (4)the multivariate skew student's t-distribution.

To get a better sense of the role of learning about parameter uncertainty on return predictability, we compare three different return problems corresponding to three different subjective data generating processes. (1) Learning with parameter uncertainty. It is important to take the uncertainty (or estimation risk) in the estimation into account. A natural way to do this is to use the Bayesian concept of a posterior distribution $p(\theta_q|g^t)$, which summarizes the uncertainty about the parameters given the data observed so far. Integrating over this distribution, we obtain the predictive distribution for dividend and consumption growths. This distribution is conditioned only on the sample observed, and not on any fixed $\theta_q = (\mu, \Sigma, \Lambda, v)$. With learning, the conditional joint posterior distribution of growths changes from one period to the next, as new information is incorporated into the investor's beliefs each period. (2) Learning without parameter uncertainty (anticipated utility). We estimate the parameters $\theta_q = (\mu, \Sigma, \Lambda, v)$ by incorporating the skew student's t-distribution into the data observed so far. We then have $\hat{\theta}_q$, which are known and fixed at their estimated values at each period. The Euler function can be written as $E_t[f(g_{t+1})\, p(g_{t+1}|g^t, \hat{\theta}_q)] = 1$. (3) No learning where we assume the representative investor know the true parameter values, which are estimated from the whole sample period. Figure 2.1 shows the estimated sequential consumption growths' conditional return, volatility, skewness, and kurtosis, respectively. From these results, the introduction of learning and parameter uncertainty provide better explanations for equity premium puzzles than those neither learning nor parameter uncertainty.

2.5.6 Predictive Power of Moments on Actual Equity Premium

In order to study the predictive power of higher moments on actual equity premium, we regress the actual equity premium on the expected mean, variance, skewness, and kurtosis of consumption growth.

$$r_{ep,t+1} = \alpha_0+\alpha_1 E_t[(g_{t+1}-\bar{g^t})]+\alpha_2 E_t[(g_{t+1}-\bar{g^t})^2]+\alpha_3 E_t[(g_{t+1}-\bar{g^t})^3]+\alpha_4 E_t[(g_{t+1}-\bar{g^t})^4]+\epsilon \tag{2.32}$$

where the real equity premium $r_{ep,t+1} = r_{m,t+1} - r_{f,t+1}$. $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, and α_4 are the coefficients for the intercept, mean, variance, skewness and kurtosis. From the R^2 in Table 2.5, we see that under the almost- normal case, both the mean-variance criterion and the one adding the third and fourth moments provide poor predictability for the future equity premium. Under large departure from normality case, the mean-variance criterion still fail to predict future equity premium. But considering the third and fourth moments, the optimization process provides a better predictability for future equity premium. We also find that the coefficients of first and third moments have negative signs and the second and fourth moments have positive signs in the regression. The possible explanation is that better average forecast and increased upside potential will decrease the future equity premium. Skewness has a strong negative relation with subsequent returns, which indicates investors prefer positive skewness. In contrast, more uncertain growth (volatility) and fatter tail (kurtosis) will require an increase in future equity premium. We further consider different combinations of moments in Table 2.6. By comparing the R^2 for different combinations of moments, we can see that the third and fourth moments have higher predictive power about equity premium than that of second moment.

2.6 Conclusion

We incorporate skewness into a student's t-distribution to model the dynamics of consumption and dividend growth, and this setting can yield reasonable equity premium, risk-free rate, and excess volatility. Higher moments method is used to solve the undefined MGF of student's t-distribution. We further consider the parameter uncertainty and use Bayesian learning to update investor's beliefs. We find that (1) the introduction of skewness and fat-tail determine an increase in the average equity risk premium, which is around 50% higher than the skew normal or student's t-distribution, and even 90% higher than that in the absence of skewness and fat-tail dynamics, (2) the average risk free rate seems to be unaffected by the introduction of skewness and fat-tail dynamics, (3) the coefficients of first and third moments have negative signs, while the second and fourth moments have the positive sign in the regression, (4) the skewness and kurtosis have significant predictive power about equity premium, and (5) introduction of learning and parameter uncertainty provide better explanations for equity premium puzzles.

Appendix A: Kurtosis of Growth Dynamics

We derive the 4th moment (kurtosis) for the growth dynamics:

$$
\begin{aligned}
k_g^4 = \quad & E[(g - E[g])^4] \\
= \quad & [v/(v-2)][\Omega \otimes \mu \otimes \mu' + \mu \otimes \Omega \otimes \mu' + \mathrm{vec}(\Omega) \otimes \mu' \otimes \mu' + \mu' \otimes \Omega \otimes \mu \\
& + \mu \otimes \mu \otimes \mathrm{vec}(\Omega)' + \mu \otimes \mu' \otimes \Omega] + \mu \otimes \mu' \otimes \mu \otimes \mu' \\
& + [v^2/(v-2)/(v-4)][(I_4 + K_4)(\Omega \otimes \Omega) + \mathrm{vec}(\Omega)\mathrm{vec}(\Omega)'] \\
& + c[\lambda \otimes \mu' \otimes \mu \otimes \mu' + \mu \otimes \lambda' \otimes \mu \otimes \mu' \mu \otimes \mu' \otimes \lambda \otimes \mu' + \mu \otimes \mu' \otimes \mu \otimes \lambda'] \\
& + [cv(v-3)][\lambda \otimes \Omega \otimes \mu' + \mathrm{vec}(\Omega) \otimes \lambda' \otimes \mu' + \mathrm{I}_2 \otimes \lambda' \otimes \Omega \otimes \mu \\
& + \lambda' \otimes \Omega \otimes \mu + \lambda \otimes \mathrm{vec}(\Omega)' \otimes \mu + \Omega \otimes (\mathrm{I}_2 \otimes \lambda') \otimes \mu + \mu' \otimes \lambda \otimes \Omega \\
& + \mu' \otimes (\mathrm{vec}(\Omega) \otimes \lambda') + \mu' \otimes (\mathrm{I}_2 \otimes \lambda) \otimes \Omega + \mu \otimes \lambda' \otimes \Omega + \mu \otimes \lambda \otimes \mathrm{vec}(\Omega)' \\
& + \mu \otimes (\Omega(I_2 \otimes \lambda')) - \lambda \otimes \lambda' \otimes \lambda \otimes \mu' - \lambda' \otimes \lambda \otimes \lambda' \otimes \mu - \mu' \otimes \lambda \otimes \lambda' \otimes \lambda \\
& - \mu \otimes \lambda' \otimes \lambda \otimes \lambda']. \ \texttt{if}\ (v > 4)
\end{aligned}
$$

Appendix B: Nonlinear Function $f(g_{t+1})$

We derive the nonlinear function form of $f(g_{t+1})$ corresponding to $E_t[f(g_{t+1})|\Phi_t] = 1$. With respect to Equations (2.6 and 2.8), we have

$$\ln\, f_1(g_{t+1}) = \theta\, \ln\beta + \theta\, k_{c,0} + \theta(k_{c,1} z_{c,t+1} - z_{c,t}) + \theta(1 - \frac{1}{\psi}) g_{c,t+1}$$

$$
\begin{aligned}
\ln\, f_2(g_{t+1}) = \quad & \theta \ln\, \beta + (\theta - 1)(k_{c,0} + k_{c,1} z_{c,t+1} - z_{c,t}) + k_{m,0} + k_{m,1} z_{m,t+1} - z_{m,t} \\
& + (\theta - 1 - \frac{\theta}{\psi})\, g_{c,t+1} + g_{d,t+1}
\end{aligned}
$$

Assuming that the consumption/price and dividend/price can be approximated as a class of parameterized function by projection methods in Equations (2.10 and 2.11), we can derive the nonlinear functions $\ln f_1(g_{t+1})$ and $\ln f_2(g_{t+1})$ as following:

$$\begin{aligned}
\ln f_1(g_{t+1}) = \quad & \theta \ln\beta + \theta\, k_{c,0} + \theta(1-\frac{1}{\psi})g_{c,t+1} \\
& +\theta\, k_{c,1} \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{1,kl} T_{k-1}(2\frac{g_{c,t+1}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t+1}-g_{d,max}}{g_{d,max}-g_{d,min}}+1) \\
& -\theta \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{1,kl} T_{k-1}(2\frac{g_{c,t}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t}-g_{d,max}}{g_{d,max}-g_{d,min}}+1)
\end{aligned}$$

$$\begin{aligned}
\ln f_2(g_{t+1}) = \quad & \theta \ln\beta + (\theta-1)k_{c,0} + k_{m,0} + (\theta - 1 - \frac{\theta}{\psi})\, g_{c,t+1} + g_{d,t+1} \\
& +(\theta-1)k_{c,1} \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{1,kl} T_{k-1}(2\frac{g_{c,t+1}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t+1}-g_{d,max}}{g_{d,max}-g_{d,min}}+1) \\
& -(\theta-1) \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{1,kl} T_{k-1}(2\frac{g_{c,t}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t}-g_{d,max}}{g_{d,max}-g_{d,min}}+1) \\
& +k_{m,1} \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{2,kl} T_{k-1}(2\frac{g_{c,t+1}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t+1}-g_{d,max}}{g_{d,max}-g_{d,min}}+1) \\
& - \sum_{k=1}^{N}\sum_{l=1}^{N} \varphi_{2,kl} T_{k-1}(2\frac{g_{c,t}-g_{c,max}}{g_{c,max}-g_{c,min}}+1) T_{l-1}(2\frac{g_{d,t}-g_{d,max}}{g_{d,max}-g_{d,min}}+1)
\end{aligned}$$

Bibliography

Aas, K., and I.H. Haff, 2006, The generalized hyperbolic skew Student's t-distribution, *Journal of Financial Econometrics*, 4(2), 275–309.

Abel, A.B., 2002, An Exploration of the effects of pessimism and doubt on asset returns, *Journal of Economic Dynamics and Control*, 26, 1075–1092.

Azzalini, A., and A. Capitanio, 2003, Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society*, B65, 367–389.

Bakshi, G., and G. Skoulakis, 2010, Do subjective expectations explain asset pricing puzzles, *Journal of Financial Economics*, 98, 117–140.

Bansal, R., D. A. Hsieh, and S. Viswanathan, 1993, A new approach to international arbitrage pricing, *Journal of Finance*, 48, 1719–1747.

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

Bansal, R., R.F. Dittmar, and L. Christian, 2005, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance*, 60, 1639–1672.

Barro, R.J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics*, 121, 823–867.

Barsky, R.B., and J.B. DeLong, 1993, Why does the stock market fluctuate, *Quarterly Journal of Economics*, 108, 291–312

Bauwens, L., and S. Laurent, 2005, A new class of multivariate sk ew densities, with application to generalized autoregressive conditional heteroscedasticity models, *Journal of Business and Economic Statistics*, 23, 346–354

Beeler, J., and J.Y. Campbell, 2012, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, *Critical Finance Review*, Vol. 1 : No 1, 141–182.

Bossaerts, P., 1995. The econometrics of learning in financial markets, *Econometric Theory*, 11, 151–189

Branco, M.D., and D.K. Dey, 2001, A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, 79, 99–113

Brandt, M., Q. Zeng, and L. Zhang, 2004, Equilibrium stock return dynamics under alternative rules of learning about hidden states, *Journal of Economic Dynamics and Control*, 28, 1925–1954.

Brav, A., and J.B. Heaton, 2002, Competing theories of financial anomalies, *Review of Financial Studies*, 15(2), 575–606

Brennan, M.J., and Y. Xia, 2001, Stock return volatility and equity premium, *Journal of Monetary Economics*, 47, 249–283

Campbell, J.Y., 1996, Understanding risk and return, *Journal of Political Economy*, 104, 298–345.

Campbell, J.Y., and J. H. Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy*, 107, 205

Campbell, J.Y., and R. J. Shiller, 1989, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, 195–228.

Cecchetti, S., P.S. Lam, and N. Mark, 2000, Asset pricing with distorted beliefs: are equity returns too good to be true, *American Economic Review*, 90, 787–805

Chen, H., S. Joslin, and N. Tran, 2012, Rare disasters and risk sharing with heterogeneous beliefs, *Review of Financial Studies*, 25(7), 2189–2224.

Cochrane, J.H., 2008, The dog that did not bark: a defense of return predictability, *Review of Financial Studies*, 21, 1533–1575.

Colacito, Ric., E. Ghysels, and J. Meng, 2012, Skewness in expected macro fundamentals and the predictability of equity returns: evidence and theory, working paper.

Darlington, R.B., 1970, Is kurtosis really peakedness, *The American Statistician*, 24, 2, 19–22.

Epstein, L., and S. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica*, 57, 937–969.

Fernandez, C., and M.F.J. Steel, 1998, On Bayesian modelling of fat tails and skewness, *Journal of the American Statistical Association*, 93, 359–371.

Gabaix, X., 2012, Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance, *The Quarterly Journal of Economics*, 127, 645

Geweke, J., 2001, A note on some limitations of CRRA utility, *Economics Letters*, 71, 341–345.

Gvozdeva, E., and P. Kumar, 2012, Unknown consumption and financial risk and asset pricing puzzles, working paper.

Hansen, B.E., 1994, Autoregressive conditional density estimation, *International Economics Review*, 35, 3, 705–730.

Harvey, C.R., and A. Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance*, 55, 1263–1295.

Jones, M.C., and M.J. Faddy, 2003, A skew extension of the t-distribution with applications, *Journal of the Royal Statistical Society*, Series B, 65, 159–174.

Judd, K., 1992, Projection methods for solving aggregate growth models, *Journal of Economic Theory*, 58, 410–452.

Kane, A., 1982, Skewness preference and portfolio choice, *Journal of Financial and Quantitative Analysis*, 17, 15–25.

Kocherlakota, N., 1996, The equity premium: it's still a puzzle, *Journal of Economic Literature*, 34, 42–71.

Kraus, A., and R.H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance*, 31, 1085–1100.

LeRoy, S., and R. Porter, 1981, The present-value relation: tests based on implied variance bounds, *Econometrica*, 49, 555–574.

Lewellen, J., and J. Shanken, 2002, Learning, asset pricing tests, and market efficiency, *Journal of Finance*, 57, 1113–1146.

Ljungqvist, L., and T. Sargent, 2004, Recursive Macroeconomic Theory, Cambridge, MA: MIT Press.

Longstaff, F.A., and M. Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics*, 74, 401–421.

Lucas, R., 1978, Asset prices in an exchange economy, *Econometrica*, 46, 1429–1445.

Mehra, R., 2008, The equity premium puzzle: a review, *Foundations and Trends in Finance*, 2, 1–81.

Mehra, R., and E. Prescott, 1985, The equity premium: a puzzle, *Journal of Monetary Economics*, 15, 145–161.

Rietz, T., 1988, The equity risk premium: a solution, *Journal of Monetary Economics*, 22, 117–131.

Sahu, S.K., D.K. Dey, and M.D. Branco, 2003, A new class of multivariate skew distributions with applications to Bayesian regression models, *The Canadian Journal of Statistics*, 31, 129–150.

Shiller, R., 1981, Do stock prices move too much to be justified by subsequent changes in dividends, *American Economic Review*, 71, 421–436.

Theodossiou, P., 1998, Financial data and the skewed generalized t-distribution, *Management Science*, 44 (12-1), 1650–1661.

Timmermann, A., 1993, How learning in financial markets generates excess volatility and predictability in stock prices, *Quarterly Journal of Economics*, 108, 1135–1145.

Veronesi, P., 2000, How does information quality affect stock returns, *Journal of Finance*, 55, 807–837.

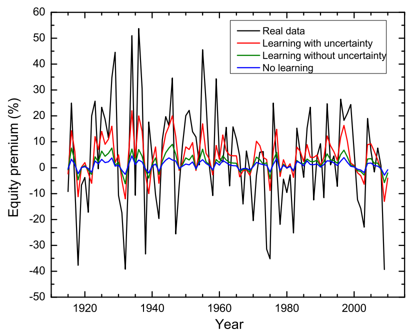
Wachter, J.A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility, *Journal of Finance*, forthcoming.

Weil, P., 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics*, 24, 401–421.

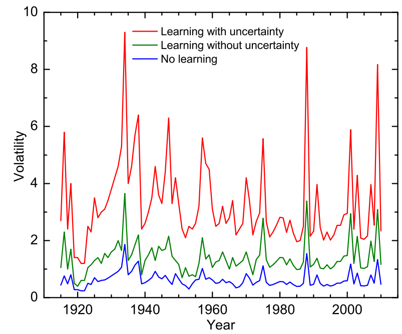
Weitzman, M., 2007, Subjective expectations and asset-return puzzles, *American Economic Review*, 97, 1102–1130.

Figure 2.1 **Sequential equity premium, conditional volatility, skewness and kurtosis of consumption growths**

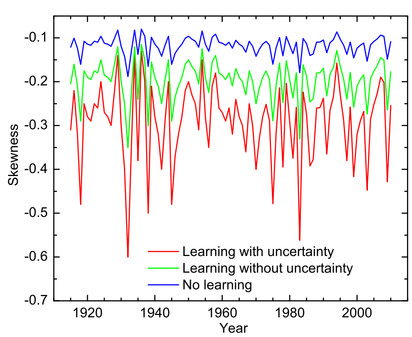
Sequential equity premium, conditional volatility, skewness and kurtosis of consumption growths estimated by models of (1) learning with parameter uncertainty, (2) learning without parameter uncertainty, and (3) no learning and no parameter uncertainty. The period is from 1915 to 2010.



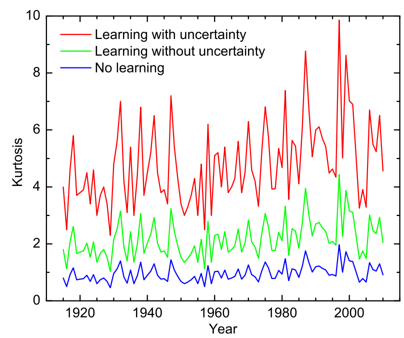
(a) Sequential equity premium.



(b) Conditional volatility of consumption growths.



(c) Conditional skewness of consumption growths.



(d) Conditional kurtosis of consumption growths.

Table 2.1 **Summary Statistics**
Entries are statistics computed from annual observations for the U.S. economy (1891-2010). Mean is the sample mean, Std is the standard deviation, Skewness is the standard measure of skewness, and Kurtosis is the standard measure of excess kurtosis. Consumption growth is $\log(C_t/C_{t-1})$, where C is real per capita consumption. Dividend growth is $\log(D_t/D_{t-1})$, where D is real dividend. Returns are gross real returns and the excess return is the difference between the returns on equity and the 1-year bond. The 1-year bond is the Treasury security of maturity closest to 1 year. Equity is the S&P 500. Consumption and dividend data are from Shiller (2010).

	Mean(%)	Std (%)	Skewness	Kurtosis
Consumption growth	2.00	3.52	-0.3519	4.09
Dividend growth	1.06	11.60	-0.6512	7.13
Risk free rate	1.91	1.33	1.1392	5.37
Return on equity	7.44	18.67	-0.0654	2.98
Excess return on equity	5.53	18.73	0.0085	2.98

Table 2.2 **Baseline Calibration**
Based on the first 25 observations, we calibrate 48 parameters in Equation (2.22) and list the main parameters.

	Value
Risk aversion γ	8
Intertemporal elasticity of substitution ψ	1.2
Discount factor β	0.988
Average consumption growth μ_c	0.025
Average dividend growth μ_d	0.012
Average consumption growth volatility $\Sigma_{1,1}$	0.0015
Average dividend growth volatility $\Sigma_{2,2}$	0.013
Correlation between consumption and dividend growth ρ	0.35
Skewness for consumption growth λ_c	-0.32
Skewness for dividend growth λ_d	-0.68
Degree freedom for consumption growth distribution v_c	9
Degree freedom for dividend growth distribution v_d	11

Table 2.3 **Moments of Consumption and Dividend Growths**
This table reports the unconditional mean, volatility, skewness, and kurtosis for real U.S log-consumption growth (in Panel A) and log-dividend growth (in Panel B) computed using annual data from 1915 to 2010, and using the Model of skew student's t-distribution (simulating the model 1000 times with sample size 96 years). The column labeled S.E reports the standard errors of these moments.

Panel A: Consumption Growth

	Data Estimate	Data S.E	Model Mean	Model S.E
$\mathrm{E}(g_c)(\%)$	1.97	0.78	1.80	1.23
$\sigma(g_c)(\%)$	3.13	0.58	4.40	0.64
skew(g_c)	-0.37	0.14	-0.30	0.04
kurt(g_c)	4.86	0.49	4.22	0.67

Panel B: Dividend Growth

	Data Estimate	Data S.E	Model Mean	Model S.E
$\mathrm{E}(g_d)(\%)$	0.77	0.65	0.80	0.40
$\sigma(g_d)(\%)$	11.45	7.36	9.70	6.30
skew(g_d)	-0.70	0.37	-0.62	0.40
kurt(g_d)	7.83	1.64	6.45	0.88

Table 2.4 **Effects by Skewness and Fat-Tail**
The first column reports the statistics of interest calculated using annual U.S data from 1915 to 2010. The second to fifth column reports the results by using skew student's t (SST), skew normal (SN), student's t (ST), and normal (NM) distribution models, respectively.

	Data	SST	SN	ST	NM
$\mathrm{E}[r_m - r_f]$	5.74	6.20	3.85	4.04	2.97
		[4.55,7.85]	[3.14,4.56]	[3.25,4.83]	[1.74,4.20]
$\sigma[r_m - r_f]$	19.7	17.6	12.4	11.0	8.7
		[13.1,22.1]	[9.2,15.6]	[8.1,13.9]	[5.9,11.5]
$\mathrm{E}[r_f]$	1.91	1.73	1.70	1.68	1.75
		[1.24,2.22]	[1.36,2.04]	[1.27,2.09]	[1.38,2.12]
$\sigma[r_f]$	1.48	1.54	1.20	1.25	0.96
		[1.25,1.83]	[0.97,1.43]	[1.03,1.47]	[0.70,1.24]

Table 2.5 **Predictability of Higher Moments**

This table shows the predictive power of higher moments on the equity premium in different distributions. For each column, the depend variable is the subsequent equity premium $r_{ep,t+1}$. α_{0-4} are the coefficients for the interception and the first to fourth moments of the consumption growth. In Model A of almost normal, we fix $\binom{\lambda_c=0}{\lambda_d=0}\binom{v_c=100}{v_d=100}$. In Model B of moderate normal, we fix $\binom{\lambda_c=-0.16}{\lambda_d=-0.34}\binom{v_c=20}{v_d=20}$. In Model C of highly non-normal, we fix $\binom{\lambda_c=-0.32}{\lambda_d=-0.68}\binom{v_c=10}{v_d=10}$. The regression is described as $r_{ep,t+1} = \alpha_0 + \alpha_1 E_t[g_{t+1}] + \alpha_2 E_t[(g_{t+1} - \bar{g}^t)^2] + \alpha_3 E_t[(g_{t+1} - \bar{g}^t)^3] + \alpha_4 E_t[(g_{t+1} - \bar{g}^t)^4] + \epsilon$.

	Model A		Model B		Model C	
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Mean	-0.173	-0.180	-0.176	-0.185	-0.192	-0.190
	[0.078]	[0.082]	[0.074]	[0.093]	[0.076]	[0.080]
Variance	0.048	0.037	0.057	0.060	0.054	0.058
	[0.062]	[0.043]	[0.042]	[0.047]	[0.039]	[0.044]
Skewness		-0.120		-0.150		-0.132
		[0.092]		[0.085]		[0.048]
Kurtosis		0.65		0.77		0.83
		[0.73]		[0.51]		[0.35]
Adj. R^2	1.6%	2.2%	2.0%	2.7%	2.4%	4.8%

Table 2.6 **Predictive Regressions**

For each column, the depend variable is the subsequent equity premium $r_{ep,t+1}$. α_{0-4} are the coefficients for the intercept and the first to fourth moments of the consumption growth. The moments are obtained by using skew student's t model. The regression is described as

$r_{ep,t+1} = \alpha_0 + \alpha_1 E_t[g_{t+1}] + \alpha_2 E_t[(g_{t+1} - \bar{g^t})^2] + \alpha_3 E_t[(g_{t+1} - \bar{g^t})^3] + \alpha_4 E_t[(g_{t+1} - \bar{g^t})^4] + \epsilon.$

Coefficient	[1]	[2]	[3]	[4]	[5]	[6]	[7]
$Mean$	-0.243				-0.212	-0.196	-0.190
	[0.065]				[0.077]	[0.075]	[0.080]
$Variance$		0.080			0.074	0.067	0.058
		[0.054]			[0.043]	[0.049]	[0.044]
$Skewness$			-0.184			-0.163	-0.130
			[0.053]			[0.065]	[0.048]
$Kurtosis$				0.96			0.85
				[0.42]			[0.34]
Adj. R^2	1.5%	0.9%	1.4%	1.2%	2.4%	3.5%	4.8%

Table 2.7 **Learning and Parameter Uncertainty**

This table shows the results between actual data and that obtained by learning and parameter uncertainty. [1] the case with learning and parameter uncertainty. [2] the case with learning, but no parameter uncertainty. [3] the case with neither learning nor parameter uncertainty. The period is from 1915 to 2010.

	Data	[1]	[2]	[3]
E$[r_m - r_f]$	5.74	6.20	3.90	2.70
		[4.55,7.85]	[2.70,5.10]	[2.20,3.20]
$\sigma[r_m - r_f]$	19.7	17.9	8.4	6.3
		[13.1,22.1]	[6.2,11.0]	[4.0,8.6]
E$[r_f]$	1.91	1.73	1.84	1.88
		[1.24,2.22]	[1.35,2.34]	[1.27,2.49]
$\sigma[r_f]$	1.48	1.54	1.64	2.12
		[1.25,1.83]	[0.97,2.31]	[1.42,2.82]