THREE ESSAYS ON PRICING KERNEL IN ASSET PRICING

by

Minhao Cai

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Business Administration

Charlotte

2013

Approved by:

____

Dr.Weidong Tian

____

Dr.Chris Kirby

____

Dr.Steven Clark

____

Dr.Zongwu Cai

ABSTRACT

MINHAO CAI. Three essays on pricing kernel in asset pricing.
(Under the direction of DR.WEIDONG TIAN)

Pricing Kernel extends concepts from economics and finance to include adjustments for risk. When pricing kernel is given, by non-arbitrage theory, all securities can be priced. Searching for a proper pricing kernel is one of the most important tasks for researchers in asset pricing. In this thesis, we attempt to search a proper pricing kernel in three different scenarios.

In chapter 1, we attempt to find a robust pricing kernel for a stochastic volatility model with parameter uncertainty in an incomplete commodity market. Based on a class of s-tochastic volatility models in Trolle and Schwartz (2009), we investigate how the parameter uncertainty affects the risk premium and commodity contingent claim pricing. To answer this question, we follow a two-step procedure. Firstly, we propose a benchmark approach to find an optimal pricing kernel for the model without parameter uncertainty. Secondly, we uncover a robust pricing kernel via a robust approach for the model with parameter uncertainty. Thirdly, we apply the two pricing kernels into the commodity contingent claim pricing and quantify effect of parameter uncertainty on contingent claim securities. We find that the parameter uncertainty attributes a negative uncertainty risk premium. Moreover, the negative uncertainty risk premium yields a positive uncertainty volatility component in the implied volatilities in the option market.

In chapter 2, we propose a multi-factor model with a quadratic pricing kernel, in which the underlying asset return is a linear function of multi-factors and the pricing kernel is a quadratic function of multi-factors. The model provides a potential unified framework to

link cross sectional literatures, time-series literatures, option pricing literatures and term structure literatures. By examining option data from 2005 to 2008, this model dramatically improves the cross-sectional fitting of option data both in sample and out of sample than many standard GARCH volatility models such as Christoffersen, Heston and Nandi (2011). This model also offers explanations for several puzzles such as the U shape relationship between the pricing kernel and market index return, the implied volatility puzzle and fat tails of risk neutral return density function relative to the physical distribution.

In chapter 3, we investigate whether idiosyncratic volatility risk premium is cross-sectional variant. We use stock historical moving average price as a proxy for retail ownership and examine whether idiosyncratic volatility is correlated with stock price level. Evidence from cross-sectional regressions and portfolio analysis both suggests that low-priced stocks (high retail ownership) have a significantly higher idiosyncratic volatility risk premium than high-priced stocks (low retail ownership). Especially, evidence in subsample tests suggests that lowest-priced stocks (highest retail ownership) have a significantly positive idiosyncratic risk premium while highest-priced stocks (lowest retail ownership) have an insignificant one, which is consistent with theoretical predictions of Merton (1986) and classical portfolio theory.

INTRODUCTION

Pricing Kernel extends concepts from economics and finance to include adjustments for risk. As long as pricing kernel is given, we can price any security by non-arbitrage theory. Searching for a proper pricing kernel is the core question for Researchers in the area of asset pricing. In finance, researchers often apply different models in different financial markets. Because the pricing kernel heavily depends on the specification of an asset pricing model, there is no unified pricing kernel that can justify all securities in different financial markets until today. In this section, we firstly present two widely-used class of asset pricing models and then discuss the corresponding pricing kernels.

Volatility models are one class of the most important asset pricing models in finance. Based on the principle that ``risk require rewards'', as a measure of risk, volatility are widely used to explain realized returns and expected returns of securities in different financial markets. In general,volatility models are categorized by two groups: stochastic volatility models (See Heston (1993), Trolle and Schwartz (2009)) and GARCH models (See Heston and Nandi (2000)). These two models can both well capture the time-varying volatility process. Even though a GARCH model is a discrete-time model while a stochastic volatility model is a continuous-time model, a stochastic volatility model can often be translated into a GARCH model with infinitesimal interval. However these two groups still have significant difference in predicting future volatility. In stochastic volatility models, there is a random component which assumes that future volatility cannot be fully determined by today's information. In contrast, GARCH models assume that future volatility is already known based on today's information. To price a security, the pricing kernel needs to be specified either explicitly or implicitly. In the explicit way, we explicitly specify the function of pricing kernel. Given the pricing kernel, we can find the return process of underlying assets under risk-neutral measure and price a security. In the implicitly way, we directly specify the return process of underlying assets under risk-neutral measure without giving a function form for the pricing kernel. However the pricing kernel implicitly affect security prices by risk premiums.

In the first chapter, we dedicate to find proper pricing kernels for a class of stochastic volatility models in different scenarios and analyze the effect of parameter uncertainty on option prices in an incomplete market. In a complete market, the pricing kernel is uniquely determined by traded assets in a financial market. However, when the market is incomplete, the pricing kernel is not unique, how to select an optimal pricing kernel is a difficult issue. In the implicit way, choosing an optimal pricing kernel is equivalent to choosing optimal risk premiums. As part of the volatility is un-spanned a stochastic volatility model, we may have concern on the accuracy of parameter calibrations. In this chapter, we propose two pricing kernels in two scenarios. Firstly, we propose a benchmark approach to determine an optimal pricing kernel in an incomplete market. The benchmark approach finds an optimal pricing kernel which maximizes the Sharpe ratio for all securities. Secondly, we use a robust approach to search for an optimal pricing kernel for the model with parameter uncertainty (volatility uncertainty) in an incomplete market. By comparing security prices via the two different pricing kernels, we can find the effect of parameter uncertainty on option pricing in an incomplete market. We find that the parameter uncertainty attributes a *negative uncertainty risk premium*. Moreover, the negative uncertainty risk premium yields a *positive uncertainty volatility component* in the implied volatilities in the option market. The uncertainty volatility component can be even significant under certain circumstances.

Linear factor models are another strand of asset pricing models in finance. A linear factor model relates the return on an asset to the values of a limited number of factors, with the relationship described by a linear equation. Compared with volatility models, linear factor models are better in capturing cross-sectional variants among stocks in equity market. Since the Capital asset pricing model can not explain anomalies in the market, researches on linear factor models mainly focus on searching of new market factors. Fama and French (1993) find that SML, HML and market excess return can explain over 90 percent of the time series variation of diversified portfolios. Carhart (1997) finds that momentum factor is another important market factor which can explain why past winners continue to win in the near future.

By non-arbitrage pricing theory, the pricing kernels of linear models should be in a linear form by the principle of one price in stock market. In the bond market, the pricing kernel is often specified as a linear function of multi-factors. For example, affined term structure models (ATSMs) assume that the logarithm of pricing kernel is a linear function of latent factors. For the inter-temporal consumption based model, Harrison and Kreps (1979) and Hansen and Jagannathan (1991) have showed that the pricing kernel is a representative investor's intertemporal marginal rate of substitution of consumption. when the growth rate of consumption is proportional to the growth rate of market return, we can derive the pricing kernel as linear function of the market return from Taylor expansions. By classic economic theory, because of the diminishing marginal utility, the pricing kernel should be a decreasing function of the market return. However, empirical findings find that the pricing kernel is a ``non-monotone" function of underlying asset return and is time varying. For example, Bakshin, Madan and Panayotov (2010) use four different OTM calls which focus on claims with payout on the upside and find that the average returns contradict the downward-slope pricing kernel. To explain the pricing kernel puzzle in option market, we propose a new pricing kernel which is a quadratic function of factors in the second chapter.

Motivated by these findings, we propose a multi-factor model with a quadratic pricing kernel to explain several anomalies in equity market. In this model, the market return is a linear function of multi-factors while the logarithm of pricing kernel is a quadratic function of multi-factors. By the settings, our model can automatically explain the U-shaped relationship between pricing kernel and market return. Even though a quadratic pricing kernel has been discussed in primitive equity market and bond market, In stock market and bond market, A quadratic pricing kernel has been well studied in many literatures. Harvey and Siddique (2000) and Dittmar (2002) both develop the quadratic pricing kernel in cross section analysis. In bond market, although affine term structure models (ATSMs) are prevail, Ahn, Dittmar and Gallant (2002) propose quadratic term structure models (QTSMs) in which the pricing kernel is a quadratic function of factors. They demonstrate that the QTSMs can overcome several disadvantages of ATSMs and better fit the US data. However, a quadratic pricing kernel has never been used to price derivatives in the equity market, because the pricing kernel needs to both satisfy non-arbitrage theory and price derivatives easily. In Chapter 2, we propose a multi-factor model with a quadratic pricing kernel and present an almost closed form solution to option price. By fitting cross-sectional option data from 2005 to 2008, we find that this model can significantly outperform Christoffersen, Jacob and Heston (2012) both in sample and out of sample. The presented model not only can explain the pricing kernel puzzle, but also can explain several other empirical puzzles, such as implied volatility puzzle and fat tails of risk neutral return density function relative to the physical distribution.. The proposed model also builds linkages among the linear factor model, the pricing kernel literature and the option pricing literature.

By modern economic theory, because only systematic risk is priced, the pricing kernel only needs to adjust for systematic risk. However, this theory is based on the assumption that representative investors will hold well-diversified portfolios. It is well known that investors, especially retail investors often hold under-diversified portfolios in reality. Merton (1987) points out that idiosyncratic risk should be positively priced when representative investors hold under-diversified portfolios. The two theories gave out two different answers based on different representative investors. However, in regardless of representative investors, empirical studies often treat all the stocks the same and test whether idiosyncratic risk is priced using the full sample. In the third chapter, we attempt to fill the gap between theories and empirical studies on pricing idiosyncratic risk. We attempt to prove that idiosyncratic volatility risk premium is clientele-based and separately test whether idiosyncratic risk is priced differently for stocks with different representative investors. Because idiosyncratic risk premium is the negative correlation between the pricing kernel and idiosyncratic risk, this is equivalent to test whether stocks with different representative investors have different pricing kernels. We use stock historical moving average price as a proxy for retail ownership and examine whether idiosyncratic volatility is correlated with stock price level. Evidence from cross-sectional regressions and portfolio analysis both suggests that low-priced stocks (high retail ownership) have a significantly higher idiosyncratic volatility risk premium than high-priced stocks (low retail ownership). Especially, evidence in subsample tests suggests that lowest-priced stocks (highest retail ownership) have a significantly positive idiosyncratic risk premium while highest-priced stocks (lowest retail ownership) have an insignificant one, which is consistent with theoretical predictions of Merton (1986) and classical portfolio theory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: PARAMETER UNCERTAINTY AND VOLATILITY RISK PREMIUM

1.1 Introduction

This paper is a study into how the market incompleteness interacts with the model parameter uncertainty in a commodity market. It is well documented recently that, there is unspanned volatility in the commodity market so the stochastic volatility model is required to capture the commodity price movement (See, for instance, Trolle and Schwartz (2009, 2010), Hughen (2010).). In this chapter, we consider a class of stochastic volatility model developed in Trolle and Schwartz (2009) for the crude oil market, we investigate the following questions. What if the volatility model is not correct? How to characterize the market price of the parameter uncertainty in a robust approach? and how large the model uncertainty contributes to the derivative pricing and implied volatility?

For this purpose, we introduce a three-step procedure. In the first step, we develop a robust approach for the class of stochastic volatility model in the presence of model parameter uncertainty [1] by following a methodology developed in Boyle, Feng, Tian and Wang (2008). In essence, the robust approach is to find the robust stochastic discount factor which is the least sensitive stochastic discount factor (SDF) with respect to the perturbation of model parameters in pricing a given security. This robust approach yields a robust stochastic

[1] Model uncertainty emerges when there is not accurate estimation of model parameters and can be essential in some situations and ignoring the model uncertainty can lead to a pronounced, negative effect. See, for instance, Garlappi, Uppal and Wang (2007) for expected return uncertainty, Tian, Wang and Yan (2010) for correlation uncertainty. Recently, Lo and Mueller (2010) propose a taxonomy of uncertainty and some uncertainty might be not reducible. In this paper we follow a relatively narrow sense in that it is impossible to estimate accurately some model parameters on the volatility process.

discount factor and its corresponding (robust) risk premium when the model parameter uncertainty interacts with the stochastic volatility. As the robust risk premium estimated in this step consists of ingredients from both parameter uncertainty and market incompleteness together, the risk premium attributed by the parameter uncertainty purely is not characterized yet.

To estimate the parameter uncertainty premium, we take an indirect way in the second step. We examine a benchmark case in which model parameters are known while only stochastic volatility is persistent. We propose one approach, namely a benchmark approach, to the stochastic volatility model that leads to a benchmark SDF and a stochastic volatility risk (benchmark risk) premium. The benchmark approach, in nature, is similar to the robust approach mentioned above without uncertainty concern and can be applied to any incomplete financial asset pricing model. In this way, we divide the incompleteness and uncertainty separately so we are able to estimate the uncertainty risk premium, which is the difference between the risk premiums in the above two approaches. Precisely, the robust risk premium is decomposed as a sum of the benchmark risk premium and the uncertainty risk premium

$$\begin{aligned}\text{Robust Risk Premium} &= \text{Benchmark Risk Premium} \\ &\quad + \text{Uncertainty Risk Premium}\end{aligned}$$

where the robust risk premium is the solution in the robust approach while the benchmark risk premium is solved by the benchmark approach.

In the third step, we apply two SDFs, the robust SDF and the benchmark SDF constructed in the first two steps, into the commodity contingent claim pricing. In particular, we study the joint effects of market incompleteness and parameter uncertainty on the implied volatility of the option market. We find out that model parameter uncertainty contributes a ``uncertainty volatility component" in the implied volatility of the commodity option market. Similar to the decomposition of the robust risk premium, the implied volatility in the presence of model parameter uncertainty and stochastic volatility, namely the robust implied volatility, is decomposed as

$$\begin{aligned}\text{Robust Implied Volatility} &= \text{Benchmark Implied Volatility} \\ &\quad + \text{Uncertainty Volatility Component.}\end{aligned}$$

Following the above proposed methodology, the empirical results show that, using a recent sample of data in the commodity market, the volatility risk premiums and the uncertainty risk premium are negative. Negative volatility risk premiums consistent with recent empirical literatures such as Doran and Ronn (2008), and Christoffersen, Heston and Jacobs (2010). Doran and Ronn (2008) document that negative volatility risk premium is closely related to the disparity between risk-neutral and statistical volatility in the commodity market. Based on arbitrage-free principle, Christoffersen, Heston and Jacobs (2010) demonstrate that negative volatility risk premium is persistent in a class of stochastic volatility model. In addition to the negative volatility uncertainty premium, we also show that the uncertainty volatility component is positive. The intuition of a positive uncertainty volatility component is that a higher price is required to compensate the parameter uncertainty. Therefore, the implied volatility is increased when the parameter uncertainty is added. Under certain circumstance, the uncertainty volatility component compared to the robust implied volatility can be fairly substantial.

Our approach is related to previous researches in dealing with the commodity risk. Some authors follow the partial equilibrium approach by giving stochastic processes on the economical factors such as spot and future prices or demand, and follow an arbitrage-free argument to analyze the commodity risk. See, for instance, Gibson and Schwartz (1990), Miltersen and Schwartz (1997). As recognized in Cassus and Collin-Dufresne (2005), Hughen (2009) and Trolle and Schwartz (2009), future prices cannot fully span the volatility, hence the stochastic volatility model needed to be developed. As the commodity market becomes incomplete, it is a challenging problem to estimate the stochastic discount factor, or equivalently, the market price of risk in the commodity market. In another strand of literature, such as Routledge, Seppi and Spatt (2000), Hong (2000), and Kogan, Livdan and Yaron (2008), an equilibrium model for the term structure of future prices is developed. The commodity derivative can be priced uniquely in this equilibrium approach. The endogenous stochastic process of price process as well as the market price of risk, however, does not always fit the commodity market well. The purpose of this paper is to study the robust stochastic discount factor for a given commodity derivative contract and examine the model parameter uncertainty effect in a precise manner.

This paper also contributes to asset pricing literature. In standard asset pricing literatures on incomplete financial model, the agent wants to determine the stochastic discount factors (SDF). In an equilibrium approach the stochastic discount factor is uniquely determined if the agent's risk preference is specifically characterized. In a good-deal approach, the class of stochastic discount factors can be significantly smaller by imposing economical-based bounds on the stochastic discount factors than that in the no-arbitrage approach. See Bernardo and Ledoit (2000), and Cochrane, Saa-Requejo (2000). Other approaches to find a narrow class of stochastic discount factors we refer to Bondarenko (2003), Carr and Madan (2001). Boyle, Feng, Tian and Wang (2008) introduce a robust approach to pricing financial securities, when the model parameters are uncertain in an incomplete financial market. A remarkable difference between the robust approach with other approaches is that the robust stochastic discount factor could depend on a derivative itself. In other words, corresponding to different derivative contract, the agent may need to use different robust stochastic discount factor to deal with both the market incompleteness and parameter uncertainty together. It makes economical sense since there is no one stochastic discount factor that can be applied to all securities in an incomplete model containing parameter estimation risk when risk preference is not imposed en-ante. This paper examines in details how the market incompleteness and parameter uncertainty jointly affect the asset pricing in the real market place by extending the robust approach in Boyle, Feng, Tian and Wang (2008).

The methodology developed in this paper has several potential applications and implications. (a). The approach can be used to analyze the scale of the implied volatility. Carr and Wu (2009) find that the realized volatility can't be as large as the implied volatility in the market place. The empirical analysis in this paper might suggest that a significant component in the implied volatility that is driven from the model parameter uncertainty. Therefore, our result could be useful to explain the gap between the realized volatility and the implied volatility. (b). The presented methodology can be used to discuss the risk management strategy for the price risk and the parameter estimation risk in one uniform setting. Boyle, Feng, Tian and Wang (2008) develop another robust approach to study the event of small probability with large price movement. By combing this robust approach and the method developed in this paper one can examine the risk management strategy that is robust to a large loss with small probability. (c). The method developed in this paper can be applied to other financial markets. For instance, Collin-Dufresne and Goldstein (2002), Li and Zhao (2006) document the unspanned stochastic volatility in the fixed income market. Therefore, one can apply our method to the fixed income market as well.

The paper is organized as follows. Section 1.3 presents a benchmark approach in an incomplete model without volatility model parameter uncertainty. We also show that this benchmark approach is identical to the optimal-variance approach that is proposed in some other contexts. Section 1.4 introduces the robust approach for incomplete model with volatility model uncertainty. We examine the robust approach for the pricing of commodity contingent claim. In Section 1.5 we calibrate the model in the crude oil market and present our empirical results. Section 1.6 provides the conclusions and all proofs are presented in Proof.

1.2 A Benchmark Approach

We present a benchmark approach to obtain the benchmark stochastic discount factor in an incomplete model. The insight is better illustrated in a single-period setting first. Consider a derivative with payoff p and whose price is unknown. There is a set of admissible stochastic discount factor $\{m_\alpha : \alpha \in \mathscr{A}\}$. In asset pricing theory (see Harrison and Kreps (1979)), corresponding to each SDF m_α, the expectation $\mathbb{E}[m_\alpha p]$ offers one arbitrage-free price of p. The price of the derivative is not unique when the derivative p is not attainable. In addition, the range between the smallest and highest arbitrage-free price of p, that is $\min \mathbb{E}[m_\alpha p]$ and $\max \mathbb{E}[m_\alpha p]$, might be too large to be used in applications.

Consider an agent who want to find a unique price to hedge the risk of choosing the ``wrong" stochastic discount factor. To do this, the agent finds an approximate derivative, written as $z_\alpha + p$, of the original derivative p for each possible SDF m_α. The choice of the approximate derivative hinges on some hedging considerations that m_α might be chosen inappropriately. The criteria of choosing the derivatives $z_\alpha + p$ will be explained shortly. Assuming these approximate derivatives $\{z_\alpha + p : \alpha \in \mathscr{A}\}$ have been chosen, the agent comes up with one quantity $A(m_\alpha)$ for each stochastic discount factor m_α. The benchmark approach is to find the robust SDF m_α through the quantities $A(m_\alpha)$.

Evidently, the approximate derivatives are not unique and we follow the robust approach proposed in Boyle, Feng, Tian and Wang (2008),[2]. To choose the approximate derivative, $z_\alpha + p$, for a given stochastic discount factor, the construction is build on two following criteria[3]. (1). The approximate derivative $z_\alpha + p$ has the same value as p under the same stochastic discount factor m_α, that is, $\mathbb{E}_P[m_\alpha(z_\alpha + p)|\mathscr{F}] = \mathbb{E}_P[m_\alpha p|\mathscr{F}]$. It can be written as

$$\mathbb{E}_P[m_\alpha z_\alpha|\mathscr{F}] = 0 \tag{1.1}$$

where $\mathscr{F}$ represents the information set and P is the subjective probability. (2). The approximate derivative $p + z_\alpha$ has the highest Sharpe ratio, given a stochastic discount factor m_α. The intuition behind the second criteria is that the agent finds the approximate derivative with the smallest risk (volatility) given the expected return from the hedging perspective of the derivative $p + z_\alpha$.

[2]For other robust approach we refer to Hansen and Sargent (2001). The ideas in the robust approach are fairly similar, to deal with the worst case.

[3]Of course, the choice of those criteria is not unique by the incompleteness market structure. We argue that this approach is similar to the robust one in which the model parameter estimation risk is a concern. See Section 2 in Boyle, Feng, Tian and Wang (2008) for the justifications.

In a precise manner, the excess return of $p+z_\alpha$ under m_α is

$$\mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha(p+z_\alpha)] = \mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha p] \tag{1.2}$$

where the first criteria is used. Hence, the Sharpe ratio of $p+z_\alpha$, under m_α, is

$$\frac{\mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha(p+z_\alpha)]}{\sqrt{Var[(p+z_\alpha)]}}.$$

Thus the criteria for finding the ``approximate derivative" is to solve the following problem

$$\max_{z_\alpha, \mathbb{E}_P[m_\alpha z_\alpha]=0} \frac{\{\mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha(p+z_\alpha)]\}^2}{Var[(p+z_\alpha)]}.to$$

Since the agent is uncertain about which SDF is appropriate for the derivative p, the benchmark approach in the incomplete market is to find m_{α^*} by

$\min_{m_\alpha} \max_{z_\alpha, \mathbb{E}_P[m_\alpha z_\alpha]=0} \frac{\{\mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha(p+z_\alpha)]\}^2}{Var[(p+z_\alpha)]}$.

The solution of this benchmark approach is presented by the next proposition.

$\max_{\mathbb{E}_P[m_\alpha z_\alpha]=0} \frac{\{\mathbb{E}_P[(p+z_\alpha)] - \mathbb{E}_P[m_\alpha(p+z_\alpha)]\}^2}{Var[(p+z_\alpha)]} = \mathbb{E}_P[m_\alpha^2] - 1.$

The benchmark approach is to find the robust SDF m_{α^*} such that

$\min_{m_\alpha} Var[m_\alpha]$.

See Proof.

This result is intuitive. In the choice of the derivative $p+z_\alpha$, as the criteria is to find the highest Sharpe ratio with corresponding SDF m_α, this criteria is equivalent to the optimal derivative choice in the mean-variance setting. Indeed, the approximate derivative $z_\alpha + p$, given one SDF m_α, can be found as follows. Given any a real number c, let

$$
\begin{aligned}
z_\alpha(c,\omega) &= \mathbb{E}_P[p] - p(\omega) + c \\
&\quad + \{\mathbb{E}_P[m_\alpha p] - \mathbb{E}_P[p] - c\} \left(\frac{Q^\alpha(\omega)}{P(\omega)} - 1\right) \left(\mathbb{E}_P\left[\left(\frac{dQ^\alpha}{dP}\right)^2\right] - 1\right)^{-1}
\end{aligned}
$$

where the probability measure Q^α is determined by $\frac{dQ^\alpha}{dP} = m_\alpha$. As shown in Proof, $z_\alpha(c)$ offers the solution of the maximum Sharpe ratio problem (1.3) for each $\mathbb{E}[z_\alpha(c)] = c$. It corresponds to the efficient frontier in the mean-variance framework. Therefore, from the perspective of the SDF m_α, the approximate derivative of the derivative p is written as a linear contract as follows:[4]

$$
\mathbb{E}_P[p] + c + \{\mathbb{E}_P[m_\alpha p] - \mathbb{E}_P[p] - c\} \left(\frac{Q^\alpha(\omega)}{P(\omega)} - 1\right) \left(\mathbb{E}_P\left[\left(\frac{dQ^\alpha}{dP}\right)^2\right] - 1\right)^{-1}.
$$

So Proposition 1.4 follows from a direct computation.

Even though the approximate derivatives are not unique, the highest Sharpe ratio of these approximate derivatives depends only on the SDF m_α. It is worth to mention that the term $\mathbb{E}_P\left[\left(\frac{dQ^\alpha}{dP}\right)^2\right] - 1$ in Proposition 1.4 has an important economical meaning. By Hansen and Jagannathan (1997) and Goetzmann, Ingersoll and Spiegel (2007), the square root of $\mathbb{E}_P\left[\left(\frac{dQ^\alpha}{dP}\right)^2\right] - 1$ is the solution of the following maximum Sharpe ratio problem for $\tilde{x}$ given the stochastic discount factor m_α, as follows: $\max_{E[m_\alpha \tilde{x}]=0} \frac{\mathbb{E}[\tilde{x}]}{\sqrt{Var[\tilde{x}]}}$ where $\tilde{x}$ is the excess return of the portfolio. Problem (1.3) is a relative version of Problem (1.4) with respect to a derivative p.

By Proposition 1.4, the solution, m_α^*, of the benchmark approach is the one that min-

[4]The optimal derivative contract that minimize the variance while the expected return is greater than a threshold is often a linear derivative contract on m_α. See Jankunas (2001).

imizes the variance of the stochastic discount factors. The benchmark SDF m_{α}^{*} does not depend on any derivative p. It also turns out that this benchmark approach leads to the same solution of the optimal-variance approach developed in Schweizer (1995, 1996) from different perspective.

We now move to the stochastic volatility model in the commodity market by using this benchmark approach.

A. Stochastic Volatility Model in Commodity Market

We consider the crude oil market and examine a three-factor stochastic model in this paper. We start with a filtered probability space $(\Omega, \mathscr{F}, P)$. $S(t)$ denotes the time t spot price of the commodity, $y(t,T)$ is the time-t instantaneous forward carrying cost at time T. Both the volatilities of $S(t)$ and $y(t,T)$ are controlled by the (unspanned) stochastic volatility factor $v(t)$, which satisfies a square-root process.[5] Precisely,

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu_s(t)dt + \sigma_s\sqrt{v(t)}dW_1(t)^P, \\ dy(t,T) &= \mu_y(t,T)dt + \sigma_y(t,T)\sqrt{v(t)}dW_2(t)^P, \\ dv(t) &= (\gamma - \kappa v(t))dt + \sigma_v\sqrt{v(t)}dW_3(t)^P, \end{aligned} \tag{1.5}$$

where $\mu_s(t)$ is the instantaneous drift of the spot price, $\mu_y(t,T)$ the instantaneous drift of the forward carrying cost $y(t,T)$, σ_s is a constant. We assume $\sigma_y(t,T)$ is deterministic. The random noise is generated by correlated Brownian motion $(W_1(t)^P, W_2(t)^P, W_3(t)^P)$ under

[5]The convenience yield for commodity is justified by the relationship between the storage and investment opportunity. See Deaton and Laroque (1996). To simplify the notation and focus on stochastic volatility and parameter uncertainty, we assume constant interest rate. See Schwartz (1997) and Trolle and Schwartz (2009) for discussions of stochastic interest rate on the commodity market.

probability measure P. We assume that

$corr(W_i(t)^P, W_j(t)^P) = \rho_{ij}$, $\forall i,\ j = 1,2,3$. $v(t)$ is the instantaneous variance, up to a positive constant, of the spot price, and $v(t)$ revers to a long-term mean of $\frac{\gamma}{\kappa}$ with a speed of κ. We assume that $2\gamma > \sigma_v^2$, so the positivity of the instantaneous variance is guaranteed as long as the initial value $v(0) > 0$. This model specification includes the one-factor volatility model presented in Trolle and Schwartz (2009) as a special case in which all market price of risks are assumed to be zero.[6]

As both the forward carrying cost and the spot volatility cannot be observed directly, the estimation of the market price of risk is crucial to price derivatives. To simplify the model settings, we assume that the market price of risk in spot price, forward carrying cost and instantaneous variance are in the form of $\lambda_i\sqrt{v(t)}, i = 1,2,3$. So we need to estimate these parameters λ_1, λ_2 and λ_3. In the sequel we examine how to choose those parameters to address the market incompleteness for this unspanned stochastic volatility model of crude oil market. [7]

We follow the benchmark approach stated above to find the benchmark SDF in this commodity model. With a particular choice of the market price of risk $\{\lambda_i\sqrt{v(t)} : i = 1,2,3\}$ at any instant time t, there exists a stochastic discount factor[8]

$$Z_T^{\lambda_1,\lambda_2,\lambda_3} \equiv exp\left\{-\sum_i^3 \int_0^T \lambda_i\sqrt{v(s)}dW_i(s) - \frac{1}{2}\int_0^T \sum_i^3\sum_j^3 \rho_{ij}\lambda_i\lambda_j v(s)ds\right\}. \tag{1.6}$$

[6] In fact, Trolle and Schwartz (2009) also introduce a two factor volatility model in the commodity market. Based on their extensive comparison analysis, the one factor volatility model explains the commodity market fairly well. Therefore, we only use the one factor volatility model.

[7] Our market price of risk specifications are standard and can be used to explain some option anomalies. For instance, Christoffersen, Heston and Jacobs (2010) examine the equity index option market. See also Bikbov and Chernov (2009), Broadie, Chernov and Johannes (2007), Cheredito, Filipovic and Kimmel (2007) for the discussion of the market price of risk specification.

[8] By model assumption, the Novikov condition holds, i.e., $\mathbb{E}\left[exp\left\{\frac{1}{2}\int_0^T \sum_i^3\sum_j^3 \rho_{ij}\lambda_i\lambda_j v(s)ds\right\}\right] < \infty$. Then $\mathbb{E}[Z_T^{\lambda_1,\lambda_2,\lambda_3}] = 1$.

Let $Q = Q^{\lambda_1,\lambda_2,\lambda_3}$ be the corresponding risk-neutral martingale measure of SDF $Z_T^{\lambda_1,\lambda_2,\lambda_3}$. Then the movements of $\{S(t), y(t,T), v(t)\}$ are represented by the following processes:

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \delta(t)dt + \sigma_s\sqrt{v(t)}dW_1(t)^Q, \\ dy(t,T) &= (\mu_y(t,T) - \lambda_2\sigma_y(t,T)v(t))dt + \sigma_y(t,T)\sqrt{v(t)}dW_2(t)^Q, \\ dv(t) &= (\gamma - \kappa v(t) - \lambda_3\sigma_v v(t))dt + \sigma_v\sqrt{v(t)}dW_3(t)^Q, \end{aligned} \tag{1.7}$$

where $\delta(t) = \mu_s(t) - \lambda_1\sigma_s v(t)$, $\delta(t)$ denotes the instantaneous carrying cost under Q. Moreover, $y(t,t) = \delta(t)$ by its definition. By Girsanov's theorem, the process $\{W_i^Q(t)\}$ defined by $W_i(t)^Q = W_1(t)^P + \int_0^t \lambda_i\sqrt{v(s)}ds$ is a Brownian motion under the measure Q.

Let $P(S,y,v,t)$ be the crude oil contingent claim price under the risk-neutral martingale measure Q, and assume it is a twice continuously differentiable function of the spot price, forward carrying cost and the volatility, then

$$\begin{aligned} &\frac{1}{2}P_{SS}S^2\sigma_s^2 v + \frac{1}{2}P_{yy}\sigma_y^2 v + \frac{1}{2}P_{vv}\sigma_v^2 v + P_{Sy}S\sigma_S\sigma_y\rho_{12} + P_{Sv}S\sigma_S\sigma_v v\rho_{13} + P_{yv}y\sigma_y\sigma_v v\rho_{23} \\ &\quad + P_S S\delta(t) + P_y(\mu_y - \lambda_2\sigma_y v) + P_v(\gamma - \kappa v - \lambda_3\sigma_v v) - P_t - rP = 0. \end{aligned} \tag{1.8}$$

Any oil contingent claim satisfies the above equation with some specific boundary condition. For instance, let $F(t,T)$ be the time t price of a future contract maturing at time T, it can be shown that (See Proof. It includes Trolle and Schwartz (2009), Proposition 1, as a special case.),

$$\frac{dF(t,T)}{F(t,T)} = \sqrt{v(t)}\left[\sigma_s dW_1(t)^Q + \int_t^T \sigma_y(t,u)du dW_2(t)^Q\right] \tag{1.9}$$

and

$$\mu_y(t,T) = -v(t)\sigma_y(t,T)\left[\rho_{12}\sigma_s + \int_t^T \sigma_y(t,u)du - \lambda_2\right]. \tag{1.10}$$

Under the risk-neutral probability measure Q, the forward carry of cost satisfy

$$\begin{aligned} y(t,T) &= \int_0^t -v(u)\left[\sigma_y(u,T)\left(\rho_{12}\sigma_s\int_u^T \sigma_y(u,T)du - \lambda_2\right) + \lambda_2\sigma_y(u,T)\right]du \\ &\quad + y(0,T) + \int_0^t \sigma_y(u,T)\sqrt{v(t)}dW_2(u)^Q \end{aligned} \tag{1.11}$$

By the above formulas (1.10)-(1.11), the market price of risk $\lambda_2\sqrt{v}$ of the forward carrying cost can be estimated by the future prices of the market. However, under most circumstances, this parameter λ_2 cannot be estimated accurately and it is even more challenging to estimate the other parameters λ_1 and λ_3.

As the commodity market is incomplete, the estimated parameters $\{\lambda_1, \lambda_2, \lambda_3\}$ are not unique. We assume $(\lambda_1, \lambda_2, \lambda_3) \in \mathscr{A} \subseteq \mathbb{R}^3$, a exogenously given bounded closed subset of $\mathbb{R}^3$. By Proposition 1.4, the benchmark approach as illustrated above solves $\min_{(\lambda_1,\lambda_2,\lambda_3)\in\mathscr{A}} \mathbb{E}[Z_T^2]$. Let $(\lambda_1^*\sqrt{v}, \lambda_2^*\sqrt{v}, \lambda_3^*\sqrt{v})$ be the minimal-variance market price of risk in this model. The analytical expression of $\mathbb{E}[Z_T^2]$ is available and written as follows (see Proof for its derivation):

$$\mathbb{E}[Z_T^2] = e^{A(T)-B(T)v_0} \tag{1.12}$$

where

$$A(t) = \frac{k_1 k_2^*}{\sigma_v^2} t + \frac{k_1 \ln[1 + \tan[\frac{1}{2}\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2}(t + C_1)]^2]}{\sigma_v^2} + C_2 \tag{1.13}$$

and

$$B(t) = \frac{-k_2^* + tan[\frac{1}{2}t\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2} + \frac{1}{2}C_1\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2}]\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2}}{\sigma_v^2} \tag{1.14}$$

and $k_1 = \gamma, k_2^* = k_2 + 2\lambda_3\sigma_v$, $\lambda = \sum_i^3 \sum_j^3 \rho_{ij}\lambda_i\lambda_j$, and parameters C_1, C_2 are given below (under the condition that $2\lambda\sigma_v^2 > (k_2^*)^2$) :

$$C_1 = \frac{2\arctan[\frac{k_2^*}{\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2}}]}{\sqrt{2\lambda\sigma_v^2 - (k_2^*)^2}} \tag{1.15}$$

$$C_2 = -\frac{k_1 \ln[1 + (\frac{k_2^*}{2\lambda\sigma_v^2 - (k_2^*)^2})^2]}{\sigma_v^2}. \tag{1.16}$$

After discussing the benchmark approach for this stochastic volatility model, we now move to the robust approach in which the model parameters are not known prior.

1.3 Pricing with Estimation Risk

In this section we assume that the model parameters on the volatility factor $v(t)$ are unknown. As documented in Trolle and Schwartz (2009), Hughen (2009) for the commodity market, Collin-Dufresne and Goldstein (2002), Li and Zhao (2006) for the fixed income market, the unspanned volatility feature of $v(t)$ makes it very difficult to estimate the market price of risk and model parameters. Therefore, we follow the robust approach with modification, developed in Boyle, Feng, Tian and Wang (2008), for the pricing of a commodity contingent claim.

To interpret the estimation risk of the model parameters, we consider a class of model perturbations, defined by

$$\frac{dS(t)}{S(t)} = \mu_s(t)dt + \sigma_s\sqrt{v_\varepsilon(t)}dW_1(t)^P,$$
$$dy(t,T) = \mu_y(t,T)dt + \sigma_y(t,T)\sqrt{v_\varepsilon(t)}dW_2(t)^P,$$
$$dv_\varepsilon(t) = (\gamma - \kappa v_\varepsilon(t))dt + \sigma_v(t,T)\varepsilon\sqrt{v_\varepsilon(t)}dW_3(t)^P, \tag{1.17}$$

where ε moves between the lowest one ε_{low} and the highest one ε_{high}. Each value of ε determines a specific volatility model in this class.

The above class of model perturbations describes the possible model parameter uncertainty, index by the parameter ε. In one extreme case $\varepsilon = 1$, the model reduces to the stochastic volatility model discussed in the last section. In another extreme case $\varepsilon = 0$, the volatility $v(t)$ becomes deterministic. In general $\varepsilon \in (0,1)$, and the range $(\varepsilon_{low}, \varepsilon_{high})$ depends on available information the agent access. The more information about the unspanned volatility $v(t)$ and the more accuracy the model specification is, the smaller the range $(\varepsilon_{low}, \varepsilon_{high})$ of ε.

We discuss the robust approach for this stochastic volatility model via an expectation method. The main insight of the robust approach is as follows. For each possible market price of risk specification, characterized by $\{\alpha \equiv (\lambda_1, \lambda_2, \lambda_3) \in \mathscr{A}\}$, there exists a sequence of arbitrage-free stochastic volatility model endowed with a risk-neutral probability measure, indexed by ε. Hence, there exists an arbitrage-free price $C(p;\alpha,\varepsilon)$ for a given derivative p. By contrast to the benchmark approach of the last section, the parameter ε denotes the parameter uncertainty. If we choose the deterministic volatility model as a benchmark model in the expectation method, the difference between the arbitrage-free price $C(p;\alpha,\varepsilon)$ and $C(p;\alpha,0)$ illustrates the effect of parameter uncertainty on the option prices. The robust SDF is the one to minimize the differences $|C(p;\alpha,\varepsilon)-C(p;\alpha,0)|$ when ε varies to some extents.

We now take the deterministic volatility model as a benchmark model (in which volatility is spanned by the future prices), and derive in details the robust approach for a call option on future contract. In the deterministic volatility model,[9]

$$dv(t) = (\gamma - \kappa v(t) - \lambda_3 \sigma_v v(t))\, dt, \tag{1.18}$$

Let $\gamma = \kappa_1, \kappa + \lambda_3 \sigma_v = \kappa_2$, so $v(t)$ satisfies (to highlight the dependence to the parameters κ_1, κ_2)

$$dv(t;\kappa_1,\kappa_2) = (\kappa_1 - \kappa_2 v(t;\kappa_1,\kappa_2))\, dt, v(0;\kappa_1,\kappa_2) = \sigma_0^2 \tag{1.19}$$

where $\kappa_1 > 0, \kappa_2 > 0$. It is well known that the path integral of the instantaneous variance,

$$V(\kappa_1,\kappa_2,0) = \int_0^{T'} v(t;\kappa_1,\kappa_2) dt, \tag{1.20}$$

plays a key role in pricing the option in the model. As our purpose is to compare the benchmark approach and the robust approach, to simplify some technical results and notations,

[9]See discussions in Boyle, Feng, Tian and Wang (2008) as to justification of this kind of benchmark model.

we further assume that $\rho_{13} = 0, \rho_{23} = 0$.[10] The derivative under considered is a call option with maturity T' on a T-year future contract, strike price is K.

By Black's formula (Black (1976)), the price of this call option in the deterministic volatility model is

$$C(\sqrt{V(\kappa_1,\kappa_2,0)}) = e^{-rT'}[FN(d_1) - KN(d_2)], \tag{1.21}$$

where

$$\begin{aligned} d_1 &= \frac{\ln\frac{F}{K} + \frac{1}{2}BV(\kappa_1,\kappa_2,0)}{\sqrt{B}\sqrt{V(\kappa_1,\kappa_2,0)}}, \\ d_2 &= \frac{\ln\frac{F}{K} - \frac{1}{2}BV(\kappa_1,\kappa_2,0)}{\sqrt{B}\sqrt{V(\kappa_1,\kappa_2,0)}}, \end{aligned} \tag{1.22}$$

and

$$B = \sigma_s^2 + \left(\int_0^T \sigma_y(t,u)du\right)^2 + 2\rho_{12}\sigma_s \int_0^T \sigma_y(t,u)du. \tag{1.23}$$

By the same method, we can price the call option in a sequence of perturbation model, for each possible market price of risk. If fact, given $(\lambda_1,\lambda_2,\lambda_3) \in \mathscr{A}$, in the perturbation model,

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \delta(t)dt + \sigma_s\sqrt{v_\varepsilon(t)}dW_1(t)^Q, \\ dy(t,T) &= (\mu_y(t,T) - \lambda_2 v(t))dt + \sigma_y(t,T)\sqrt{v_\varepsilon(t)}dW_2(t)^Q, \\ dv_\varepsilon(t) &= (\kappa_1 - \kappa_2 v_\varepsilon(t))dt + \sigma_v(t,T)\varepsilon\sqrt{v_\varepsilon(t)}dW_3(t)^Q. \end{aligned} \tag{1.24}$$

[10] The discussion on the general situation is also possible, see Romano and Touzi (1997).

Therefore, the call option's arbitrage price in this perturbation model under Q is written as $\mathbb{E}^{\kappa_1,\kappa_2}\left[C(\sqrt{V(\kappa_1,\kappa_2,\varepsilon)})|\{v_\varepsilon(t)\}\right]$, where

$$V(\kappa_1,\kappa_2,\varepsilon)=\int_0^{T'} v(t;\kappa_1,\kappa_2,\varepsilon)dt. \tag{1.25}$$

$\mathbb{E}^{\kappa_1,\kappa_2}[\cdot|\{v_\varepsilon(t)\}]$ represents the conditional (risk-neutral) expectation under a specific choice of the market price of risk parameters $\{\lambda_1,\lambda_2,\lambda_3\}$ given the path $\{v_\varepsilon(t)\}$. The difference between prices of European call option in the two models is

$$Diff\equiv\mathbb{E}^{\kappa_1,\kappa_2}[C(\sqrt{V(\kappa_1,\kappa_2,\varepsilon)})|\{v_\varepsilon(t)\}]-C(\sqrt{V(\kappa_1,\kappa_2,0)}) \tag{1.26}$$

By using the second-order Taylor expansion, the difference is written as

$$\begin{aligned} Diff &= \mathbb{E}^{\kappa_1,\kappa_2}\left[C'\left(\sqrt{V(\kappa_1,\kappa_2,0)}\right)\left(\sqrt{V(\kappa_1,\kappa_2,\varepsilon)}-\sqrt{V(\kappa_1,\kappa_2,0)}\right)|\{v_\varepsilon(t)\}\right] \\ &\quad +\frac{1}{2}\mathbb{E}^{\kappa_1,\kappa_2}\left[C''\left(\sqrt{V(\kappa_1,\kappa_2,0)}\right)\left(\sqrt{V(\kappa_1,\kappa_2,\varepsilon)}-\sqrt{V(\kappa_1,\kappa_2,0)}\right)^2|\{v_\varepsilon(t)\}\right] \\ &\quad +\text{higher order terms} \end{aligned}$$

where C' and C'' denote the first and second-order derivatives of function $C(V)$. Naturally, we use the the rate of convergence at which the call option price in the model with perturbations convergence to that in the benchmark model when $\varepsilon\sigma_v\to 0$ to capture the parameter uncertainty concern on ε. Let

$$F(\kappa_1,\kappa_2)=\lim_{\varepsilon\sigma_v\downarrow 0}\frac{\mathbb{E}^{\kappa_1,\kappa_2}\left[C\left(\sqrt{V(\kappa_1,\kappa_2,\varepsilon)}\right)|\{v_\varepsilon(t)\}\right]-C\left(\sqrt{V(\kappa_1,\kappa_2,0)}\right)}{\varepsilon^2\sigma_v^2}. \tag{1.27}$$

The rate of convergence is given by

$$\begin{aligned} F(\kappa_1,\kappa_2) &= \left| \frac{C''\left(\sqrt{V(\kappa_1,\kappa_2,0)}\right)}{8V(\kappa_1,\kappa_2,0)} - \frac{C'\left(\sqrt{V(\kappa_1,\kappa_2,0)}\right)}{8(V(\kappa_1,\kappa_2,0))^{3/2}} \right| \\ &\quad \times \int_0^{T'} \frac{1}{(\kappa_2)^2}\left[1-e^{\kappa_2(s-T')}\right]^2 \left[\left(\sigma_0^2 - \frac{\kappa_1}{\kappa_2}\right)e^{-k_2 s} + \frac{\kappa_1}{\kappa_2}\right] ds. \end{aligned}$$

See Proof.

The function $F(\kappa_1,\kappa_2)$ captures the convergent speed at which the option price from the model with perturbations converges to the option price from the benchmark model when ε changes. This function depends on the model parameters κ_1 and κ_2. In our model, κ_1 is known, while κ_2 is a function of the volatility risk premium λ_3 which is not unique in an incomplete market. Note that only the variance risk of premium is involved in those parameters, the robust approach is reduced to be a one-dimensional optimization problem for λ_3. The robust approach is try to find the optimal volatility risk premium λ_3^* which satisfies:

$$\frac{\partial F(\kappa_1,\kappa_2)}{\partial \lambda_3}\Big|_{\lambda_3^*} = 0$$

By intuition, the optimal volatility risk premium λ_3^* is the position when the convergent speed function $F(\kappa_1,\kappa_2)$ most robust with respect to the volatility risk premium. If λ_3^* changes a little, $F(\kappa_1,\kappa_2)$ will change smallest accordantly. That means the difference between the option price and the benchmark price is most robust with respect to the volatility risk premium.

So far, we have illustrated two approaches to choose the benchmark SDF and robust S-

DF when the parameter uncertainty is absent or in the presence of the parameter uncertainty concern, respectively. Therefore, we are able to estimate the parameter uncertainty premium by distracting the risk premium without parameter uncertainty from the risk premium with parameter uncertainty consideration. Furthermore, the implied volatilities can be calculated using the two optimal SDFs. Hence, we can also estimate the parameter uncertainty component that contribute to the implied volatility.

We next move to the empirical studies of our approaches.

1.4 Estimation and Empirical Results

In this section we first illustrate the data and the calibration. Then we present our empirical results on the uncertainty premium by comparing the benchmark approach and the robust approach. At last, we estimate the uncertainty component in the implied volatility.

A. The Sample

We estimate the model specification based on a data set of sweet crude oil future contracts on the NYMEX. The raw data is collected from Bloomberg, from February 13rd, 2006 to July 9th, 2010, on settlement prices, open interests, and daily volume for all available futures. The options data are used to compare the market implied volatility with the implied volatility derived from the model. The time to expiration of futures contract ranges from one month to one year. To minimize the effect of liquidity risk and interest rate on estimation, we use futures contract with one month, two months and three months to expiration to calibrate the parameters in the model.

The historical volatilities of the front month futures price in the sample are displayed in Figure 1.The historical volatilities range from 16.7% to 120%. As observed, the historical volatilities during the 2008 oil crisis time is the most volatile time period.

B. Parameter Calibration

In calibration we follow the methodology presented in Trolle and Schwartz (2009). We first express futures price as a linear combination of several states, then we use Kalman filter in conjunction with EM algorithm to estimate all model parameters. While EM algorithm is robust to the initial value of the data, this method, however, is sensitive to the initial values of all parameters. Therefore, we choose different initial values of parameters to obtain stable model parameters to minimize the calibration risk.

Several assumptions are made for calibration as in Trolle and Schwartz (2009). We specify $\sigma_y(t,T)$ as time-homogeneous with respect to time to maturity $T-t$, and $\sigma_y(t,T) = \alpha e^{-\beta(T-t)}$. Intuitively, the long-term volatility of carrying costs should be less than the short-term volatility. Moreover, the initial forward cost of the carrying curve is flat. Therefore, $y(0,t) = \zeta$ and $lnF(0,T) - lnF(0,t) = \zeta(T-t)$.

Under the specification of $\sigma_y(t,T)$, the futures price can be expressed as follows (See Proof.):

$$lnF(t,T) = \zeta(T-t) + s(t) + \frac{\alpha}{\beta}\left(1 - e^{-\beta(T-t)}\right)x(t) + \frac{\alpha}{2\beta}\left(1 - e^{-2\beta(T-t)}\right)\phi(t) \quad (1.28)$$

where the state variables $\{x(t), \phi(t), s(t), v(t)\}$ satisfy

$$\begin{cases} dx(t) = \left[-\beta x(t) - (\frac{\alpha}{\beta} + \rho_1 2\sigma_s - \lambda_2)v(t)\right]dt + \sqrt{v(t)}dW_2(t)^Q \\ d\phi(t) = \left[-2\beta\phi(t) + \frac{\alpha}{\beta}v(t)\right]dt \\ ds(t) = \left[\zeta + \alpha x(t) + \alpha\phi(t) - \frac{1}{2}\sigma_s^2 v(t)^2\right]dt + \sigma_s v(t)dW_1(t)^Q \\ dv(t) = (\gamma - kv(t) - \lambda_3\sigma_v v(t))dt + \sigma_v\sqrt{v(t)}dW_3(t)^Q, \end{cases} \quad (1.29)$$

The estimations of all the parameters are reported in Table 1. Compared with the estimations in SV1 model of Trolle and Schwartz (2009), our estimations are similar to theirs. Specially ρ_{12} is significant negative, which means that the crude oil price is inversely correlated with the carrying costs. When carrying costs increase, investors do not want to keep the crude oil, the demand will decrease.

C. Empirical Findings

We report two types of numerical results in this section. The first type is about the risk premiums. Firstly, We estimate the volatility risk premium by using a benchmark approach when there is no parameter uncertainty. We also estimate the robust volatility risk premium by a robust approach when parameter uncertainty is pervasive. At last, we estimate the uncertainty risk premium in the unspanned stochastic volatility model.

The second type of numerical results is on the implied volatility. As the implied volatilities are calculated via both approaches, we estimate the uncertainty component in the implied volatility, termed as ``uncertainty volatility component". In what follows we first summarize the main empirical findings and next present further details for each finding in this section.

(1) Parameter uncertainty contributes to an uncertainty risk premium (URP), and the uncertainty risk premium is negative in many situations.

(2) Parameter uncertainty generates a positive uncertainty volatility component (UVC) in the implied volatility.

(3) The effect of parameter uncertainty on the implied volatility is more significant when the initial volatility is lower.

D. Uncertainty Risk Premiums

Table 1.2 reports the volatility risk premium via the benchmark approach and the robust approach for futures option with different strike price K, different initial volatility v_0 and different set of time to maturity of futures option and futures contract. The option parameters are $F = 50$ and the option moneyness $\frac{K}{F}$ is chosen between 80% and 120%. The range of the volatility risk premium is $[-2,2]$ which is large enough for us to get an global optimal estimation of the volatility risk premium. As Trolle and Schwartz (2009) consider two initial values $v_0 = 1$ and $v_0 = 5$, to be consistent with the parameters in Trolle and Schwartz (2009), we investigate $v_0 = 1,2,5$ in both benchmark and robust approach. [11] We choose relative long time to maturity of futures option and futures contract with several considerations. Firstly, when the time to maturity of futures option and futures contract are long, we are worried about the parameter uncertainty of the model. We all know that futures option and futures contract with short time to maturity are liquid in the market. Because of the illiquid trading of futures option and futures contract with long time to maturity, we do not have enough information to precisely estimate those long-term contracts. Compared with those contracts with short time to maturity, parameter uncertainty is a important issue when we price those long-term contract. Secondly, from proposition 2, the robust approach only can be used when the volatility of $v(t)$, $\sigma_v\varepsilon$ is very small. Intuitively, long-term volatility of $v(t)$ should be smaller than short-term volatility of $v(t)$. Using the robust approach to choose an optimal volatility risk premium for a long-term contract is more comfortable. [12] To get a comprehensive understanding of the potential effect of parameter uncertain-

[11] Other initial values of v_0 and contract parameters (T', T) are implemented while not reported here. Those numerical results are available from authors and they all support the above mentioned empirical findings.

[12] Although interest rate may have an unnegligiable impact on the long-term contracts, in this paper, we only focus on the effect of parameter uncertainty on option price. A consideration of a stochastic interest rate model maybe an extension of this model.

ty on pricing derivatives, we consider five future options: $(T',T) = (1,2), (1.5,2)$, $(1,4)$, $(1.5,4)$ and $(2,4)$where T is the time to maturity of the future contract and T' is the option's maturity.

In Table 1.2, optimal volatility risk premiums via the benchmark approach with different moneyness, different expirations and different initial volatility are reported in the last three columns. We denote the optimal volatility risk premiums by the benchmark approach as the benchmark volatility risk premium. All those benchmark volatility premiums are positive. We also can find that the benchmark volatility risk premium decreases as the initial volatility v_o increases.[13] We define the optimal volatility risk premium via the robust approach as the robust volatility risk premium which are displayed in the left part of table 2. All those robust volatility risk premiums are negative. The robust volatility risk premium increases in value but decreases in absolute value as the initial volatility increases.

Recent literatures have documented that the volatility risk premiums are negative from both the equity market as well as the commodity market. By checking the performance of delta-hedged portfolios in equity market, Bakshi and Kapadia (2003) find a negative sign of volatility risk premium. Doran and Ronn (2008) find that the negative volatility risk premium can explain the disparity between risk-neutral and statistical volatility in both equity and commodity-energy markets. Carr and Wu (2009) employ a model-free method by directly checking the difference between the realized variance and the synthetic variance swap rate, and find a negative volatility risk premium in equity market. Finally, Christoffersen, Heston and Jacobs (2010) derive a theoretical result that the volatility risk premium is negative when a class of market price of volatility risk is considered. However, there is no clear explanation on the negative volatility risk in those empirical and theoretical findings.

[13]($(T',T) = (2,4)$ is a special case, the benchmark volatility risk premiums hit the boundary.

In our paper, the benchmark volatility risk premium is positive while the robust volatility risk premium is negative. The difference between the benchmark and robust volatility risk premium is the effect of parameter uncertainty on the chosen of volatility risk premium. It is highly possible that the negative sign of volatility risk premium is caused by parameter uncertainty or is at least partly caused by parameter uncertainty. The intuition is straight forward. Because of the extra risk of parameter uncertainty in pricing a security, we need to compensate the extra risk. When the volatility premium decreases, the stochastic volatility will not converge to the mean very fast. Even when the volatility risk premium is negative enough in magnitude, the drift of the stochastic volatility will increase. The option price will increases as the volatility increases. That means we need to pay extra money to buy an option contract which has already taken account of the model uncertainty.

By comparing the robust volatility risk premium parameter λ_3 from the benchmark approach with that from the robust approach, we also see that the robust volatility risk premium parameter λ_3 from robust approach is smaller than that from the benchmark approach. Therefore, the uncertainty risk premium (URP), which is defined as the difference between the volatility risk premiums in both approaches, is also negative. Moreover, the magnitude of the uncertainty risk premium relative to the stochastic volatility risk premium is fairly large. Because the benchmark volatility risk premium is decreasing and the robust volatility risk premium is increasing as the initial volatility increases, the URP decreases as the initial volatility increases.

E. Uncertainty Volatility Component

After obtaining the optimal market risk premiums from the benchmark approach and the robust approach, we now compute the implied volatilities of the European call option with different strike price using these market risk premiums. As the parameters $\{\lambda_1, \lambda_2\}$ are not involved in the future option formula (see equation (1.23)), we make use of the market volatility risk premium parameter λ_3 estimated in Table 1.2. Then we compare the implied volatilities from these two approaches to illustrate the parameter uncertainty component in the implied volatility.

We report the implied volatilities from both approaches and the difference in both absolute and relative aspects. Figure 1.3, Panel A-B-C display the implied volatilities from the benchmark approach and the robust approach when $T = 2, T' = 1$, and different initial volatility value v_0. The solid line represents the implied volatilities from the robust approach, while the dash line represents the implied volatilities from the benchmark approach. Similarly, in Figure 1.4 to Figure 1.7, Panel A-B-C displays the implied volatilities when $(T', T) = (1.5, 2), (1, 4), (1.5, 4)$ and $(2, 4)$ when $v_0 = 1, 2 and 5$ respectively.

As demonstrated, there are some interesting relationships between the implied volatilities of both approaches. Therefore, the uncertainty component in the implied volatility can be estimated and investigated numerically. First, the implied volatilities from the robust approach are higher than those from the benchmark approach. Therefore, the uncertainty volatility component is positive in the implied volatility. This empirical result is intuitive appealing. As the agent faces the parameter uncertainty, a higher price is required to compensate the uncertainty concern. Therefore, the implied volatility is increased at the presence of parameter uncertainty.

Second, the implied volatilities depend on the initial volatility v_0. We find that the higher the initial volatility v_0 is, the higher the expected implied volatility will be. In fact,

$$\mathbb{E}[v_t] = (v_0 - \frac{k_1}{k_2})e^{-k_2 t} + \frac{k_1}{k_2}, \tag{1.30}$$

thus the expected volatility $\mathbb{E}[v_t]$ is positively proportional to the initial volatility v_0.

We now see the relative difference of implied volatility and how much the uncertainty contribute in proportion to the implied volatility. We use a proportion to describe how large the proportion of implied volatilities comes from parameter uncertainty. The proportion of the uncertainty premium to the implied volatility of the benchmark approach, is defined as $(\mathrm{IV}^R - IV^B)/IV^B, where\mathrm{IV}^R$ is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. These proportions are shown in Figure 1.3 - Figure 1.7, Panel D-E-F, for $(T', T) = (1,2), (1.5,2), (1,4), (1.5,4)$ and $(2,4)$, respectively.

As displayed in Figure 1.3-1.7, increase implied volatility is associated with parameter uncertainty. Actually, the percentage of the uncertainty component moves between 6% and 70%. These uncertainty components are fairly large that the effect of model uncertainty in pricing derivatives could be substantial. Consequently, the option would be undervalued significantly without incorporating the uncertainty component. Carr and Wu (2009) document empirically that the realized volatilities are lower than the implied volatilities from the market data and illustrate this phenomenon by the demand of hedging downside risk from the market. Given the significant effect of parameter uncertainty on the implied volatilities in the commodity market as we have shown above, the parameter uncertainty might explain the difference between the realized volatility and the implied volatility from the real market data.

Finally, the effect of the model uncertainty is more significant when the initial volatility lower. Indeed, when the initial volatility $v_0 = 1$, in Figure 1.3-1.7, the proportion reaches its highest value control for T and T'. Intuitively, when the initial volatility is low, the expected volatility of underlying asset should also be low. Because option price is positive related with the volatility of underlying asset, the option price is relative lower when the initial volatility is lower. Control for the absolute effect of parameter uncertainty on option price, the relative effect of parameter uncertainty will be maximized as the option price is lowest.

1.5 Conclusion

We have examined a robust approach in a general class of stochastic volatility models with model (parameters) uncertainty in a commodity market. We document that the concern on the model parameters estimation risk adds significant component on the commodity contingent claim and the implied volatility. We demonstrate the existence of a negative uncertainty risk premium and this negative model uncertainty premium contributes a positive component in the implied volatility. In other words, the implied volatility is large when the parameter uncertainty is strong.

The robust approach developed in this paper can be used in other financial markets, such as the fixed income market and the equity market. The presented methodology extends the one developed in Boyle, Feng, Tian and Wang (2008) in which a robust approach is first proposed for a class of stochastic volatility model. When a large price movements with small probability event is considered, the robust approach based on large deviation principle in Boyle, Feng, Tian and Wang (2008), can be used to study the robust risk management strategy for commodity contingent claim. We leave this topic for further study.

Proof of Proposition 1.4.

This problem is reduced to a sequence of optimization problem in the mean-variance framework as follows:

$$\min_{\mathbb{E}_P[z_\alpha|\mathscr{F}]=c,\mathbb{E}_P[m_\alpha z_\alpha|\mathscr{F}]=0} Var[(p+z_\alpha)|\mathscr{F}]$$

where c is one number. We omit the script $\mathscr{F}$ in the following if no confusions. Let

$$\mathscr{L}=\frac{1}{2}\sum P(\omega)(p(\omega)+z_\alpha(\omega))^2-\lambda\left(\sum P(\omega)z_\alpha(\omega)-c\right)-\gamma\sum Q^\alpha(\omega)z_\alpha(\omega).$$

Solving the first-order equations we obtain that $\lambda+\gamma=\mathbb{E}_P[p]+c, \lambda+\gamma\sum\frac{Q^\alpha(\omega)^2}{P(\omega)}=\mathbb{E}_P[m_\alpha p]$, and

$$z_\alpha(\omega)=\lambda+\gamma\frac{Q^\alpha(\omega)}{P(\omega)}-p(\omega). \tag{1.31}$$

Solving λ, γ, we obtain

$$\begin{aligned} z_\alpha(c,\omega) &= \mathbb{E}_P[p]-p(\omega)+c \\ &\quad +\{\mathbb{E}_P[m_\alpha p]-\mathbb{E}_P[p]-c\}\left(\frac{Q^\alpha(\omega)}{P(\omega)}-1\right)\left(\mathbb{E}_P\left[\left(\frac{dQ^\alpha}{dP}\right)^2\right]-1\right)^{-1}. \end{aligned}$$

Then, the approximate derivative under the SDF m_α is

$$z_\alpha(c)+p=\mathbb{E}_P[p]+c+\frac{\mathbb{E}_P[m_\alpha p]-\mathbb{E}_P[p]-c}{S^2}\left(\frac{dQ^\alpha}{dP}-1\right)$$

where $S^2 = \mathbb{E}_P[(\frac{dQ^\alpha}{dP})^2] - 1$. Moreover, it is straightforward to see that

$$\frac{(\mathbb{E}_P[p+z_\alpha] - \mathbb{E}_P[m_\alpha(p+z_\alpha)])^2}{Var[p+z_\alpha]} = S^2. \tag{1.32}$$

The proof is finished.

Proof of Proposition 1.3.

To simplify notation, we will suppress the arguments, κ_1 and κ_2, of $v(t;\kappa_1,\kappa_2,\varepsilon)$,

$$v(t,\varepsilon) = v(t,0) + \sigma_v \varepsilon \int_0^t e^{\kappa_2(s-t)} \sqrt{v(s,\varepsilon)} dW_3(s)^Q \tag{1.33}$$

By stochastic Fubini's theorem, we have

$$V(\varepsilon) = V(0) + \varepsilon \sigma_v y, \tag{1.34}$$

where

$$y = \int_0^{T'} \frac{1}{\kappa_2}[1 - e^{\kappa_2(s-T')}] \sqrt{v(s,\varepsilon)} dW_3(s)^Q \tag{1.35}$$

let $h(V) = C(\sqrt{V})$, then

$$h'(V) = C'(\sqrt{V}) \frac{1}{2\sqrt{V}}, \tag{1.36}$$

$$h''(V) = \frac{C''(\sqrt{V})}{4V} - \frac{C'(\sqrt{V})}{4V^{\frac{3}{2}}}, \tag{1.37}$$

and

$$h'''(V) = C'''(\sqrt{V})\frac{1}{8V^{\frac{3}{2}}} - C(\sqrt{V})''(\sqrt{V})\frac{3}{8V^2} + C(\sqrt{V})'\frac{3}{8V^{\frac{5}{2}}}. \quad (1.38)$$

Moreover, we have

$$C'(\sqrt{V}) = \frac{\sqrt{B}F}{\sqrt{2\pi}}e^{-\frac{d_1^2}{2}},$$
$$C''(\sqrt{V}) = \frac{1}{\sqrt{2\pi}}Fe^{-\frac{d_1^2}{2}}\left[\frac{(ln\frac{F}{K})^2}{\sqrt{B}V^{\frac{3}{2}}} - \frac{1}{4}B^{\frac{3}{2}}\sqrt{V}\right], \quad (1.39)$$

$$\begin{aligned} C'''(\sqrt{V}) &= \frac{1}{\sqrt{2\pi}}Fe^{-\frac{d_1^2}{2}}[\frac{\left(ln\frac{F}{K}\right)^4}{B^{\frac{3}{2}}V^3} - \frac{1}{2}\frac{\sqrt{B}}{V}\left(ln\frac{F}{K}\right)^2 \\ &\quad + \frac{1}{16}B^{\frac{5}{2}}V - \frac{3\left(ln\frac{F}{K}\right)^2}{\sqrt{B}V^2} - \frac{1}{4}B^{\frac{3}{2}}]. \end{aligned} \quad (1.40)$$

When $ln\frac{F}{K} \neq 0$[14], it is readily seen that

$$\lim_{\sigma\to 0}\frac{C^{(k)}(\sigma)}{\sigma^n} = \lim_{\sigma\to\infty}\frac{C^{(k)}(\sigma)}{\sigma^n} = 0, \quad (1.41)$$

for $k = 1,2,3$ and $n \geq 1$. Hence

$$\lim_{V\to 0}h'''(V) = \lim_{\sigma\to\infty}h'''(V) = 0. \quad (1.42)$$

Since the function $h(V) \in C^\infty((0,\infty))$, there exists a positive real number K such that

$$|h'''(V)| \leq K, \quad (1.43)$$

[14] When $ln\frac{F}{K} = 0$, the proof is similar to the argument in Boyle, Feng, Tian and Wang (2008).

for all $V > 0$, By Taylor expansion, we have

$$\begin{aligned} h(V(\varepsilon)) &= h(V(0)) + h'(V(0))(V(\varepsilon) - V(0)) + \frac{1}{2}h''(V(0))\,(V(\varepsilon) - V(0))^2 \\ &\quad + \frac{1}{6}h'''(\theta_\varepsilon)\,(V(\varepsilon) - V(0))^3 \end{aligned} \tag{1.44}$$

where $|\theta_\varepsilon - V(0)| \leq |V(\varepsilon) - V(0)|$. Taking expectation yields

$$\begin{aligned} \mathbb{E}[h(V(\varepsilon))] &= h(V(0)) + h'(V(0))\mathbb{E}[V(\varepsilon) - V(0)] \\ &\quad + \frac{1}{2}h''(V(0))\mathbb{E}[(V(\varepsilon) - V(0))^2] \\ &\quad + \frac{1}{6}\mathbb{E}[h'''(\theta_\varepsilon)\,(V(\varepsilon) - V(0))^3]. \end{aligned} \tag{1.45}$$

By Cauchy-Schwartz inequality,

$$\mathbb{E}[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3]^2 \leq \mathbb{E}[h'''(\theta_\varepsilon)^2]\mathbb{E}[(V(\varepsilon) - V(0))^6]. \tag{1.46}$$

As $h'''(V)$ is bounded, we obtain

$$\mathbb{E}[h'''(\theta_\varepsilon)\,(V(\varepsilon) - V(0))^3] \leq K\sqrt{\mathbb{E}[(V(\varepsilon) - V(0))^6]} \tag{1.47}$$

It is easy to show that

$$\begin{aligned} \mathbb{E}[V(\varepsilon) - V(0)] &= 0, \\ \mathbb{E}[(V(\varepsilon) - V(0))^2] &= (\varepsilon\sigma_v)^2\mathbb{E}[y^2], \\ \mathbb{E}[(V(\varepsilon) - V(0))^6] &= o((\varepsilon\sigma_v)^6), \end{aligned} \tag{1.48}$$

then

$$\begin{aligned}\mathbb{E}\left[C\left(\sqrt{V(\varepsilon)}\right)\Big|\mathscr{F}\right] &= C\left(\sqrt{V(0)}\right)+(\varepsilon\sigma_v)^2\left[\frac{C''\left(\sqrt{V(0)}\right)}{8V(0)}-\frac{C'\left(\sqrt{V(0)}\right)}{8(V(0))^{3/2}}\right]\mathbb{E}[y^2] \\ &+o((\varepsilon\sigma)^2).\end{aligned}$$

Finally,

$$y=\int_0^{T'}\frac{1}{\kappa_2}[1-e^{\kappa_2(s-T')}]\sqrt{v(s,\varepsilon)}dW_3(s)^Q \tag{1.49}$$

then

$$\mathbb{E}[y^2]=\int_0^{T'}\frac{1}{(\kappa_2)^2}\left[1-e^{\kappa_2(s-T')}\right]^2\left[(\sigma_0^2-\frac{\kappa_1}{\kappa_2})e^{-k_2 s}+\frac{k_1}{k_2}\right]ds \tag{1.50}$$

Then we derive the expression of $F(\kappa_1,\kappa_2)$ as stated in the Proposition. □

Derivation of Equations (1.9)-(1.11).

Let $Y(t,T)=\int_t^T y(t,u)du$, then by Ito's lemma,

$$\begin{aligned}dY(t,T) &= \left[-\delta(t)+\int_t^T(\mu_y(t,u)-\lambda_2\sigma_y(t)v(t))du\right]dt \\ &+\sqrt{v(t)}\int_t^T\sigma_y(t,u)dudW_2(t)^Q.\end{aligned} \tag{1.51}$$

The future price is $F(t,T)=S(t)e^{Y(t,T)}$. By Ito's lemma again,

$$\frac{dF(t,T)}{F(t,T)}=\frac{dS(t)}{S(t)}+dY(t,T)+\frac{1}{2}[dY(t,T),dY(t,T)]_t+\frac{[dS(t),dY(t,T)]_t}{S(t)}. \tag{1.52}$$

By using the spot price process and the forward carry cost process, we obtain

$$\begin{aligned}\frac{dF(t,T)}{F(t,T)} &= \int_t^T [\mu_y(t,u) - \lambda_2\sigma_y(t,u)v(t)]\,dudt \\ &+v(t)\left[\frac{1}{2}\left(\int_t^T \sigma_y(t,u)du\right)^2 + \rho_{12}\sigma_s\int_t^T \sigma_y(t,u)du\right]dt \\ &+\sigma_s\sqrt{v(t)}dW_1(t)^Q + \sqrt{v(t)}\int_t^T \sigma_y(t,u)dudW_2(t)^Q. \end{aligned} \tag{1.53}$$

Setting the drift term equal to zero under the risk neutral measure $Q^{\lambda_1,\lambda_2,\lambda_3}$, (1.9) is proved. Moreover, by differentiating with respect to T in the drift term of the last equation, we get

$$\mu_y(t,T) = -v(t)\sigma_y(t,T)[\rho_{12}\sigma_s + \int_t^T \sigma_y(t,T)du - \lambda_2].$$

For the continence yield, we note that

$$dy(t,T) = [\mu_y(t,T) - \lambda_2\sigma_y(t,T)]v(t)dt + \sigma_y(t,T)\sqrt{v(t)}dW_2(t)^Q. \tag{1.54}$$

Take the integration from 0 to t on both side, we get

$$\int_0^t dy(u,T) = \int_0^t (\mu_y(u,T) - \lambda_2\sigma_y(u,T)v(u)du + \int_0^t \sigma_y(u,T)\sqrt{v(u)}dW_2(u)^Q. \tag{1.55}$$

Therefore the equation (1.11) is proved.

Derivation of Equations (1.12)-(1.16).

Write

$$\begin{aligned}\mathbb{E}[Z_T^2] &= \mathbb{E}\left[exp\left(-2\sum_i^3\int_0^T \lambda_i\sqrt{v(s)}dW_i(s) - \int_0^T\sum_i^3\sum_j^3\rho_{ij}\lambda_i\lambda_j v(s)ds\right)\right] \\ &= \mathbb{E}^{Q^*}\left[exp\left(\int_0^T\sum_i^3\sum_j^3\rho_{ij}\lambda_i\lambda_j v(s)ds\right)\right].\end{aligned}$$

where Q^* is the measure under which $W_i(t)^{Q^*} = W_i(t)^P + \int_0^t 2\lambda_i\sqrt{v(s)}ds$ is the Brownian motion. Under Q^* measure, $dv(s) = (\kappa_1 - (\kappa_2 + 2\lambda_3)v(s))ds + \sigma_v\sqrt{v(s)}dW_3^*(s)$, Let $\kappa_1 = \gamma, \kappa_2^* = \kappa_2 + 2\lambda_3$. By the Feynman-Kac theorem,$\mathbb{E}\left[Z_t^2\right]$ satisfies the following PDE given that $v(0) = x$,

$$u(t,x)_t + (\kappa_1 - \kappa_2^* x)u(t,x)_x + \frac{1}{2}\sigma_v^2 x u(t,x)_{xx} + \lambda x u(t,x) = 0, u(0,x) = 1. \tag{1.56}$$

It is standard from affine model (see Duffie, Singleton and Pan (2000)) that the solution is in the form of $u(t,x) = e^{A(t)-xB(t)}$, and $A(t), B(t)$ are determined by the following two ordinary differential equations:

$$\frac{\partial A(t)}{\partial t} - B(t)\kappa_1 = 0,\ \frac{\partial B(t)}{\partial t} = B(t)\kappa_2^* + \frac{1}{2}\sigma_v^2 B(t)^2 + \lambda. \tag{1.57}$$

In our numerical examples $2\lambda\sigma_v^2 > \kappa_2^{*2}$. It is straightforward to derive $A(t)$ and $B(t)$ as expressed in the text under the condition that $2\lambda\sigma_v^2 > \kappa_2^{*2}$.

Derivation of Equations (1.28)-(1.29)

With the specification $\sigma_y(t,T) = \alpha e^{-\beta(T-t)}$ and using the equation (1.11), we have

$$\begin{aligned} y(t,T) &= y(0,T) + \int_0^t v(u)[\frac{\alpha^2}{\beta}e^{-2\beta(T-u)} + (\lambda_2 - \frac{\alpha}{\beta} - \rho_{12}\sigma_s) \\ &\quad \alpha e^{-\beta(T-u)} - \lambda_2 e^{-\beta(T-u)}]du + \int_0^t \alpha e^{-\beta(T-u)}\sqrt{v(u)}dW_2(u)^Q. \end{aligned}$$

Alternatively,

$$y(t,T) = y(0,T) + \alpha e^{-\beta(T-t)}x(t) + \alpha e^{-2\beta(T-t)}\phi(t) \tag{1.58}$$

where the state variables $\{x(t), \phi(t)\}$ satisfy

$$dx(t) = \left[-\kappa x(t) - (\frac{\alpha}{\beta} + \rho_{12}\sigma_s - \lambda_2)v(t)\right]dt + \sqrt{v(t)}dW_2(t)^Q \tag{1.59}$$

$$d\phi(t) = \left[-2\beta\phi(t) + \frac{\alpha}{\beta}v(t)\right]dt \tag{1.60}$$

subject to $x(0) = \phi(0) = 0$,

Moreover, denote $s(t) = lnS(t)$, we have

$$ds(t) = [y(0,t) + \alpha x(t) + \alpha\phi(t) - \frac{1}{2}\sigma_s^2 v(t)^2]dt + \sigma_s v(t)dW_1(t)^Q \tag{1.61}$$

Moreover,

$$\begin{aligned} lnF(t,T) &= lnF(0,T) - lnF(0,t) + s(t) + \frac{\alpha}{\beta}(1 - e^{-\beta(T-t)}x(t) \\ &+ \frac{\alpha}{2\beta}(1 - e^{-2\beta(T-t)})\phi(t). \end{aligned}$$

We now assume that the initial forward cost of carry curve is flat, $y(0,t) = \zeta$ and $lnF(0,T) - lnF(0,t) = \zeta(T-t)$. Then we have

$$lnF(t,T) = \zeta(T-t) + s(t) + \frac{\alpha}{\beta}(1 - e^{-\beta(T-t)})x(t) + \frac{\alpha}{2\beta}(1 - e^{-2\beta(T-t)})\phi(t). \quad (1.62)$$

The proof is finished.

Table 1.1: Parameters calibration

This table reports the estimated model parameters in the model. $\sigma_y(t,T) = \alpha e^{-\beta(T-t)}$. The estimation procedure is similar to Trolle and Schwartz (2009).

α	0.6388
σ_s	0.5367
ρ_{12}	-0.4712
β	0.7704
ζ	-0.0036
σ_v	2.8931
κ	1.1376
γ	0.5401

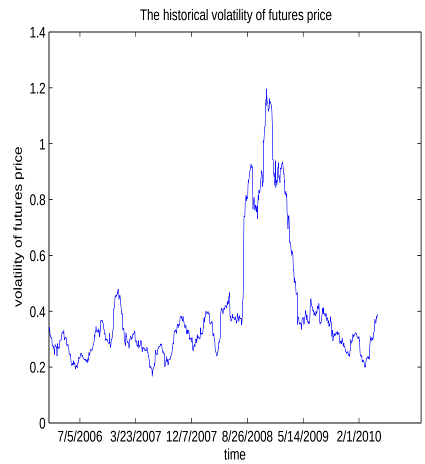

Figure 1.1: Historical future price volatility

This figure displays the historical future price volatility by using the historical front month futures contract data from Feb 2006 to July 2010. We use the front month futures contract in the database to calculate historical volatilities which actually are 30-day historical volatilities.

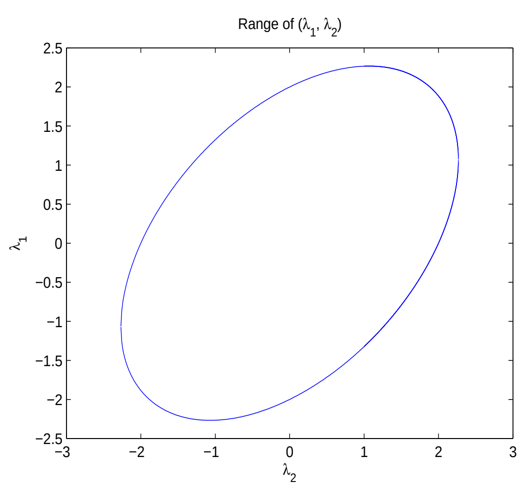

Figure 1.2: Region of risk premiums

This figure displays the region of $\{\lambda_1, \lambda_2\}$ which is bounded by $\lambda_1^2 + 2\rho_{12}\lambda_1\lambda_2 + \lambda_2^2 \leq 4$. $\rho_{12} = -0.4712$ which is estimated from Table 1.1. The range of λ_3 is $-2 \leq \lambda_3 \leq 2$.

Table 1.2: Volatility and uncertainty risk premium

This table reports the optimal volatility risk premiums $\{\lambda_3\}$ chosen by the robust approach and the benchmark approach for future options with different strikes and different expiration of futures and future options. In this table $(T,T') = (2,1),(2,1.5),(4,1)(4,1.5)(4,2)$, respectively, and the moneyness, defined as $\frac{K}{F}$, runs from 0.8 to 1.2. The initial value $v_0 = 1,2,5$, including the situations studied in Trolle and Schwartz (2009). URP represents the uncertainty risk premium. URP are positive numbers as shown in this table.

	$(T,T') = (2,1)$						Benchmark		
K/F	$v_0 = 1$	**URP**	$v_0 = 2$	**URP**	$v_0 = 5$	**URP**	$v_0 = 1$	$v_0 = 2$	$v_0 = 5$
0.80	-1.97	2.73	-1.63	2.04	-1.2	1.38	0.76	0.41	0.18
0.85	-1.97	2.78	-1.62	2.04	-1.2	1.38	0.81	0.42	0.18
0.90	-1.96	2.80	-1.62	2.05	-1.2	1.39	0.84	0.43	0.19
0.95	-1.96	2.82	-1.62	2.05	-1.2	1.39	0.86	0.43	0.19
1.00	-1.96	2.83	-1.62	2.06	-1.2	1.39	0.87	0.44	0.19
1.05	-1.96	2.83	-1.62	2.06	-1.2	1.39	0.86	0.44	0.19
1.10	-1.96	2.81	-1.62	2.05	-1.2	1.39	0.85	0.43	0.19
1.15	-1.96	2.79	-1.62	2.05	-1.2	1.39	0.83	0.43	0.19
1.20	-1.97	2.77	-1.62	2.04	-1.2	1.38	0.80	0.42	0.18
	$(T,T') = (2,1.5)$						Benchmark		
K/F	$v_0 = 1$	**URP**	$v_0 = 2$	**URP**	$v_0 = 5$	**URP**	$v_0 = 1$	$v_0 = 2$	$v_0 = 5$
0.80	-1.28	3.28	-1.08	2.95	-0.8	1.95	2	1.87	1.15
0.85	-1.28	3.28	-1.08	2.98	-0.8	1.95	2	1.90	1.15
0.90	-1.28	3.28	-1.08	3	-0.8	1.95	2	1.92	1.15
0.95	-1.28	3.28	-1.08	3.01	-0.8	1.95	2	1.93	1.15
1.00	-1.28	3.28	-1.08	3.01	-0.8	1.95	2	1.93	1.15
1.05	-1.28	3.28	-1.08	3.01	-0.8	1.95	2	1.93	1.15
1.10	-1.28	3.28	-1.08	3	-0.8	1.95	2	1.92	1.15
1.15	-1.28	3.28	-1.08	2.99	-0.8	1.95	2	1.91	1.15
1.20	-1.28	3.28	-1.08	2.97	-0.8	1.95	2	1.89	1.15
	$(T,T') = (4,1)$						Benchmark		
K/F	$v_0 = 1$	**URP**	$v_0 = 2$	**URP**	$v_0 = 5$	**URP**	$v_0 = 1$	$v_0 = 2$	$v_0 = 5$
0.80	-1.79	2.44	-1.49	1.88	-1.06	1.25	0.65	0.39	0.19
0.85	-1.79	2.49	-1.49	1.89	-1.06	1.25	0.70	0.40	0.19
0.90	-1.79	2.52	-1.49	1.90	-1.06	1.25	0.73	0.41	0.19
0.95	-1.79	2.54	-1.49	1.91	-1.06	1.25	0.75	0.42	0.19
1.00	-1.79	2.55	-1.49	1.91	-1.06	1.25	0.76	0.42	0.19
1.05	-1.79	2.55	-1.49	1.91	-1.06	1.25	0.76	0.42	0.19
1.10	-1.79	2.53	-1.49	1.90	-1.06	1.25	0.74	0.41	0.19
1.15	-1.79	2.51	-1.49	1.90	-1.06	1.25	0.72	0.41	0.19
1.20	-1.79	2.48	-1.49	1.89	-1.06	1.25	0.69	0.40	0.19
	$(T,T') = (4,1.5)$						Benchmark		
K/F	$v_0 = 1$	**URP**	$v_0 = 2$	**URP**	$v_0 = 5$	**URP**	$v_0 = 1$	$v_0 = 2$	$v_0 = 5$
0.80	-1.19	3.19	-0.99	2.60	-0.7	1.25	2	1.61	0.19
0.85	-1.19	3.19	-0.99	2.63	-0.7	1.80	2	1.64	1.10
0.90	-1.19	3.19	-0.99	2.65	-0.7	1.80	2	1.66	1.10
0.95	-1.19	3.19	-0.99	2.66	-0.7	1.81	2	1.67	1.11
1.00	-1.19	3.19	-0.99	2.66	-0.7	1.81	2	1.67	1.11
1.05	-1.19	3.19	-0.99	2.66	-0.7	1.81	2	1.67	1.11
1.10	-1.19	3.19	-0.99	2.65	-0.7	1.81	2	1.66	1.11
1.15	-1.19	3.19	-0.99	2.64	-0.7	1.80	2	1.65	1.10
1.20	-1.19	3.19	-0.99	2.62	-0.7	1.80	2	1.63	1.10
	$(T,T') = (4,2)$						Benchmark		
K/F	$v_0 = 1$	**URP**	$v_0 = 2$	**URP**	$v_0 = 5$	**URP**	$v_0 = 1$	$v_0 = 2$	$v_0 = 5$
$[0.8, 1.20]$	-0.89	2.89	-0.76	2.76	-0.55	2.55	2	2	2

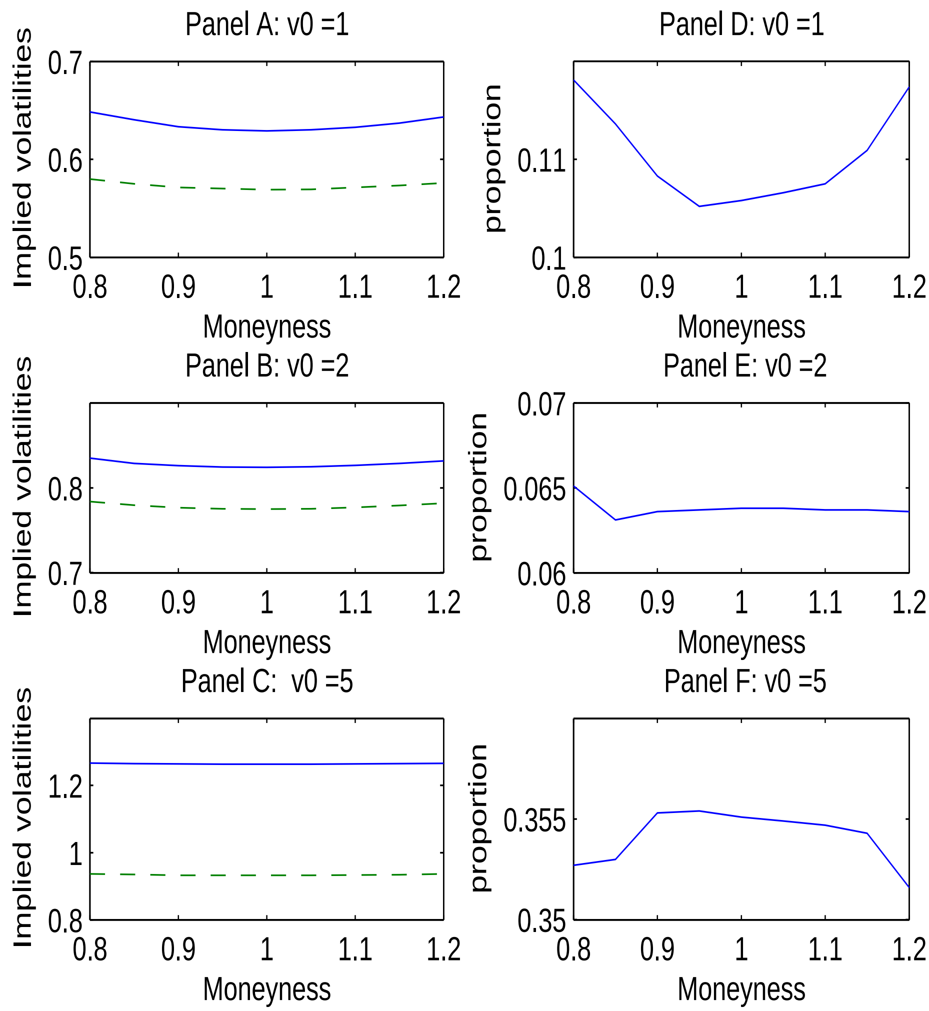

Figure 1.3: Implied volatilities via robust and benchmark approach

The graph displays the implied volatilities from the robust approach and benchmark approach, with different initial volatility v_0. The solid line represents the implied volatilities from the robust approach while the dash line describes the implied volatilities from the benchmark approach. The input parameters are $T = 2$, $T' = 1$ and $F = 50$. The option is computed where the variance is expressed via (1.25). Panel A-B-C display the implied volatilities for $v_0 = 1, v_0 = 2, v_0 = 5$, respectively. Panel D-E-F display the percentage, defined as $(IV^R - IV^B)/IV^B$, of the uncertainty component with respect to the implied volatility from the benchmark approach. IV^R is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. The initial volatility $v_0 = 1, 2$ and $v_0 = 5$ in Panel D-E-F, respectively.

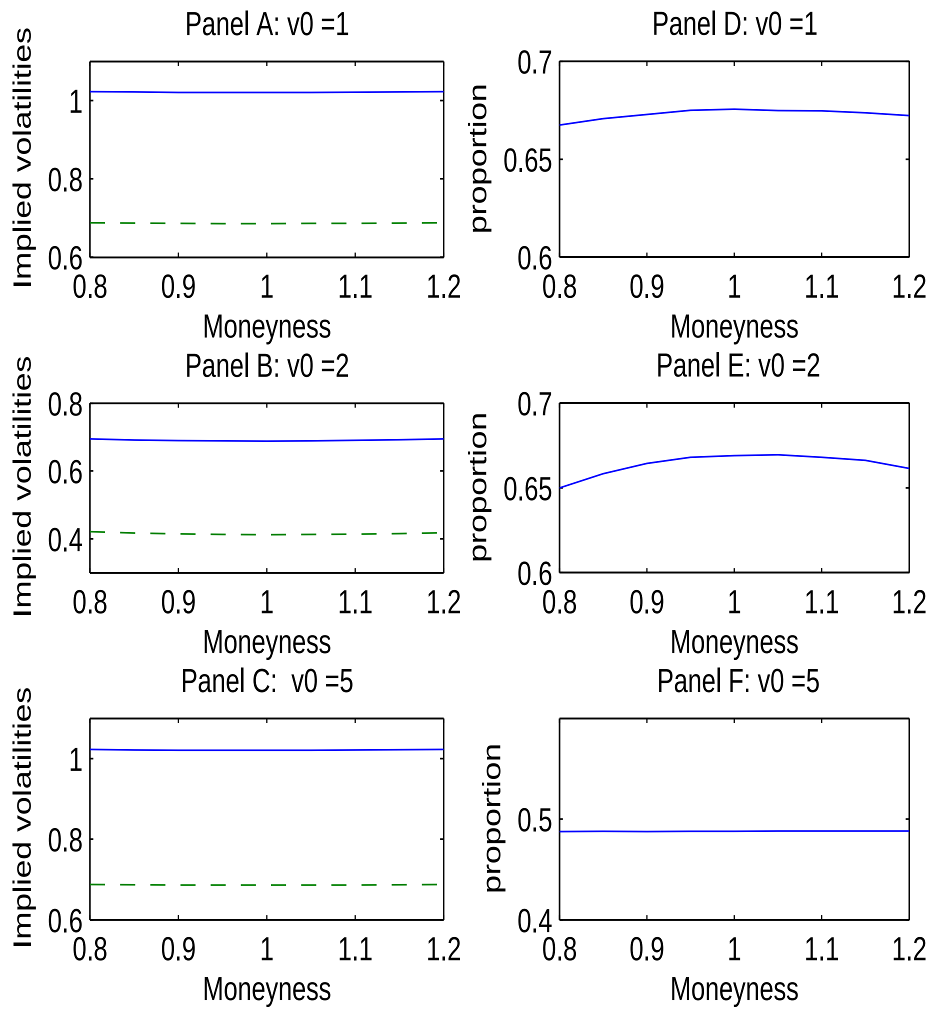

Figure 1.4: Implied volatilities via robust and benchmark approach
The graph displays the implied volatilities from the robust approach and benchmark approach, with different initial volatility v_0. The solid line represents the implied volatilities from the robust approach while the dash line describes the implied volatilities from the benchmark approach. The input parameters are $T = 2$, $T' = 1.5$ and $F = 50$. The option is computed where the variance is expressed via (1.25). Panel A-B-C display the implied volatilities for $v_0 = 1, v_0 = 2, v_0 = 5$, respectively. Panel D-E-F display the percentage, defined as $(IV^R - IV^B)/IV^B$, of the uncertainty component with respect to the implied volatility from the benchmark approach. IV^R is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. The initial volatility $v_0 = 1, 2$ and $v_0 = 5$ in Panel D-E-F, respectively.

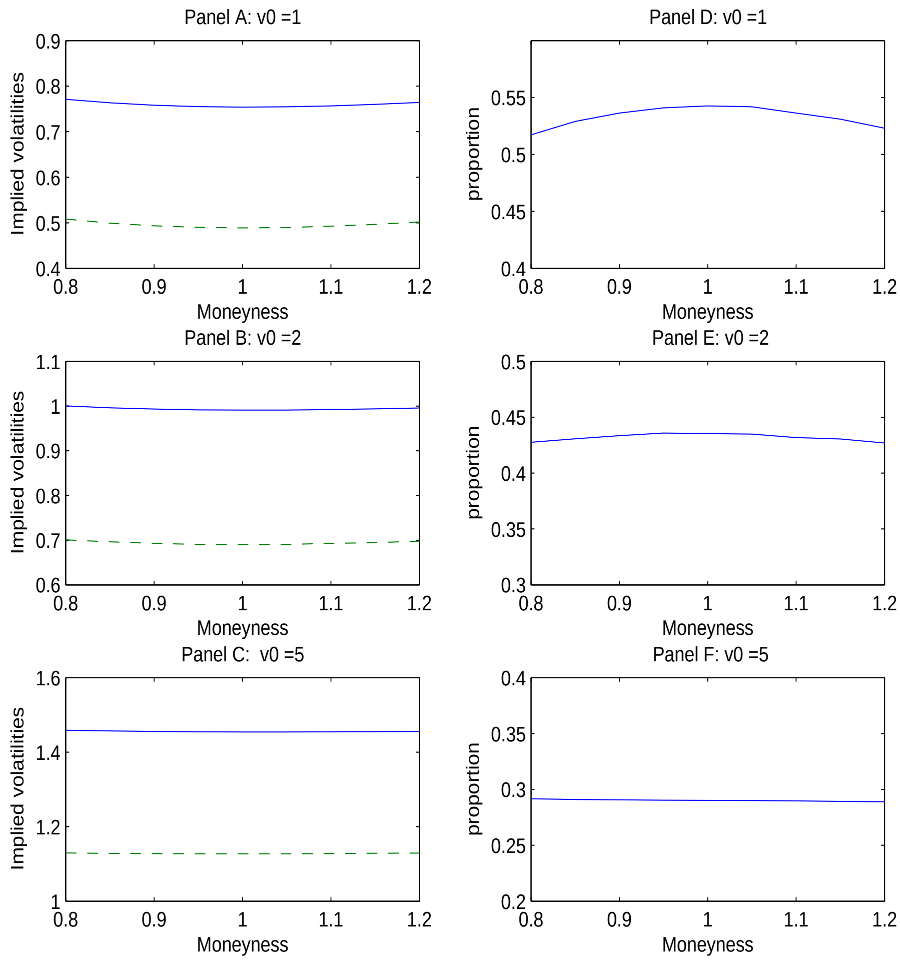

Figure 1.5: Implied volatilities via robust and benchmark approach

The graph displays the implied volatilities from the robust approach and benchmark approach, with different initial volatility v_0. The solid line represents the implied volatilities from the robust approach while the dash line describes the implied volatilities from the benchmark approach. The input parameters are $T = 4$, $T' = 1$ and $F = 50$. The option is computed where the variance is expressed via (1.25). Panel A-B-C display the implied volatilities for $v_0 = 1,2,5$, respectively. Panel D-E-F display the percentage, defined as $(IV^R - IV^B)/IV^B$, of the uncertainty component with respect to the implied volatility from the benchmark approach. IV^R is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. The initial volatility $v_0 = 1,2$ and $v_0 = 5$ in Panel D-E-F, respectively.

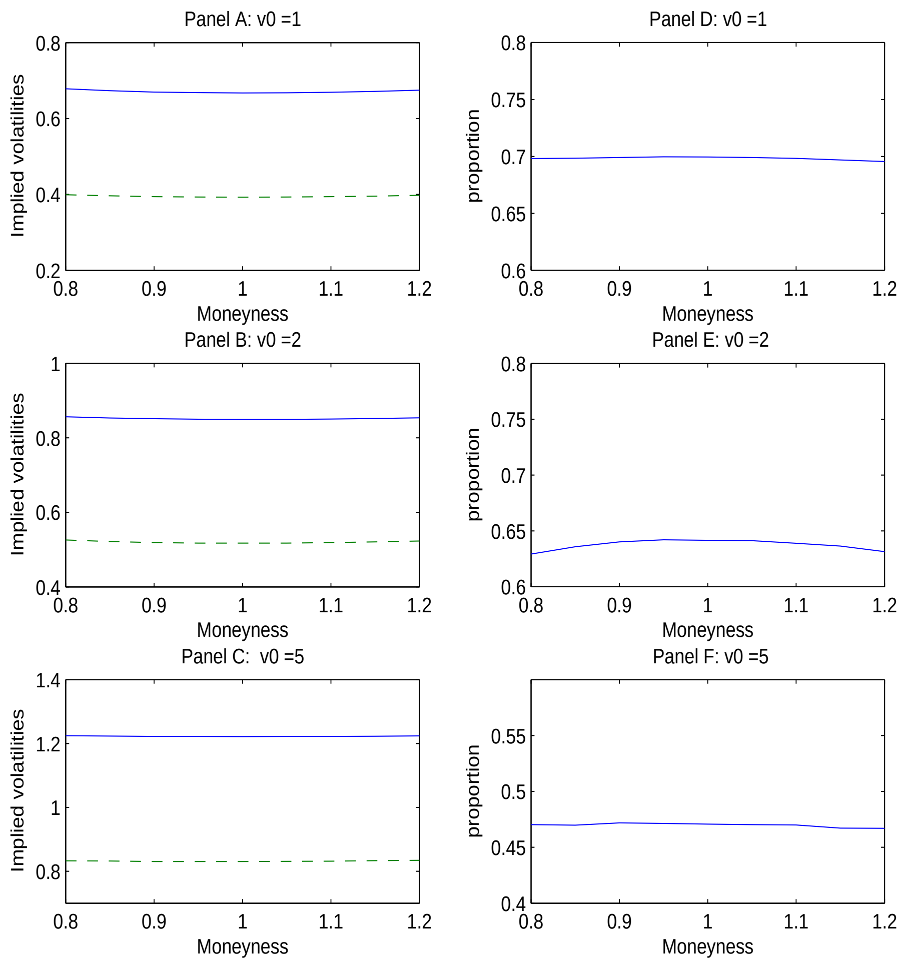

Figure 1.6: Implied volatilities via robust and benchmark approach

The graph displays the implied volatilities from the robust approach and benchmark approach, with different initial volatility v_0. The solid line represents the implied volatilities from the robust approach while the dash line describes the implied volatilities from the benchmark approach. The input parameters are $T = 4$, $T' = 1.5$ and $F = 50$. The option is computed where the variance is expressed via (1.25). Panel A-B-C display the implied volatilities for $v_0 = 1, 2, 5$, respectively. Panel D-E-F display the percentage, defined as $(IV^R - IV^B)/IV^B$, of the uncertainty component with respect to the implied volatility from the benchmark approach. IV^R is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. The initial volatility $v_0 = 1, 2$ and $v_0 = 5$ in Panel D-E-F, respectively.

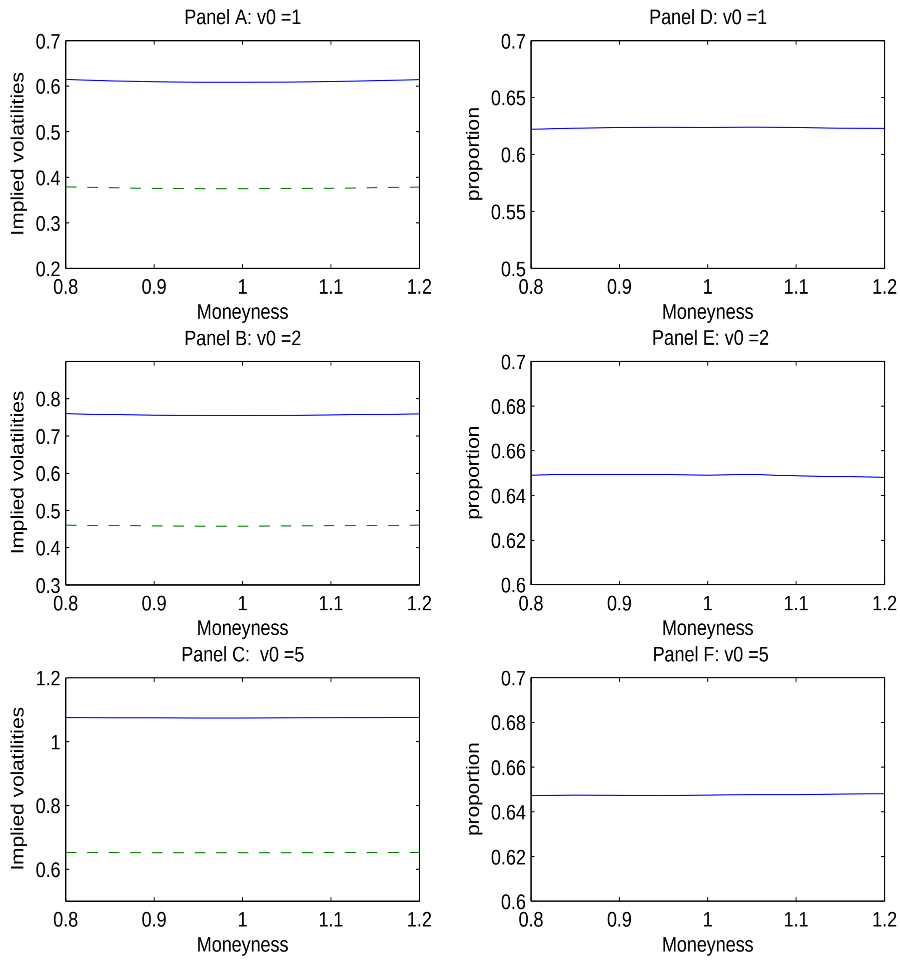

Figure 1.7: Implied volatilities via robust and benchmark approach

The graph displays the implied volatilities from the robust approach and benchmark approach, with different initial volatility v_0. The solid line represents the implied volatilities from the robust approach while the dash line describes the implied volatilities from the benchmark approach. The input parameters are $T = 4$, $T' = 2$ and $F = 50$. The option is computed where the variance is expressed via (1.25). Panel A-B-C display the implied volatilities for $v_0 = 1, 2, 5$, respectively. Panel D-E-F display the percentage, defined as $(IV^R - IV^B)/IV^B$, of the uncertainty component with respect to the implied volatility from the benchmark approach. IV^R is the implied volatility from the robust approach and IV^R is the implied volatility from the benchmark approach. The initial volatility $v_0 = 1, 2$ and $v_0 = 5$ in Panel D-E-F, respectively.

CHAPTER 2: A FACTOR MODEL WITH QUADRATIC PRICING KERNEL

2.1 Introduction

Pricing equity and its financial derivative is a core question in finance. Researches on this topic are broad and extensive. We all know that the price of a derivative depends on its underlying asset, but researches on equity return and its derivative usually follow different model or framework. There are two main frame works in literatures. Firstly, factor models are widely accepted in cross sectional literature, where researches focus on explaining the cross sectional dispersion across individual stocks. By setting a time series process of risk factors, factor models are also showed up in time series literatures, where researches focus on the variation of market return across time. Secondly, stochastic volatility model and GARCH volatility model to some degree dominate literatures in derivative market. In volatility model, the time series variation of return is explained by the time varying volatility. However, factor models are rarely used to price option price because it is hard to specify a reasonable pricing kernel. For volatility model, on the other hand, the one factor (volatility) is not good enough to explain the cross sectional anomalies in equity market.

To price a primitive security, we concentrate our attention on its first order information only and neglect its higher order information in a discrete time model . Accurately estimate the return of a primitive security enables us to focus on its first order information as completely as possible, while higher order information is often neglect by literatures. A usual way in literature to check the viability of a model of estimating the return of a primitive security is by examining the cross section return, $E_t(B_t r_{t+1} - B_t p_t) = 0$, where B_t denotes the weights of different primitive securities in a portfolio and p_t is the price vector of different primitive securities. The higher order information of $E_t(r_{t+1})$ is not examined or required to know in this case. To check the time series return of a primitive security in a specific model, discrete time models dominate this area because of its simplicity to implement and estimate period by period[1].

To price a primitive security's derivative, we need to know a suitable pricing kernel or the specification of a model under risk neutral probability measure and higher order information of its underlying asset's return, while the mean of an underlying asset can not tested by derivative data. By transfer the underlying asset return in a real probability measure to the return under risk neutral probability measure, we do not need to worry about how to find the correct discount rate of its payoff in each time period. Under risk neutral probability measure, the drift of a primitive security's return is always equal to risk free rate. Thus derivative can not tell us the first order information of its underlying asset under real probability measure. Because the payoff of a derivative is a nonlinear function of the return, we need to know detailed higher order information of its primitive underlying asset to better price a derivative. Thus higher order information of underlying asset under risk neutral measure can significantly effect derivative pricing. For instance, volatility model which captures the time varying volatility are better price options than Black Scholes model in equity market.

To sum up, derivative pricing models focus on capturing higher order information of a primitive security return under risk neutral measure, while primitive security pricing mod-

[1]The return of a primitive security depends on macro economic factors, all these information can only be acquired monthly, quarterly or yearly. a discrete time model is much easier to incorporate all these information

els mainly focus on estimating the mean of a primitive security return under real probability measure. To build up a unified model to explain both primitive equity market and its derivative market, a suitable model is required to capture the first order information of equity return under physical measure and higher order information of underlying asset's return under risk neutral measure.

Linear factor model is the most influential model in pricing a primitive security. Literatures on this topic are prolific. A general specification is $Er_{t+1} - r_f = \theta(u_t) f_{t+1}$. f_{t+1} are the risk factors, $\theta(u_t)$ are the factor loadings and u_t are instruments. Factors are generally specified as observed macro factors or mimicking market risk factors. For example, Fama-French three factors (Fama and French (1992)), momentum factor (Jegadeesh and Titman (1993)) and coskewness factor (Harvey and Siddique (2000)) are all important factors in pricing a primitive security. Although these models can explain the price of a primitive security well, there are several limitations restrict its appeal to derivative pricing.

Firstly, there is no consensus on the number of factors or the type of factors in literature. From arbitrage pricing theory (APT) or Merton's ICAPM, factors represent systematic risk or future investment opportunities, all idiosyncratic risk is diversified. Many researches use macro economic factors to directly represent the systematic risk. For example, Petkova and Zhang (2005) use the dividend yield, the default premium, the term yield premium and the one-month Treasury bill rate as the risk factors. We can also follow the principle of ``one price" to construct the risk factors. Systematic risk can be proxied by market mimicking portfolios such as Fama-French 3 factors and Momentum factors. Besides that, several recent papers provide evidence to show that idiosyncratic risk is also priced (Fu (2010)). Thus it is not clear what type of risk factors and how many risk factors should be included.

Correctly defining the factors are crucial in explaining a primitive security for a linear factor model.

Secondly, these factors are only related with a primitive security but they can not predict its price (See Brenna, Wang and Xia (2005)). Without specifying a complete process of each factor, we can only do in sample test. These may arise the possibility of factor dredging and overfitting the data (See Famma (1991) and Lo and Mackinlay (1995)).

Thirdly, such linear factor models fail to fit the volatility of primitive security well. It is commonly know that the volatility of a primitive security is time varying. Based on the linear factor model, the volatility is a constant. This restricts its application to derivative pricing. Even for a linear factor model with time varying beta, without correctly specifying the factors, Ghysels (1998) shows that pricing errors with constant beta models are smaller than with conditional CAPMs.

Fourthly, the specification of a linear factor model is independent of a pricing kernel. Pricing securities in primitive market does not depend on the pricing kernel. However, if we want to price any derivative by a linear factor model, a suitable pricing kernel need to be specified. Many literatures often assume that the pricing kernel is a linear function of risk factors for simplicity. Given other standard assumptions, the linear pricing kernel implies that an asset has a constant volatility, which can not fit the cross sectional option data well. Specifying the pricing kernel is also not voluntary. A correct pricing kernel with the specification of an asset's return should always follow the standard asset pricing theory to prevent any arbitrage opportunity. In other words, the specification of an asset's return in a linear factor model itself restricts the specification of the pricing kernel. Thus searching for a good pricing kernel for a linear factor model is difficult.

In option pricing literatures, there are two important strands, stochastic volatility models and GARCH option pricing models. Both stochastic volatility models and GARCH option pricing models can capture the time varying volatility of an underlying asset's return. Compared with stochastic volatility model, the volatility in GARCH option pricing model is determined and spanned, while in a stochastic volatility model, the volatility includes an unpredictable part and is also unspanned. So GARCH option pricing models is easier to implement and requires no filtration in estimation. In this paper, we only focus on discussing GARCH option pricing models. To price a derivative, the price process of an underlying asset under physical measure is not important as long as the drift term of its underlying asset price process is the risk free rate under the risk neutral probability measure. Most GARCH models do not directly specify a pricing kernel, instead they specify the model under the risk neutral probability measure directly (Glosten, Jagannathan and Runkle (GJR),Duan (1995), Heston and Nandi (2000) and Adesi, Engle and Mancini (2008)). This approach may appear to be a pure fitting exercise, concerns of economic constrains may arise. Mainly focus on explaining cross sectional option data, GARCH option pricing model often gives a much simpler specification of an asset's return under physical measure and contains less economic intuitions than the linear factor model. The drift term in GARCH option pricing moels are often assumed to be a linear function of volatility. This specification makes the GARCH option pricing model less attractive than a linear factor model in cross sectional literatures or time series literatures in primitive market.

In this paper, we build up a multi-factor model with a quadratic pricing kernel. This model links cross sectional literatures, time series literatures, option pricing literatures and term structure literatures. Following standard setting of factor models, the return of underlying asset is a linear function of multi-risk factors. The coefficients of risk factors are time varying and are proxied by past risk factors. This specification makes ICAPM model and conditional CAPM model nested within our framework. This lends our model capability in explaining the cross sectional variation across individual stock.

To describe the time varying dispersion of an underlying asset's return, we assume that risk factors follow OU processes. This leads our model to an intertemperal model setting. The OU process is a standard setting extending factor models to time series literatures. Unlike other linear factor models in time series literature, the volatility in our model is time varying and its dynamic process is similar to that in standard GARCH option pricing models. In this paper, we also show that our model can capture the asymmetric effect of innovation on volatility.

In this paper, we explicitly specify the logarithm of pricing kernel as a quadratic function of risk factors. Given the quadratic pricing kernel, we can easily make a linear transformation of the process of risk factors from physical measure to risk neutral measure. Thus the moment generating function of an underlying asset's return in our model is an exponential function of risk factors. We can easily price an option based on the almost closed form solution of GARCH option pricing models in Heston and Nandi (2000). Because the pricing kernel is new, in this paper we focus on the ability of our model fitting cross sectional option data. Compared with two benchmark models, new Heston and Nandi model and ad-hoc Black Scholes model, we find that our model gets lower RMSE than the other two models both in sample and out of sample.

There are four main contributions in this paper.

1. This is the first one to explicitly present a quadratic multi-factor pricing kernel. This

quadratic multi-factor pricing kernel can make a linear transformation for risk factors from physical measure to risk neutral measure if the risk factors follow OU processes. When the market index return is a standardized linear function of risk factors with time varying coefficients, the moment generating function of market index return in our model can be written as an exponential function of risk factors. Heston and Nandi (2000) gives our an almost closed form solution to GARCH option pricing models when the moment generating function is known. In our model, the volatility at time t+1 is known at time t. Thus in nature, our model is still a GARCH type option pricing model. We can easily find cross sectional option price by Fast Fourier Transform technique or numerical integration approach. This quadratic pricing kernel extends the linear factor model to option pricing literatures.

2. The simple one factor model with quadratic pricing kernel outperforms the new Heston and Nandi model (Christoffersen, Heston and Nandi (2011)) and ad-hoc Black Scholes model both in sample and out of sample in fitting cross sectional option data. In this paper, we estimate three models, the new Heston and Nandi model, the ad-hoc Black Scholes model and one factor model from option data on *SP* 500 index from 2005 to 2007. In Christoffersen, Heston and Jacob (2011), the new Heston and Nandi model is proved to be better fitting than Heston and Nandi (2000) by incorporating a new variance risk premium. This new Heston and Nandi model is considered as a more challenging benchmark model. The RMSE of our one factor model is much lower than the RMSE of new Heston and Nandi model or ad-hoc Black Scholes model overall. Especially in 2007, the RMSE is 50 cents lower than the RMSE of new Heston and Nandi model. We also perform out of sample test to check whether our

model is overfitting the data. Our one factor model still gets lower RMSE than other models in forecasting the option data 52 weeks ahead.

3. The model provides a possible solution to several stylized facts simultaneously. Empirical studies find the following stylized facts, (a) implied volatility is higher than physical volatility, (b) the density function of market index return under risk neutral measure has fatter tails than that under physical measure and (c) there exists a U shape relationship between the pricing kernel and the market index return.

4. This model provides a potential unified frame work, which can connects cross-sectional literatures, time series literatures, option pricing literatures and term-structure literatures together. In our model, individual stock return is a linear function of risk factors while the coefficients of risk factors are time varying and proxied by past risk factors. This specification is consistent with the standard ICAPM model and the conditional CAPM models. It enables our model can explain the cross sectional variation across individual stock. By assuming the risk factors follow OU process, the volatility of individual stock return or market index return is time varying and is similar to GARCH models. This enables our model capture the time varying variation of an asset's return. A quadratic pricing kernel joint with the specification of underlying asset's return and risk factor implies that the moment generating function of underlying asset is an exponential function of risk factors. We can easily price option using the almost closed form solution of GARCH option pricing models in Heston and Nandi (2000). Besides that, our model links to the term-structure literatures. The quadratic term structure model (QTSM) is one of the most important strand in term structure literatures. The risk free rate is assumed as a quadratic function of risk factors while the risk fac-

tors follow OU processes. The specification of risk free rate is consistent with our specification and can be easily incorporated in our model.

The paper is organized as follows. Section 2 presents the multi-factor model with a quadratic pricing kernel and give out a detailed option price formula. In section 3, we compare our model with standard GARCH model from the perspective of volatility process. In section 4, we calibrate our one factor model and other two benchmark models, new Heston and Nandi model and ad-hoc Black Scholes model. We also compare these three models by in-sample and our of sample tests. Section 5 gives our possible solution to several stylized facts by our model. Section 6 discusses the link of our model to cross sectional literatures, time-series literatures, and term structure literatures. Section 7 provides the conclusions and all proofs are presented in Appendix.

2.2 Model

In this section, we firstly present a general linear multi-factor model with a quadratic pricing kernel. The logarithm of conditional market return is a general linear function of risk factors, while the coefficients and the intercept are a linear function and a quadratic function of past risk factors respectively. This specification make many return specifications in cross sectional literatures are nested within our model. We assume that the logarithm pricing kernel is a quadratic function of risk factors. This specification automatically leads the pricing kernel to be time-varying and has non-monotonic relationship with return. Given the assumption that the risk factors follow OU processes, European option price can be easily derived and estimated. Secondly, we compare our model with other GARCH option pricing models. Unlike standard GARCH option pricing models, we separate the effect of past innovation on volatility from the residual. We present the conditional variance in our model can capture more information than that in classical GARCH models or Heston and Nandi (2000). Besides that, our model can also capture the asymmetric effect of innovations on volatility.

2.2.1 Model Specifications

Researches in equity market mainly try to answer the following three questions, the dispersion across stocks (Cross sectional analysis, referred as CS), variation across time (Time series analysis, we refer it as TS) and derivative pricing.

However, researches in these areas follow different strands, there is no unified model can be applied in all of these three areas or link literatures in these areas together. In this paper, we propose a multi-factor model with a quadratic pricing kernel. Many models, such as CAPM, conditional CAPM and conditional ICAPM are nested within the specification of the in cross sectional analysis. More importantly, given that the risk factors following OU processes and a quadratic pricing kernel, this model can explain the time variation of index return and options on index cross sectionally. We are the first one to directly extend the linear factor model to option pricing in derivative market. In this section, we will first discuss the specification of return, and then the quadratic pricing kernel. Given model specifications, we present an explicit transformation of an innovation from physical measure to risk neutral measure. Moment generating function of the underlying asset's return is also presented.

Return Across Stocks

Explaining the dispersions of cross sectional returns in equity market is one of the most important topics. Unconditional CAPM may be the simplest and most influential model in cross sectional literatures. CAPM implies that cross sectional stock returns can be explained by one factor, the return of market portfolio. However, CAPM can not explain several anomalies, such as small firms earn more than big firms on average, value firms perform better than growth firm etc (See Fama and French 1989, 1992 and 1993). To explain these anomalies, more systematic factors (states) are added to explain the cross sectional anomalies. There are two theories issue the fishing license to researchers who dedicate to find new risk factors. One of the two is the Arbitrage pricing theory (APT) in Ross (1976). Based on the principle of ``one price", APT provides a theoretical background for a group of linear factor models for one time period. risk factors in these models are often represented as the excess returns of market mimicking portfolios, such as Fama-French factors, momentum factor, etc. However, the APT does not preclude arbitrage over dynamic portfolios. The intertemporal CAPM (ICAPM) model (Merton, 1973) issues another ``fishing license" to authors who are in search of risk factors for a dynamic linear factor model, factors[2] in which are not necessary to be excess returns of market mimicking portfolios. In general, these risk factors can be divided into three groups, excess return of mimicking market portfolios, macroeconomic factors and firm-specific factors. Popular risk factors formed by mimicking market portfolios are size factor, book to market ratio factor, momentum factor, liquidity factor, etc. Many macroeconomic factors can also explain the cross sectional return in equity market, such as one month Treasury bill rate (Fama and Schwert, 1977; Campbell, 1991; Hodrick, 1992), Yield spread, default spread, dividend yield (see, e.g., Hahn and Lee 2006, Campbell and Cochrane, 1999, Constantinides and Duffie, 1996, Petkova and Zhang, 2005) etc. Many firm specific risk factors such as firm size, book to market ratio (Fama and French 1992, 1993), liquidity factor (Pastor and Stam-

[2]In ICAPM model, it defines those risk factors which can predict future return of individual stock as states, to conform with the terminology with APT, in this paper, we do not distinguish them.

baugh, 2003), aggregate pricing-earnings ratio (Campbell and Shiller, 1988b; Campbell and Vuolteenaho, 2004), etc, also determine individual stock return. Literatures in cross sectional analysis do not clearly distinguish APT from ICAPM (See Cochrane, 2005). Because ICAPM provides much more straight interpretation of risk factors than APT, we will interpret risk factors from ICAPM.

Merton (1973) ICAPM is to solve an optimal portfolio choice problem of a representative investor in continuous time. In equilibrium, the relationship between expected return and risk is given in a simplified case:

$$\mu_i - r = \gamma\sigma_{im} + \gamma_z\sigma_{iz} \tag{2.1}$$

where γ denotes the parameter of relative risk aversion; σ_{im} and σ_{iz} denote the covariances between the return on asset i and the market return and state variable, respectively. γ_z denotes the risk priced associate with state variable. The approximate discrete time version of equation (1) is in Cochrane (2005, Chapter 9).

$$E_t R_{i,t+1} - R_{f,t+1} = \gamma Cov_t(R_{i,t+1}, R_{m,t+1}) + \gamma_z Cov_t(R_{i,t+1}, \Delta Z_{t+1}) \tag{2.2}$$

where $R_{i,t+1}$ is the return on asset i between t and $t+1$; $R_{f,t+1}$ denotes the risk-free rate at time t; $R_{m,t+1}$ is the market return; and ΔZ_{t+1} denotes the innovation or change in the state variable. Unfortunately, the innovation or change in the state variable are latent and unobservable. Literatures often project $Cov_t(R_{i,t+1}, \Delta Z_{t+1})$ linearly to risk factors. Any risk factors based on the ICAPM model are proxies of innovations in state variable or the change of state variable. For example, Campbell (1996) implies that factors in ICAPM should be related to innovations in the state variables that forecast future investment opportunities; Liew and Vassalou (2000) relates FF factors to future rates of economic growth; both Lettau and Ludvigson (2001) and Bassalou (2003) show that FF factors include the information of GDP growth. However, it is not free to add risk factors based on the theory of ICAPM model. Risk factors in ICAPM model should satisfy other constraints[3]. Paulo Maio and Pedro Santa-Clara (2012) find that only Fama-French three factors and Carhart momentum factor meet ICAPM restrictions when checked with 25 portfolios sorted by size and book to the market ratio. Even though researches in these risk factors are extensive, there has not been any consensus on the number of risk factors or what risk factors should be included in ICAPM[4].

Because the state variables are not clear and we also do not know what risk factors are best proxy for those state variables. In this paper, we do not identify risk factors and only assume that the return of individual stock is driven by N latent risk factors X_t which are proxy of innovations or change in the state variable. Identifying risk factors are left for future works.

Classic ICAPM assumes that when $Cov_t(R_{i,t+1}, R_{m,t+1})$ and $Cov_t(R_{i,t+1}, \Delta Z_{t+1})$ are linearly projected on risk factors, all the betas are constant. In ICAPM, the return of individual stock can be written as a linear function of multi-risk factors[5].

$$log(R_{i,t+1}) = \beta_{0,i} + \beta_{1,i} X_{t+1} + \varepsilon_{i,t+1} \tag{2.3}$$

[3] First of all, if a state variable forecasts positive (negative) changes in investment opportunities in time series regressions, its innovation should earn a positive (negative) risk price in the cross-sectional test of the respective multi-factor model. Second, the market price of risk must be economically plausible as an estimate of the coefficient of relative risk aversion (RRA).

[4] Merton (1973) does not identify any state variables.

[5] From here, we refer it as multi-factor model

However, It is important to specify a time varying coefficient in a linear factor model. Bollerslev, Engle and Wooldridge (1988), Lettau and Ludvigson (2001) both find that time-varying drift and factor loadings in a linear factor model can significantly improve its power in explaining both the cross-sectional dispersion and the time series dispersion in primitive equity market.

Individual stock return in a general setting for multi-factor model can be expressed as:

$$ln(R_{i,t+1}) = \beta_{0,i}(u_t,t) + \beta_{1,i}(u_t,t)X_{t+1} + \varepsilon_{i,t+1} \tag{2.4}$$

$R_{i,t+1}$ is the return of individual stock i, $\beta_{0,i}(u_t,t)$ is the intercept, and $\beta_{1,i}(u_t,t)$ is the factor loadings for individual stock i. X_{t+1} is a vector of j risk factors $X_1, X_2, ..., X_j$. $\varepsilon_{i,t+1}$ is the idiosyncratic risk. u_t is the instruments which help to identify the beta of risk factors. The drift $\beta_{0,i}(u_t,t)$ and the coefficient of risk factors,$\beta_{1,i}(u_t,t)$ can both be time varying. For a well diversified portfolio, idiosyncratic risk can be diversified away. The return of a well diversified portfolio I (e.g., size portfolios, book to the market ratio portfolios and SP500 index) can be written as:

$$ln(R_{I,t+1}) = \beta_{0,I}(u_t,t) + \beta_{1,I}(u_t,t)X_{t+1} \tag{2.5}$$

Following the standard linear multi-factor model, we assume that the return of individual stock in our model follows:

$$ln(R_{i,t+1}) = X_t' b_i X_t + c_i' X_{t+1} + d_i' X_t + X_t' g_i X_{t+1} + r_{t,t+1} + f_i \tag{2.6}$$

Many model specifications of return in literature are nested within equation (6), such as conditional CAPM model, ICAPM model and conditional ICAPM model.

The aggregate return in the market follows:

$$ln(R_{t,t+1}) = X_t' b X_t + c' X_{t+1} + d' X_t + X_t' g X_{t+1} + r_{t,t+1} + f \tag{2.7}$$

The dimension of c, d and g is N by 1, N by 1, N by N respectively. $r_{t,t+1}$ denotes the risk free rate from time t to $t+1$. f is a constant. It is straight forward to rewrite the return of a well diversified portfolio into a linear function of multi risk factors X_{t+1} at time t.

$$ln(R_{t,t+1}) = X_t' b X_t + +d' X_t + r_{t,t+1} + f + (X_t' g + c') X_{t+1} \tag{2.8}$$

We can directly match the return function in equation (7) with the general setting of linear factor model in equation (5),

$$\beta_{0,I}(u_t, t) = \beta_{0,I}(X_t, t) = X_t' b X_t + +d' X_t + r_{t,t+1} + f$$

$$\beta_{1,I}(u_t, t) = \beta_{1,I}(X_t, t) = X_t' g + c'$$

in which past risk factors are identified as instruments u_t. We use a quadratic function of X_t to approximate $\beta_{0,I}(u_t, t)$ and also use a linear function of X_t to approximate $\beta_{1,I}(u_t, t)$. Using linear function of past risk factors as instruments is widely accepted in literatures of conditional CAPM model (See Petkova and Zhang (2005)). Because $\beta_{0,I}(u_t, t)$ and $\beta_{1,I}(u_t, t)$ depend on past risk factor X_t, they are time varying and path dependent.

Quadratic Pricing Kernel

To capture the time series variation of return, we need to specify the process of risk factors across time In the multi-factor model.

In continuous time version of ICAPM model, many papers assume that state variables follow Ornstein-Uhlenbeck processes (OU). For example, Kim and Omeberg (1996) assume an OU process for Sharpe ratio. Brennan, Wang and Xia (2004) also assume OU processes the maximum Sharpe ratio and the real interest rate.

$$dX_i(t) = k(\mu_{X_i} - X_i(t))dt + \sigma_{X_i} dZ_{X_i(t)} \tag{2.9}$$

For state variable X_i, μ_{X_i} is a parameter controlling for the mean, k is a parameter controlling for the reversion speed. σ_X is a parameter controlling for volatility.

The discrete time version of OU process is AR(1) process. In order to construct a tractable valuation model, we assume that a well diversified portfolio return and its derivatives are both effected by N common risk factors $X_{i,t}, i = 1, 2, ...N$ at time t in an equilibrium economy. All these common risk factors follow AR(1) processes.

$$X_{i,t+1} = \omega_i + \phi_i X_{i,t} + \Sigma \varepsilon_{i,t+1}, \tag{2.10}$$

where $\varepsilon_{i,t+1}$ $N(0,1)$ and ε_{t_1} is independent of ε_{t_2} if $t_1 \neq t_2$.

This setting can be easily extended if proper state variables are identified. However, in literature, studies on linear factor models are often constrained in the primitive security market. It is difficult to extend the linear factor model to derivative markets without a proper pricing kernel (Stochastic discount factor).

In literature, there are two approaches to capture the pricing kernel. Firstly, we can directly specific the pricing kernel as a function of risk factors (See Harvey and Siddigue (2000) and Dittmar (2000)). Also we can use non-parametric approach or semi-parametric approach to estimate it (See Jackwerth (2000) and Ronsenberg and Engle (2002)). Secondly, we can indirectly specify the pricing kernel by specify the form of models in risk neutral probability measure (See Heston (1993), Heston and Nandi (2000)). In this paper, we follow the first way to examine the relationship between the return and pricing kernel.

Researches on the form of pricing kernel in literatures are extensive. Many classical linear factor models propose that the pricing kernel is a linear function of risk factors. In the unconditional CAPM, the representative agent's derived utility function may be restricted to forms such as quadratic or logarithmic which guarantee that the pricing kernel is linear in the value weighted portfolio of wealth. Independent of any representative agent's utility function, APT (Ross, 1976) indirectly assumes that the pricing kernel is a linear function of risk factors, proxied by the excess return of market mimicking portfolios. However, recently there are two main important findings about the pricing kernel in equity market. Firstly, the pricing kernel is time varying. Ronsenberg and Engle (2002) find that a time varying pricing kernel significantly improved the hedging performance compared with that by a time-invariant pricing kernel. Jackwerth (2000) also finds that the risk aversion function is positive and decreasing precrash, while it is partially negative and increasing postcrash. In option market, Bliss and Panigirtzoglou (2004) finds that the estimated coefficient of risk aversion decline with the forecast horizon and is higher during periods of low volatility. Secondly, the pricing kernel is an non-monotonic function of market return. Harvey and

Siddigue (2000) and Dittman (2000) proxy the pricing kernel as a quadratic function and a cubic function of market return respectively. They both find that models based on nonlinear pricing kernel can significantly improve the linear pricing kernel in explaining cross sectional stock returns. In option market, Bakshin, Madan and Panayotov (2010) use a model free approach to detect the shape of pricing kernel. They find that the average returns of claims with payout on the upside contradict the implications of downward-sloping pricing kernel, but can be consistent with a U-shaped pricing kernels.

For the inter-temporal consumption based model, Harrison and Kreps (1979) and Hansen and Jagannathan (1991) have showed that the pricing kernel m_{t+1} is the investor's intertemporal marginal rate of substitution of consumption, $\delta \frac{U'(C_{t+1})}{U'(C_t)}$, where δ is the subjective rate of time preference. More importantly, m_{t+1} is nonnegative under the condition of non-arbitrage. Without specifying the utility function, many literatures assume that the pricing kernel can be approximated as a linear function of consumption growth by Taylor expansion, $m_{t+1} = a_t + b_t \Delta c_{t+1}$ (See Lettau and Ludvigson (2001)). a_t and b_t are parameters (potentially time-varying) and Δc_{t+1} is consumption growth, which is equal to $ln(C_{t+1}/C_t)$. Literatures often assume that the growth rate of consumption is proportional to the growth rate of market return (See e.g. Lettau and Ludvigson (2001), Harvey and Siddigue (2000) and Dittman (2000)). The pricing kernel can be written as $m_{t+1} = a_t^* + b_t^* lnR_{t,t+1}$, where $lnR_{t,t+1}$ is the natural logarithm of market return from time t to time $t+1$. The form of pricing kernel directly relates the representative's utility function. For example, a logarithmic or quadratic utility function can lead the pricing kernel to a linear function of value weighted portfolio of wealth. This specification can not guarantee that the pricing kernel follows the non-arbitrage constraint (nonnegative). Besides that, as discussed above, recent evidence shows that the pricing kernel is a non-monotone function of market return. Thus a logarithmic or quadratic utility function may not be suitable. In our paper, we also do not directly specify any utility function. To capture the two facts, pricing kernel is nonnegative and the monotone relationship between pricing kernel and return, we assume that the logarithm of pricing kernel is a quadratic function of log return of market index. This specification not only guarantees that the pricing kernel is nonnegative, but also can capture the monotone relationship between the pricing kernel and the market return.

$$ln(m(t,t+1)) = A_t^* + B_t^* lnR_{t,t+1} + C_t^*(lnR_{t,t+1})^2 \tag{2.11}$$

where A_t^*, B_t^* and C_t^* can be time varying. This pricing kernel is similar to the one discussed in Harvey and Siddique (2000) and Dittmar (2002). Harvey and Siddique (2000) propose that the pricing kernel is a quadratic function of market return.

$$m(t,t+1) = a_t + b_t R_{t,t+1} + c_t R_{t,t+1}^2 \tag{2.12}$$

while in Dittmar (2002), the pricing kernel is a cubic function of market return.

$$m(t,t+1) = a_t + b_t R_{t,t+1} + c_t R_{t,t+1}^2 + d_t R_{t,t+1}^3 \tag{2.13}$$

Our specification is just transform theirs from discrete interest version into a compounding interest version.

Put equation (7) back to equation (11), after simplification, the logarithm of pricing kernel can be expressed as[6]:

$$ln(m(t,t+1)) = X'_{t+1}A_tX_{t+1} + X'_tB_tX_t + C'_tX_{t+1} + D'_tX_t + X'_tG_tX_{t+1} - r_{t,t+1} + F_t \quad (2.14)$$

For simplicity, in this paper, we only focus on a special case of the pricing kernel, in which A_t, B_t, C_t, G_t and F_t are all constant. The logarithm of the pricing kernel $log\frac{M(t+1)}{M(t)}$[7] is a time varying quadratic function of risk factors X_{t+1}.

$$ln(m(t,t+1)) = X'_{t+1}AX_{t+1} + X'_tBX_t + C'X_{t+1} + D'X_t + X'_tGX_{t+1} - r_{t,t+1} + F \quad (2.15)$$

where the dimensions of A, B, C, G and F are N by N, N by N, N by 1, N by 1, N by N and 1 respectively. Without loss any generality, we assume that G is a symmetric matrix. This specification links literatures in term structure. the quadratic term structure model (QTSM) (See Ahn, Dittmar and Gallant (2002)) is one of the most important models in term structure literatures. In QTSM, the interest is a quadratic function of risk factors, $r_{t,t+1} = \alpha + \beta' X(t) + X(t)\phi X(t)$. The quadratic specification not only makes the risk free rate be positive all the time, but also captures several puzzles not explained by the affine term structure models (ATSM) (See Ahn, Dittmar and Gallant (2002), Leippold and Wu (2002)). A discrete time version of the pricing kernel in this group is just a special case of our pricing kernel, in which B and G are both zero matrix. Leippold and Wu (2002) also proves that, in QTSMs, specifying a OU process for states is a necessary condition, which conforms with our setting for the risk factors. We can easily justificator our model by incorporating the risk free rate as a quadratic function of risk factors. If α, β, ϕ is identified in term structure

[6] higher order term of $X_t^2X_{t+1}^2$ and X_t^4 is neglect for simplicity

[7] From now on, we use $m(t,t+1)$ as $\frac{M(t+1)}{M(t)}$

model, we can simply insert $r_{t,t+1} = \alpha + \beta' X(t) + X(t)\phi X(t)$ back to equation (16).

$$\begin{aligned} ln(m(t,t+1)) &= X'_{t+1}AX_{t+1} + X'_t(B+\phi)X_t + C'X_{t+1} + (D'+\beta')X_t + X'_tGX_{t+1} + F \\ &\quad +\alpha \end{aligned} \tag{2.16}$$

This new pricing kernel can be used in equity market and bond market simultaneously. In this paper, we focus on equity market, for simplicity, interest rate $r_{t,t+1}$ is given at time t.

Model Identification

Given the pricing kernel in equation (14), we can easily change the real probability measure into risk neutral probability measure with an explicit form.

Given that

$$ln(m(t,t+1)) = X'_{t+1}AX_{t+1} + X'_tBX_t + C'X_{t+1} + D'X_t + X'_tGX_{t+1} - r + F$$

There is a linear transformation between the disturbance ε^P_{t+1} under real probability measure P and the disturbance ε^Q_{t+1} under real probability measure Q

$$\varepsilon^P_{t+1} = \mu_{t+1} + W^{\frac{1}{2}}\varepsilon^Q_{t+1} \tag{2.17}$$

where

$\mu_{t+1} = W'\Sigma[2A\omega + C + (2A\phi + G)X_t]$ and $W = (I - 2\Sigma' A\Sigma)^{-1}$

Under real probability, the disturbance ε^P_{t+1} from risk factors X_{t+1} belongs to standard multi-normal distribution. It is equivalent to a linear function of disturbance ε^Q_{t+1} under risk neutral probability. Even though both ε^P_{t+1} and ε^Q_{t+1} belongs to multi-normal distribution, when we convert ε^P_{t+1} to risk neutral probability measure, not only its drift changes, but also its volatility will change. However, when A is a zero matrix, the logarithm of pricing kernel becomes a linear function of risk factors. The volatility of ε^P_{t+1} under risk neutral probability measure, $W^{\frac{1}{2}}$ becomes an identical matrix and is the same as its volatility under real probability measure.

Accordingly, under risk neutral probability measure, risk factors become

$$\begin{aligned} X_{t+1} &= \omega + \Sigma W \Sigma'(2A\omega + C) + (\phi + \Sigma W \Sigma'(2A\phi + G))X_t + \Sigma W^{\frac{1}{2}} \varepsilon^Q_{t+1} \\ &= \omega^* + \phi^* X_t + \Delta^* \varepsilon^Q_{t+1} \end{aligned}$$

Without loss any generality, we can always make an affine change such that $\Sigma = I, \omega = 0$ and ϕ is a symmetric matrix. This transformation implies that $A = 0.5(I - \Omega^{*-1})$, $C = \Omega^{*-1}\omega^*$, and $G = \Omega^{*-1}\phi^* - \phi$.

To ensure no arbitrage in this setting, $m(t,t+1)$ and $R_{t,t+1}$ is required to satisfy two Euler equations.

$$E_t[m(t,t+1)] = e^{-r_{t,t+1}} \tag{2.18}$$

$$E_t[e^{R_{t,t+1}} m(t,t+1)] = 1 \tag{2.19}$$

Equation (15) meet the requirement that the return a risk free asset is 1 under risk neutral probability measure. Equation (16) ensures that the return of any risk asset should be 1 under risk neutral probability measure. the logarithm of $E_t[e^{R_{t,t+1}}m(t,t+1)]$ and $E_t[e^{R_{t,t+1}}m(t,t+1)]$ are all quadratic functions of $X(t)$. To validate the two Euler equations all the time, the coefficients of X_t, $X_t'X_t$ and the constant term should all be zero matrix or zero. Thus we will have three constraints for each Euler equation. $E_t[m(t,t+1)] = e^{-r_{t,t+1}}$ implies that

$$B + \phi' A\phi + G\phi + \frac{1}{2}(2\phi' A + G)\Omega^*(2A\phi + G) = 0 \tag{2.20}$$

$$\phi' C + D + (2\phi' A + G)\Omega^* C = 0 \tag{2.21}$$

$$F - \frac{1}{2}ln|\Omega^{*-1}| + \frac{1}{2}C'\Omega^* C = 0 \tag{2.22}$$

$E_t[R_{t,t+1}m(t,t+1)] = 1$ implies that

$$b + g\phi^* + \frac{1}{2}g\Omega^* g = 0 \tag{2.23}$$

$$d + \phi^{*'} c + g\omega + g\Omega^* c = 0 \tag{2.24}$$

$$f + c'\omega^* + \frac{1}{2}c'\Omega^* c = 0 \tag{2.25}$$

From these six equations in both propositions, all the parameters in our model can be expressed in terms of 6 parameters, $\theta = \{\omega^*, \phi^*, \Omega^*, c, g, \phi\}$. The expression of all the parameters are in appendix.

By putting all the parameters in term of θ back to equation (6) and (16), we get

$$
\begin{aligned}
ln(R_{t,t+1}) &= r_{t,t+1} - (c'\omega^* + 0.5c'\Omega^* c) - [c'(\phi^* - \phi) + \omega^* g + c'\Omega^* g]X_t \\
&\quad + X_t'[g(\phi - \phi^* - 0.5\Omega^* g)]X_t + (c' + X_t' g)\varepsilon_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
ln(m(t,t+1)) &= -0.5\omega^{*'}\Omega^{*-1}\omega^* + 0.5log|\Omega^{*-1}| - r_{t,t+1} + (\phi^{*'} - \phi')\Omega^{*-1}\omega^* X_t \\
&\quad + X_t'(\phi'\Omega^{*-1}\phi^* - 0.5\phi'\Omega^{*-1}\phi - 0.5\phi^{*'}\Omega^{*-1}\phi^*)X_t \\
&\quad + [X_t'(\phi^* - \phi)\Omega^{*'} - \omega^{*'}\Omega^*]\varepsilon_{t+1} + 0.5\varepsilon_{t+1}'(I - \Omega^{*-1})\varepsilon_{t+1}
\end{aligned}
$$

To get some insights of pricing kernel and market return under risk neutral probability measure, we focus on four parameters $\phi, \omega^*, \phi^*, \Delta^*$, which determine the process of risk factors under physical measure and risk neutral measure.

$$X(t+1) = \phi X(t) + \varepsilon_{t+1}^P$$

$$X(t+1) = \omega^* + \phi^* X(t) + \Delta^* \varepsilon_{t+1}^Q$$

These two equations describe the processes of risk factors under real probability measure P and risk neutral probability measure Q respectively. when A is a zero matrix, Ω^*, the variance of disturbance under risk neutral probability measure P, will become an identical matrix ($\Omega^* = I$). $\varepsilon'_{t+1}(I - \Omega^{*-1})\varepsilon_{t+1}$ will disappear. This means higher order of disturbances will not be priced when the logarithm of pricing kernel is a linear function of risk factors. This is equivalent to say that, when the logarithm of pricing kernel is a linear function of risk factors, investors will only care about the uncertainty of risk factors (the disturbance), however they do not care about how uncertainty the uncertainty of risk factors (higher order information of disturbances) is.

However, If A is not a zero matrix, the logarithm of pricing kernel is a quadratic function of risk factors, investors will not only care about the uncertainty (disturbance), but also how uncertainty the uncertainty of risk factors is, higher order information of disturbance is priced by investors.

If ϕ and ϕ^*, auto correlations under P and Q, are the same, all the terms having X_t will disappear, then the pricing kernel does not depend on past state X_t. This is equivalent to say that the pricing kernel at time t is independent of past information when auto correlations of risk factors are the same under P and Q.

g is critical in determining the conditional volatility of return. When g is a zero matrix, the conditional variance of return is equivalent to $c'c$, which is a constant. However, in reality, the conditional volatility of market return should be time varying. Thus it is important to specify g as a non-zero matrix to account for time varying conditional volatility of market return.

Conditional on time t, the logarithm return $lnR_{t,t+1}$ is a linear function of disturbance ε_{t+1}. By the help of $\varepsilon'_{t+1}(I - \Omega^{*-1}\varepsilon_{t+1})$, the logarithm of pricing kernel becomes a quadratic function of $lnR_{t,t+1}$. In a simplified one factor model, in which $dim(X_t) = 1$, $g = 0$ and

$\phi = \phi^*$, we can easily express the logarithm of pricing kernel as a quadratic function of return.

$$ln(R_{t,t+1}) = -0.5c^2 + r_{t,t+1} + c\varepsilon_{t+1}$$

$$ln(m(t,t+1)) = -r_{t,t+1} - \omega\varepsilon_{t+1} + (1-\Omega)\varepsilon_{t+1}^2$$

It is straight forward to express the logarithm of pricing kernel as a quadratic function of return,

$$\begin{aligned} ln(m(t,t+1)) &= -0.5\omega^2 - r_{t,t+1} - \omega(\frac{lnR_{t,t+1} - r_{t,t+1}}{c} - \omega + 0.5c) \\ &\quad +(1-\Omega^*)(\frac{lnR_{t,t+1} - r_{t,t+1}}{c} - \omega + 0.5c)^2 \end{aligned}$$

Characteristic Function and Derivative Price

Because the payoff of derivatives are usually nonlinear function of underlying asset's return, we need to know the moment generating function of the underlying asset's return to exactly price its derivatives. The moment generating function (in the risk neutral measure) is exponential quadratic in the factors:

$$E_t^Q[(\frac{S_T}{S_t})^{\varphi}] = exp(\alpha_t + \beta_t' X_t + X_t' \gamma_t X_t),$$

where

$$\begin{aligned}\alpha_t &= r_{t,t+1}+\varphi f+\alpha_{t+1}-\frac{1}{2}log|I-2\Delta^{*'}\gamma_{t+1}\Delta^*|+\omega^{*'}\gamma_{t+1}\omega^*+\omega^*(\varphi c+\beta_{t+1})\\ &\quad +\frac{1}{2}(\varphi c'+\beta'_{t+1}+2\omega^{*'}\gamma_{t+1})(I-2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi c+\beta_{t+1}+2\gamma_{t+1}\omega^*)\end{aligned}$$

$$\begin{aligned}\beta_t &= \varphi(\phi^{*'}c+d+g\omega^*)+\phi^{*'}\beta_{t+1}+2\phi^{*'}\gamma_{t+1}\omega^*\\ &\quad +(2\phi^{*'}\gamma_{t+1}+\varphi g)(I-2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi c+\beta_{t+1}+2\gamma_{t+1}\omega^*)\end{aligned}$$

$$\gamma_t=\varphi b+\varphi g\phi^*+\phi^{*'}\gamma_{t+1}\phi^*+\frac{1}{2}(2\phi^*\gamma_{t+1}+\varphi g)(I-2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(2\phi^*\gamma_{t+1}+\varphi g)'$$

α_t, β_t and γ_t can be computed iteratively from the initial conditions that $\alpha_T=0, \beta_T=0$ and $\gamma_T=0$.

The European put option can be calculated from call-put parity. Even though we have 6 free parameters, $\{\phi^*, c, g, \phi, \omega^*, \Omega^*\}$, European style options only depend on five parameters, $\{\phi^*, c, g, \omega^*, \Omega^*\}$. Since now, we call $\{\phi^*, c, g, \omega^*, \Omega^*\}$ are ``risk neutral parameters'' and $\phi$ is ``real parameter''.

In literatures,there are many different GARCH models to price options on market index, such as GJR model (See Glosten, Jagannathan, and Runkle (1993)), HN model (See Heston and Nandi (2000)), Inverse GARCH model (See Christoffersen, Heston, and Jacobs (2006)), etc. In Heston and Nandi (2000), they provide a general closed form solution of GARCH models given moment generating function of underlying assets return. The moment generating function in this group[8] can be written as $\Psi = S_t^{\phi} exp(A_t + B_t \sigma_{t+1}^2)$, where σ_{t+1}^2, known at time t, is the variance of market index return at time t+1. Our model also belongs to GARCH type model, in which the variance of market index return at time t+1 is $(c + gX(t)'(c + gX(t))$. Thus the moment generating function in our model can always be written as $\Psi = S_t^{\phi} exp(A_t + B_t \sigma_{t+1}^2) + C_t \sigma_{t+1}$. Option price in our model not only depends on the variance but also the volatility, which allows our model providing more flexibility in option pricing than the other GARCH models.

Besides that our model can capture more information in the market.

For those GARCH type models, option price at time t+1 depends only on one variable σ_{t+1}^2 known at time t. Option price at time t+1 does not depends on the path of market index from time t to t+1. This seems counter intuition. For example, there are two cases from time t to t+1. In the first case, the market index does not change at all. In the second case, the market index will increase at first and then drop back to the original level. These popular GARCH models say that option price is the same in these two cases at the end of time t+1, because the one to one relationship between the return and the variance. However even though our model is a GARCH type model, the variance of market index return is a function of multi-factors. One market return can not determine a unique set of risk factor value. This means that our model can use different set of risk factor value to represent the two cases even though the market return is the same. It is equivalent to say that our model can ``price'' more information at time t+1 in option price than those popular GARCH models.

Because our model prices ``more information'' and are more flexible than those popular GARCH models, we expect that our model can calibrate option price in equity market more accurately and more stable.

[8] Here we only consider GARCH(1,1) type forms of these models.

2.2.2 Comparison with GARCH Option Pricing Models

There is overwhelming empirical evidence supporting that modeling time varying volatility is critical in modeling the market index return and pricing market index option. The success of stochastic volatility model and GARCH option pricing model in pricing option in equity attribute their abilities to capture the time varying volatility.

In stochastic volatility model, the time varying volatility needs to be filtered by some econometric approach. However, the time varying volatility in GARCH option pricing model can be clearly and directly filtered from physical return.

In general, return process and volatility process under physical measure for a standard GARCH option pricing model can be written as:

$$lnR_{t,t+1} = \mu - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\varepsilon_{t+1} \tag{2.26}$$

$$\sigma_{t+1}^2 = f(\sigma_t, Z_t, \sigma_t\varepsilon_t) \tag{2.27}$$

Where $logR_{t,t+1}$ is the natural logarithm of the underlying asset's return, σ_{t+1}^2 is the variance of return conditional on time t, Z_{t+1} is a random noise at time t+1, the process of conditional return variance, $f(\sigma_t, \varepsilon_t, \sigma_t\varepsilon_t)$ is a function of conditional variance at time t, σ_{t-1}, random noise at time t, ε_t, and the return residual, $\sigma_t\varepsilon_t$. For example, in the GARCH(1,1) model,

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\sigma_t^2 + \alpha_2(\sigma_t\varepsilon_t)^2$$

To accommodate the asymmetrical effect of shocks on volatility, Engle and Ng(1993) starts a group of nonlinear asymmetric GARCH models ((See Heston and Nandi(2000), Christoffersen, Heston, and Jacobs (2006), and Glosten, Jagannathan, and Runkle (1993)). In Engle and Ng(1993),

$$\sigma_{t+1}^2 = \omega + \alpha(\varepsilon_t - \theta\sigma_t)^2 + \beta\sigma_t^2$$

Besides that, there are several other GARCH option pricing models close to classical GARCH option pricing models. In Heston and Nandi(2000),

$$\sigma_{t+1}^2 = \omega + \alpha(\varepsilon_t - \theta\sigma_t)^2 + \beta\sigma_t^2$$

Thus it is important for a model to capture the two features of underlying asset's volatility, time varying and asymmetric effect of innovations on volatility.

Time Varying Volatility

Even though our model is a discrete time risk factor model, the volatility at time t+1 is known at time t, thus it is still a GARCH option pricing model. Compared with standard option pricing models, our model has some difference in specifying the volatility process. To easily compare with other GARCH models, we use our one factor model as an example.

Taylor(1986) and Schwert (1989) proves that GARCH model can be equivalent to a group of absolute GARCH models, in which the process of standard deviation instead of variance is proposed. A GARCH(1,1) model for the standard deviation σ_t can be written as

$$\sigma_{t+1} = \omega + \alpha\sigma_t|Z_t| + \beta\sigma_t \tag{2.28}$$

By constraining that $\omega > 0, \alpha > 0$ and $\beta > 0$, the volatility σ_{t+1} is guaranteed to be non-negative.

The conditional variance of market index return in our model at time t+1 is σ^2_{t+1},

$$\sigma^2_{t+1} \quad = \quad (c + gX_t)^2 = (c + g\phi X_{t-1} + g\varepsilon_t)^2$$

In our one factor model, the conditional volatility is

$$\sigma_{t+1} \quad = \quad |c + g\phi X_{t-1} + g\varepsilon_t| = |(1-\phi)c + \phi\sigma_t + g\varepsilon_t|$$

GARCH(1,1) model assumes that past volatility and return residual have constant effect on the volatility. However, in our model, we also assume that past volatility has constant effect on the volatility. Further more, we separate random noise from the return residual and assumes that it has a constant effect on the volatility.

g is a key parameter in our model. If $g = 0$, $\sigma^2_{t+1} = c^2$ which is a constant. g is the link connects conditional volatility with past information (risk factor X_{t-1}), which makes the conditional volatility time varying.

Asymmetric Effect of Innovation on Volatility

Engle an Ng (1991) finds that negative return increase future volatility by a larger amount than positive returns of same magnitude on average. By controlling for the asymmetric effect of innovation on volatility, Heston and Nandi (2000) is very successful in capturing cross sectional option prices.

The asymmetric version of the standard deviation of GARCH model is given by,

$$\sigma_{t+1} = \omega + \alpha\sigma_t f(Z_t) + \beta\sigma_t \tag{2.29}$$

where

$$f(Z_t) = |Z_t - b| - c(Z_t - b) \tag{2.30}$$

This indicates that when $Z_t \geq b$, $\sigma_{t+1} = \omega + \alpha(1-c)\sigma_t Z(t) + (\beta - \alpha(1-c)b)\sigma_t$; while when $Z_t < b$, $\sigma_{t+1} = \omega + \alpha(c-1)\sigma_t Z(t) + (\beta - \alpha(c-1)b)\sigma_t$. When Z_t increases from negative to positive, the volatility process will change, which leads Z_t with different sign have different effect on volatility. Empirically, α, c and b are all positive, which indicates that negative innovation at time t will have larger effect on the volatility at time t+1.

In our model, we also separate the region of ε_t into two parts, in which the volatility process will change. If $0 < \phi < 1$, $c > 0$ and $g < 0$, negative innovation will have larger effect on volatility than positive innovation. All the signs or magnitude of the three parameters are consistent with our calibrations from 2005 to 2007. The only difference between our model and asymmetric GARCH model is still that we separate the noise from the return residual and directly research into its effect on volatility.

The asymmetric effect of innovation on volatility can also be explained by a negative correlation between volatility and return in our model. When there is a large negative innovation, the return will be negative. If the correlation between volatility and return is negative, the large negative innovation will drive the volatility in the next time period increase on average. When there is a large positive innovation, the return will be positive. When the correlation between volatility and return is negative, the large positive innovation will drive the volatility in the next time period decrease on average.

In our model, the daily conditional correlation between return and variance is

$$\begin{aligned} Cov_{t-1}[\sigma_{t+1}^2, R_{t-1,t}] &= 2g(c+g\phi X_{t-1})(c+gX_{t-1}) \\ &= 2g\sigma_t^2 + 2g^2(\phi-1)X_{t-1}(c+gX_{t-1}) \end{aligned} \tag{2.31}$$

For our model, the sign of $Cov_{t-1}[\sigma_{t+1}^2, R_{t-1,t}]$ depends on risk factor X_{t-1} and is not clear. However, we can calculate the unconditional expectation of $Cov_{t-1}[\sigma_{t+1}^2, R_{t-1,t}]$ to see the relationship between return and future variance on average,

$$E[cov_{t-1}(\sigma_{t+1}^2, R_{t-1,t})] = 2g(c^2 + g^2\frac{\phi}{1-\phi^2}) \tag{2.32}$$

When $0 < \phi < 1$ and $g < 0$, The conditional covariance $Cov_{t-1}[\sigma_{t+1}^2, R_{t,t+1}]$ is negative. This is consistent with the leverage effect documented in literatures (See Black (1976), Christie (1982)).

Expectation of Conditional Variance

The expectation of conditional variance of underlying asset's return in GARCH option pricing models can be represented as a weighted average between unconditional variance and past conditional variance. For example, in Heston and Nandi (2000),

$$E_{t-1}\sigma_{t+1}^2 = (\beta + \alpha\gamma^2)\sigma_t^2 + (1-\beta-\alpha\gamma^2)E\sigma_t^2 \tag{2.33}$$

The expectation of conditional variance in GARCH (1,1) is a special case in equation (34) when $\gamma = 0$. This makes the relationships among expectation of conditional variance, past conditional variance, unconditional variance constant.

However, in our one factor model, the expectation of conditional variance can be expressed as the weighted average between unconditional variance and past conditional variance plus an extra term, which makes the expectation of conditional variance in a GARCH option pricing model be a special case of ours. For example, the expectation of conditional variance in Heston and Nandi (2000) is a special case of ours when $c = 0$.

$$\begin{aligned} E_{t-1}\sigma_{t+1}^2 &= (c + g\phi X_{t-1})^2 + g^2 \\ &= \phi^2\sigma_t^2 + (1-\phi^2)E\sigma_t^2 + 2gc\phi(1-\phi)X_{t-1} \end{aligned}$$

The extra term $2gc\phi(1-\phi)X_{t-1}$ can adjust the relationships among expectation of conditional variance, past conditional variance, unconditional variance based on risk factors.

From above, we show that the volatility process in our model is similar to that in standard GARCH option pricing model. Our model not only can capture the time varying volatility of underlying asset's return, but also can capture the asymmetric effect of innovation on volatility. In standard GARCH option pricing models, conditional expectation of conditional variance is a linear function of past conditional variance and unconditional variance. In our model, the conditional expectation of conditional variance not only depends on past conditional variance and unconditional variance, but also depends on past volatility. We expect that our model can better fit the option price by incorporating more information to capture the conditional expectation of conditional variance.

2.3 Model Calibration

There are three parts in this section. First of all, we will discuss the methodology of estimating three models, our single factor model, new Heston and Nandi model in Christoffersen, Heston and Jacob (2011)[9], and ad-hoc Black Scholes model. Secondly, we discuss the data. Thirdly, we will make an in-sample and out of sample comparison of these models for option data and underlying market index return.

2.3.1 Methodology of Calibrating Option Pricing Models

The conditional volatility at time $t+1$ of market return in our model is known at t, thus our model is still belong to GARCH type option pricing model. In literature, calibrating GARCH option pricing model usually follows three ways. The first way is to use maximum likelihood estimation on return data. See for instance Engle (1982), Bollerslev (1986), Christoffersen and Jacob (2004) and Chirstoffersen, Heston and Jacobs (2006). However, Christoffersen, Heston and Jacob (2011) note that when valuing options using GARCH models, an additional price of volatility risk can not be estimated from physical return data. Thus effectively this additional risk premium is often set to zero in GARCH option pricing models. Intuitively, in standard GARCH option pricing models, conditional volatility is set as the same under both physical measure and risk neutral measure, while the distribution of conditional volatility under the two measures are different. Christoffersen, Heston and Jacob (2011) propose a new pricing kernel, the conditional volatility under risk neutral probability measure may differ from that under physical measure. The volatility risk premium is an additional parameter to control the difference between conditional volatility under the two measures.

[9]Because Heston and Nandi (2000) is a special case of Christoffersen, Heston and Jacob (2011), we refer it as new Heston and Nandi model from here.

The second way is to estimate GARCH option pricing models using option data only. See for instance Aesi, Engle and Mancini (2008). Following this approach, the conditional volatility is filtered from its process under risk neutral measure. When the process under risk neutral probability measure significantly differ from that under physical measure, estimation from option data may not be consistent with physical return.

To find the estimation consistent with both option data and physical return, a third approach is widely used in literatures. See for instance, Heston and Nandi (2000) and Christoffersen and Jacob (2004). This approach firstly filter the volatility from physical return and then estimate the model using option data by minimizing the loss function of option data only. However, as discussed by Christoffersen, Heston and Jacob (2004), this approach still does not directly specify a loss function of return.

In this paper, we follow the approach proposed by Christoffersen, Heston and Jacob (2011). This approach is similar to the third one. In the first step, we still filter the value of risk factors from physical return. In the second step, we do not only focus on fitting option data, instead we try to estimate our model to fit both physical return and option data. Thus in the second step, we will minimize the sum of the loss function of physical return and the loss function of option data.

Similar to standard approach, we use the log likelihood of the physical return of market index as its loss function. In our model, the conditional daily density of daily return is normal, so that

$$lnR_{t,t+1} = \frac{1}{\sqrt{2\pi}(c+gX_t)^2} exp\left(-\frac{(lnR_{t,t+1}-(b+g\phi)X_t^2-(c\phi+d)X_t-f)^2}{2(c+gX_t)^2}\right) \quad (2.34)$$

The return log likelihood is:

$$LnL^R \propto -\frac{1}{2}\sum_{t=1}^{T} ln(c+gX(t))^2 + [lnR_{t,t+1} - (b+g\phi)X_t^2 - (c\phi+d)X_t - f]^2/(c+gX_t)^2 \tag{2.35}$$

For the option data, we assume that the estimation error $e^c_{i,t} = c_{i,t} - \hat{c}_{i,t}$ follows normal distribution with mean zero and variance σ_c^2. $c_{i,t}$ is the individual option price i at time t and $\hat{c}_{i,t}$ is the estimated individual option price i at time t.

The conditional daily density of estimation error of option data is:

$$e^c_t = \frac{1}{\sqrt{2\pi}\sigma_c} exp\left(-\frac{(e^c_t)^2}{2\sigma_c^2}\right) \tag{2.36}$$

where e^c_t is a vector which represents the estimation error of option data at time t. $\hat{\sigma_c^2}$ is the least square estimator from FOC. The log likelihood of options is

$$LnL^C \propto -\frac{N}{2} - \frac{1}{2N} log\frac{SSE^c}{N} \tag{2.37}$$

where $SSE^c = \sum_{t=1}^{T}\sum_{i=1}^{n_t} {e^c_{i,t}}^2$ is the sum square error of the option contracts, N is the total number of option contracts from time t to T and n_t is the number of option contracts at time t.

In our model, we have 6 free parameters, $\Theta = \{\phi, \phi^*, \omega^*, \Delta^*, c, g\}$. However, the option price in our model only depends on five parameters $\Theta^R = \{\phi^*, \omega^*, \Delta^*, c, g\}$. Intuitively, option price in our model depends on the current risk factors and distribution of innovation under risk neutral probability measure. In our model, current risk factors are the same under physical measure and risk neutral measure. However, the future processes of risk factors under physical measure and risk neutral measure are different. Thus ϕ, which describes the distribution of innovation under physical measure is irrelevant to option price. When we measure the loss function of physical return, a function depends on the distribution of innovations under physical measure, ϕ is relevant.

To fit the option data and physical return simultaneously, we follow a two step procedure. In the first step, given a set value of Θ^R, we get the value of risk factors from physical return. To get a reasonable initial value of risk factor, we set the initial value as the unconditional mean at first and filter the value of initial risk factor from physical return 250 trading days ahead. In the second step, given the set value of Θ and risk factors, we search for an optimal set of Θ to maximize the sum of the two loss function, one from physical return and the other one from option data.

To compare with our model, we calibrate two benchmark models, Christoffersen, Jacob and Heston (2011) and ad-hoc Black Scholes model. Christoffersen, Jacob and Heston (2011) provide a more general more than Heston and Nandi (2000). Compared with Heston and Nandi (2000), their model is proved to better fitting option data by providing a new parameter, volatility risk premium. Their model is a more challenging benchmark model.

The processes of return and volatility under risk neutral measure in Christoffersen, Jacob and Heston (2011) is described as:

$$ln(S(t)) = ln(S(t-1)) + r - \frac{1}{2}h^*(t) + \sqrt{h^*(t)} + \sqrt{h^*(t)}z^*(t)$$
$$h^*(t) = \omega^* + \beta h(t-1) + \alpha^*(z^*(t-1) - \gamma^*\sqrt{h^*(t)})^2 \tag{2.38}$$

where $z^*(t)$ has a standard normal distribution and we can easily match the parameters under risk neutral measure to the parameters under physical measure.

$$z(t) = z^*(t) - \lambda$$
$$h(t) = h^*(t)(1-2\alpha\zeta)$$
$$\omega = \omega^*(1-2\alpha\zeta)$$
$$\alpha = \alpha^*(1-2\alpha\zeta)^2$$
$$\gamma = \gamma^* + \phi \tag{2.39}$$

For Heston and Nandi (2000), it is a special case when $(1-2\alpha\zeta) = 1$. To estimate Christoffersen, Jacob and Heston (2011), we get the conditional volatility from physical return and estimate the physical return and option data simultaneously. To get a reasonable initial value of conditional volatility, the initial value of conditional volatility is also filtered from physical return 250 trading days ahead.

The loss function of physical return in Christoffersen, Jacob and Heston (2011) is also its log likelihood.

$$LnL^R \propto -\frac{1}{2}\sum_{t=1}^{T}\left\{ln(h(t))+(R(t)-r-\lambda h(t))^2/h(t)\right\} \tag{2.40}$$

We also use the same loss function of options for their model for comparison, the loss function of option is equation (51). Heston and Nandi (2000) derive an almost closed-form expression of European option prices by the inversion of the characteristic function technique. If we known the moment generation function of underlying asset price under risk neutral probability measure, European call option price can always be expressed as:

$$\begin{aligned} C_t &= \frac{1}{2}+\frac{e^{-r\tau}}{\pi}\int_0^\infty \Re\left[\frac{K^{-i\phi}f^*(i\phi+1)}{i\phi}\right]d\phi \\ &\quad -Ke^{-r\tau}\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \Re\left[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}\right]d\phi\right) \end{aligned} \tag{2.41}$$

where

$$f^*(\phi)=E_t^Q S_T^\phi = S_t^\phi exp(\alpha_t+\beta_t' X_t + X_t'\gamma_t X_t)$$

In our model, the moment generating function of underlying asset price under risk neutral probability measure is listed in proposition 4. For the new Heston and Nandi model, detailed information about the moment generating function of underlying asset price are listed in Christoffersen, Heston and Jacob (2011).

We use both Fast Fourier Transform theory (FFT) and numerical integration (NI) to computer the integral in equation (60) for our model and new Heston and Nandi model. To calculate the integrals by numerical approach, we evaluate the integrand function on 5,000 equally spaced mid-pints from the interval $(0,100)$. It turns our that option prices calculated by the FFT and NI are very close.

The second benchmark approach is the ad-hoc Black Scholes model. It is well documented that the ad-hoc Black Scholes model outperforms the deterministic volatility option pricing model (See Dumas, Fleming, and Whaley (1998)), because it allow cross sectional options have different implied volatility. The ad-hoc Black Scholes model assumes that the volatilities of the cross sectional options is a second order linear function of the strike price and time to maturity:

$$\sigma_{bs} = a_0 + a_1 K + a_2 k^2 + a_3 \tau + a_4 \tau^2 + a_5 K \tau$$

We first estimate the σ_{bs} for an option with strike K and time to maturity τ by OLS regression. Then we use the estimation of σ_{bs} as the implied volatility for each option and put it back to Black Scholes formula to get the estimation of cross sectional option prices. By minimizing the sum squared error (SSE) of option data, we can find the optimal parameters in ad-hoc Black Scholes model, $a_0, a_1, a_2, a_3, a_4, a_5$. The ad-hoc Black Scholes model is a purely fitting approach, it does not provide much intuition. Here we only used it as a benchmark model to test the capability of our model to fit option data.

2.3.2 Data

Our data is from Jan 5, 2005 to Dec 31, 2008 from Option metrics. To eliminate the potential weekend effect, we only use call options on each Wednesday. All the options in the data are filtered by two criterions,

Table 2.1: Summary statistics of data

Note: This table reports summary statistics of options on SP 500 index. The second column to the sixth column reports the year, number of option contracts, minimum price, maximum price, average price and standard deviation of option prices in each year respectively.

Year	Number	Min	Max	Mean	Std
2005	3587	1	130.8	38.5	34.1
2006	4750	1	146.1	48	39.3
2007	6340	1	183.1	59.4	46.1
2005-2007	14677	1	183.1	52.7	41.6

1) The call price has to be greater than or equal to the spot price minus the present value of the remaining dividends minus the discounted strike price. This is the no-arbitrage relationship in Merton (1973).

2) Option of a particular moneyness and maturity will be represented once in the sample on any particular day.

Detailed information of options are described in Table 1.

2.3.3 In Sample and Out of Sample Model Comparison

We estimate the three models in sample by year. In each year, we estimate the option data and physical return simultaneously for one factor model and new Heston and Nandi model. For ad-hoc Black Scholes model, we estimate the model by option data. In literature, some GARCH option pricing models are estimated by weekly option data. In each week, the number of option contracts are less than 100 on average, given that each of the three models has 6 parameters, this may arise the concern of overfitting the data. All the calibrations of new Heston and Nandi model, one factor model and ad-hoc Black Scholes model are listed in Table 2, Table 3 and Table 4 respectively.

To compare the in sample fitting of option data of the three models, root mean squared error (RMSE) is an important measure. RMSE measures how much the difference between the estimated option price and real option price on average for a model. RMSEs of each model are listed in the upper part of Table 5. RMSEs of one factor model and new Heston and Nandi model are much smaller than those of ad-hoc Black Scholes Model for all the three years. In 2005, the RMSE of one factor model is very close to that of New Heston and Nandi model. However, in 2006 and 2007, the RMSE of one factor model is much lower. Especially in 2007, the RMSE of one factor is over 50 cents lower than that of new Heston and Nandi model. Thus on average, the one-factor model provides a much better in sample estimation of option data.

Out of sample test is an important way to check whether a model is overfitting or not. Here we make a 52 week ahead forecasting. Given the estimation of parameters in the prior year for a model, we estimate option price and physical return in the current year. For new Heston and Nandi model and one factor model, conditional volatility and current risk factor are filtered from physical return. Detailed information of our of sample test is listed in the lower part of Table 5. The RMSE of option data is very large in 2007 for all the three models. Because the 2007-2008 financial crisis starts from the second half of 2007, we project that there is a systematic change in 2007. However, RMSE of one factor model in the out of sample test is still much lower than that of new Heston and Nandi model.

Because risk factor and conditional volatility is filtered directly from physical return in one factor model and new Heston and Nandi model respectively, the physical return is ``perfectly fitted" by the model. To differentiate their way in fitting physical return, we compare the expected returns in the two model.

Figure 1 displays the model implied expected return of one factor model and new Heston and Nandi model from 2005 to 2007. The expected return of one factor model is more volatile than the expected return of new Heston and Nandi model. We also compare the sum of absolute difference between physical return and expected return of the two models, the result is similar. Thus these two models can both capture the physical return. In literature, the time series relationship between market return and its volatility is mixed. Many empirical studies on the time series return and risk relationship based on ICAPM model find that more risk factors should be included to predict the market return. We expect that our multi-factor model can be better predicting market return than new Heston and Nandi model. Detailed discussion of the risk and return relationship by comparing the two models is in section 5.

In this section, we compared the capability of the three models in fitting option data. In the in sample fitting, the RMSE of one factor is close to the RMSE of new Heston and Nandi model in 2005, but its RMSE is much lower (50 cents lower in 2007) than the RMSE of new Heston and Nandi model in 2006 and 2007. In the out of sample test, the RMSE of one factor model is much lower than the RMSE of New Heston and Nandi model for both 2006 and 2007. To sum up, our one factor model can fit option data better than New Heston and Nandi model both in sample and our of sample in our data.

2.4 Possible Solution to Several Stylized Facts

In this section, we show that our model can explain the following stylized facts:

1. The U shape relationship between the pricing kernel and the return.

2. The implied volatility is higher than physical volatility on average.

3. density function of return under risk neutral measure has fatter tails than that under

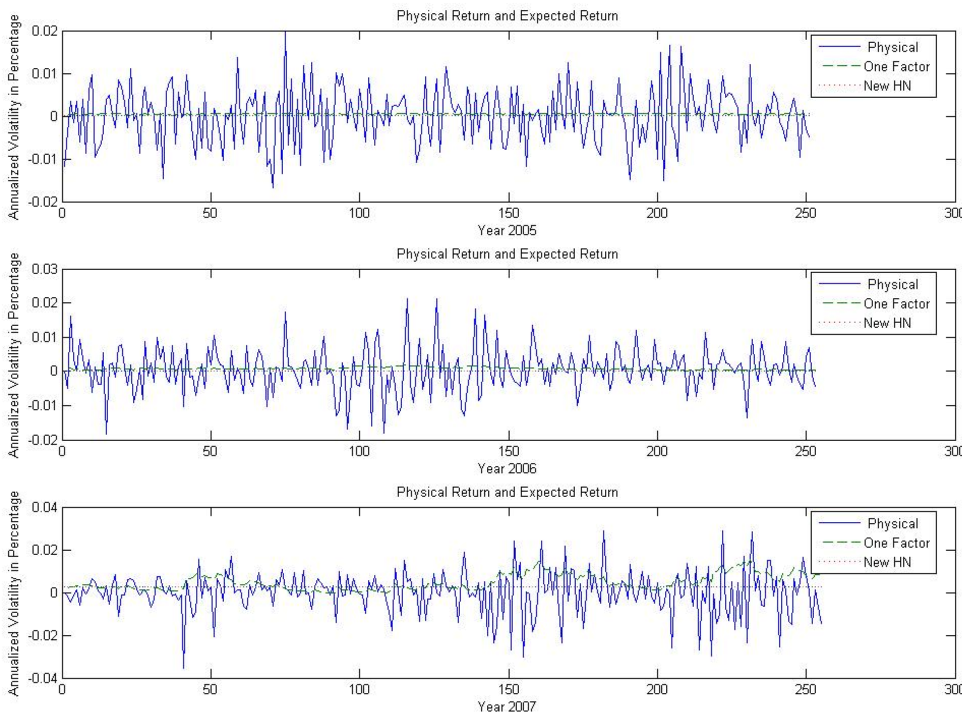

Figure 2.1: Physical return and expected return

This figure describes the physical return and model specified expected return from 2005 to 2007. '-' stands for the physical return, '- -' stands for the expected return from one factor model. '...' stands for the expected return new HN model.

Table 2.2: In sample calibration of Christoffersen, Heston and Nandi (2011)

Parameters estimates are obtained by optimizing an joint likelihood on return and options. All physical parameters can be directly match from risk neutral parameters. The initial volatility is filtered from physical return 250 trading days ahead. Detailed information of data is described in table 1. To make the volatility always be positive, we follow Christoffersen, Heston and Jacob (2011) to constrain $\omega^* = 0$.

Year	Risk Neutral Parameters					
	λ	1-2$\alpha\zeta$	ω^*	β^*	γ^*	α^*
2005	1.96	0.85	0	0.81	213.83	3.16E-06
2006	10.59	0.75	0	0.74	300.22	2.47E-06
2007	12.39	1.10	0	0.38	1041.95	0.57E-06

Table 2.3: In sample calibration of one factor model

In the one factor model, option price depends on 5 free parameters ω, ϕ^*, c, g and Δ^*, while the process of physical return depends on all the six parameters. Parameter estimation are obtained by optimizing an joint likelihood on return and options. Detailed information of data is described in table 1.

Year	Risk Neutral Parameters					
	ϕ	ω	ϕ^*	c	g	Δ^*
2005	0.98	-0.076	0.98	1.17	5.16E-03	-2.83E-04
2006	0.96	-0.091	0.97	1.21	5.58E-03	-3.88E-04
2007	0.91	-0.027	0.99	1.05	4.82E-03	-9.42E-04

Table 2.4: In sample calibration of Ad-hoc Black Scholes model

In the one factor model, option price depends on 5 free parameters ω, ϕ^*, c, g and Δ^*, while the process of physical return depends on all the six parameters. Parameter estimation are obtained by optimizing an joint likelihood on return and options. Detailed information of data is described in table 1.

Year	Risk Neutral Parameters					
	a_0	a_1	a_2	a_3	a_4	a_5
2005	2.35	-3.11E-03	1.04E-06	-1.54	0.11	1.27E-03
2006	2.83	-3.64E-03	1.20E-06	-1.46	0.43	9.95E-04
2007	0.95	-6.66E-04	9.58E-07	-1.12	0.60	6.21E-04

Table 2.5: In sample and out of sample pricing errors of different models

This table reports the root mean square error (RMSE) of options in 2005, 2006, 2007 for one factor model, new Heston and Nandi model and ad-hoc black Scholes model. Given one set of optimal parameters, sum squared error (SSE) of option prices are calculated for each model. RMSE is defined as the square root of SSE divided by the number of options. We estimate the out of sample RMSE by forecasting option price 52 weeks ahead. Risk factors in one factor model and conditional volatility in new Heston and Nandi model are filtered by physical market return to calculate the out of sample RMSE for each model.

In Sample RMSE			
Model	2005	2006	2007
One factor Model	1.48	1.84	3.29
New Heston and Nandi	1.47	2.04	3.85
Ad-hoc Black Scholes	1.74	2.50	9.70
Out of Sample RMSE			
Model	2006	2007	
One factor Model	2.14	9.13	
New Heston and Nandi	2.42	9.31	
Ad-hoc Black Scholes	4.28	14.68	

physical measure.

2.4.1 The U Shape Relationship Between the Pricing Kernel and the Return

In consumption based model, Harrison and Kreps (1979) and Hansen and Jagannathan (1991) show that the pricing kernel m_{t+1} is the investor's intertemporal marginal rate of substitution of consumption, $\frac{U'(C_{t+1})}{U'(C_t)}$. In intertemporal models, the growth rate of market consumption is often proxied by the market return in literature. If when the utility is the familiar power utility in Rubinstein's (1976), the pricing kernel will be a linear function of market return. When market return increases, the pricing kernel will decrease. This monotonic inverse relationship between the pricing kernel and market return is intuitive, when the market return increases, the required rate of return will increase, and then the pricing kernel will decrease. The monotonic inverse relationship between the pricing kernel and the market return is widely accepted in literature. The pricing kernel explicitly or inexplicit from many models are all based on the widely accepted monotonic inverse relationship between the pricing kernel and market return. Recently A lot of researches find that there exists a monotone relationship between the market return and the pricing kernel. Harvey and Sid-digue (2000) and Dittman (2000) specify the pricing kernel as a quadratic function and a cubic function of the market return in the cross section of returns. Bansal and Viswanathan (1993) and Ronsenberg and Engle (2002) using semi-parametric and parametric approach to test the function form of pricing kernel. They both find that a linear pricing kernel of market return is rejected. Recently, in the option market, Bakshin, Madan and Panayotov (2010) use a model free approach to test the function of pricing kernel in option market. They find that the pricing kernel in option market is a U shape function of the market return by testing different type of options in equity market. Christoffersen, Heston and Jacob-s (2011) also present new semi-parametric evidence to confirm the U shape relationship between the risk-neutral and physical probability densities.

In our model, we explicitly specified the market index return and the pricing kernel.

$$ln(R_{t,t+1}) = X_t' b X_t + c' X_{t+1} + d' X_t + X_t' g X_{t+1} + f \tag{2.42}$$

$$ln(m(t,t+1)) = X_{t+1}' A X_{t+1} + X_t' B X_t + C' X_{t+1} + D' X_t + X_t' G X_{t+1} + F \tag{2.43}$$

In our model, the logarithm of daily market index return is a linear function of current risk factor X_{t+1}, while the logarithm of the daily pricing kernel is a quadratic function of current risk factor X_{t+1} conditional at time t. The specifications of pricing kernel and market index return in our model explicitly provides a solution to the non-monotone relationship between the pricing kernel and market index return.

The conditional relationship between the pricing kernel and the market index return in our one factor model depends on the value of risk factor X_t, A and c. Thus the time varying relationship between the pricing kernel and the market index return[10] documented in literature can be explained by the time varying risk facto X_t. Because the unconditional expectation of X_t is 0, the unconditional relationship between the daily pricing kernel and the market index return depends on A and c. When $A > 0$ and $c > 0$, there exists a U shape relationship between the daily pricing kernel and the daily market index return on average. $A > 0$ requires that the variance under risk neutral measure of risk factor $\Omega^* > 1$ in our model. If $c > 0$ and $\Omega^* > 1$, our model can explicitly describes the unconditional U shape relationship between the pricing kernel and market index return on a daily base.

In literature, the U shape relationship between the pricing kernel and the market index return is mainly from option data. Most of them test their unconditional relationship. This is equivalent to say that the pricing kernel $\frac{M(t+1)}{M(0)}$ is a monotone function of $R_{0,t+1}$. Even though Christofferen, Heston and Jacob (2011) tests their conditional relationship, but their relationship is based on one month horizon, which is equivalent to test the conditional relationship between $\frac{M(t+30)}{M(0)}$ and $R_{0,t+1}$. Here we only give insights on that our model can explain the unconditional monotone relationship between the daily pricing kernel and daily market index return. Because the process of pricing kernel and market index return for a long horizon is very complicated, theoretically proved the unconditional U shape relationship between the pricing kernel and the market index return is very difficult. But we still can simulate their unconditional relationships on a longer horizon based on the calibrations of our model.

[10] equivalent to conditional relationship between the pricing kernel and the market index return

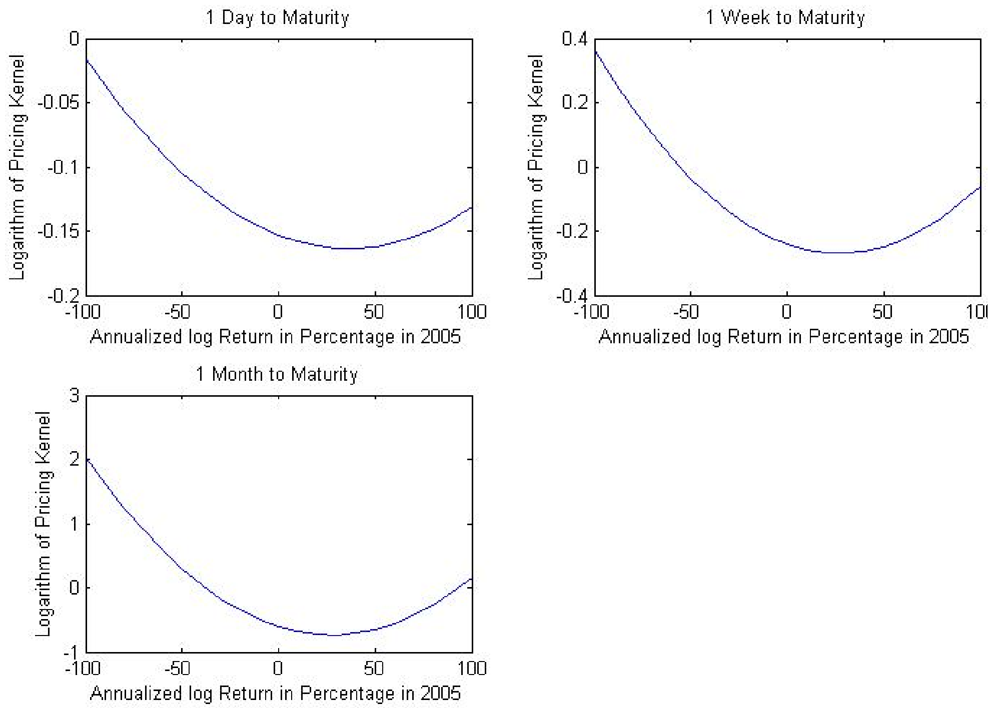

Figure 2.2: Relationship between pricing kernel and return
This figure describes the model implied relationship between the logarithm of pricing kernel and log return in different time to maturity in 2005. The pricing kernel is estimated by polynomial fitting of simulated data. For different time to maturity, we simulate 5,000 to get a robust estimation of pricing kernel. The initial value of conditional volatility is set as the unconditional expectation of volatility.

We simulate the pricing kernel and the market index return for different horizon (one day, one week and one month) based on the calibrations in different year. We then use a polynomial function to fit the natural logarithm of pricing kernel and the natural logarithm of the market return. Figure 3 to 5 describes the model implied U relationship between the logarithm of pricing kernel and the logarithm of the market index from 2005 to 2007.

2.4.2 Implied Volatility Larger than Realized Volatility

Empirical studies usually find that implied volatility from option market is larger than physical volatility on average. Figure 5 displays the annualized physical volatility filtered from GARCH (1,1) model and the annualized option implied volatility from VIX index

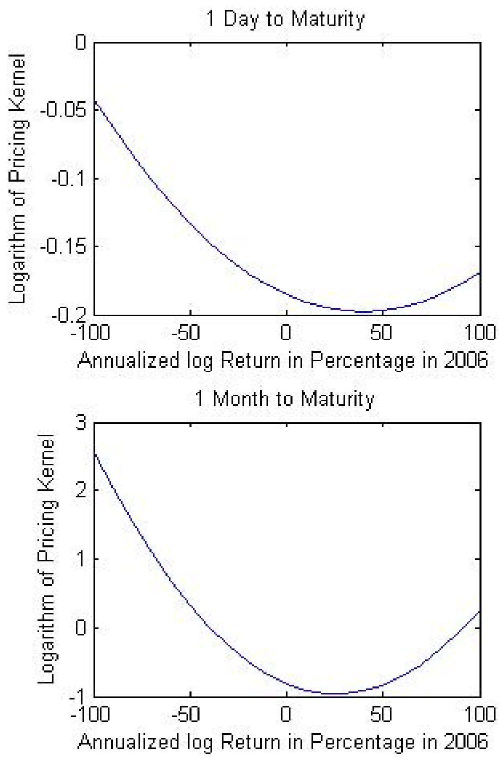

Figure 2.3: Relationship between pricing kernel and return
This figure describes the model implied relationship between the logarithm of pricing kernel and log return in different time to maturity in 2006. The pricing kernel is estimated by polynomial fitting of simulated data. For different time to maturity, we simulate 5,000 to get a robust estimation of pricing kernel. The initial value of conditional volatility is set as the unconditional expectation of volatility.

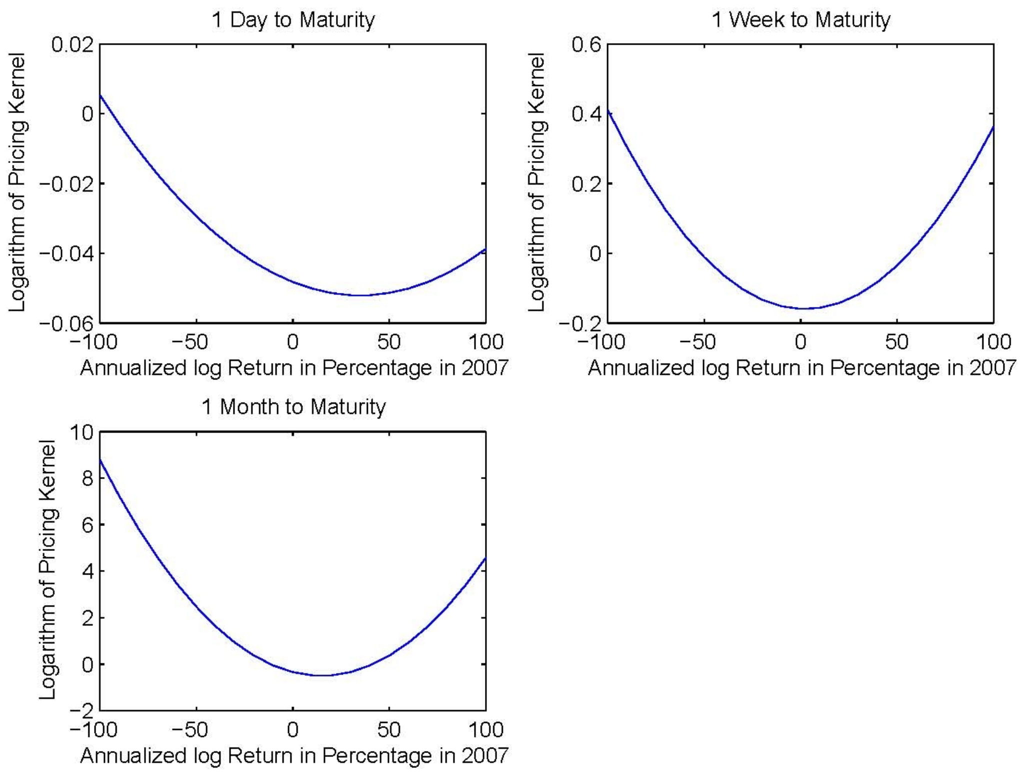

Figure 2.4: Relationship between pricing kernel and return

This figure describes the model implied relationship between the logarithm of pricing kernel and log return in different time to maturity in 2007. The pricing kernel is estimated by polynomial fitting of simulated data. For different time to maturity, we simulate 5,000 to get a robust estimation of pricing kernel. The initial value of conditional volatility is set as the unconditional expectation of volatility.

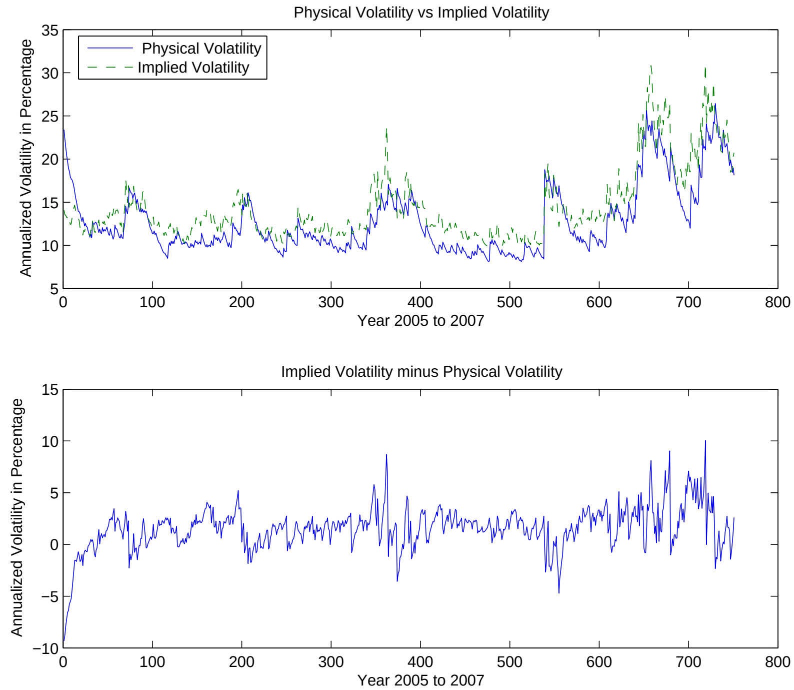

Figure 2.5: Physical volatility and option implied volatility
We plot the physical volatility and option implied volatility from 2005 to 2007. The annualized physical volatility is filtered from GARCH (1,1) model and the annualized option implied volatility from VIX index.

from 2005 to 2007. On average, implied volatility is higher than physical volatility.

In stochastic volatility model, a negative volatility risk premium is often used to explain the difference between implied volatility and physical volatility. However in discrete GARCH option pricing models, the volatility risk premium is often assumed to be zero which leads the physical volatility consistent with implied volatility. Christoffersen, Heston and Jacob (2011) propose a new Heston and Nandi model, in which a volatility risk premium is used to explain the difference between the physical volatility and implied volatility. Our model can also capture the dispersion between physical volatility and implied volatility.

Our model is a GARCH type model, at time t, the volatility of return at time t+1 is known. Conditional at time t, the variance of return at time t+1 under physical measure and risk neutral measure ($Var_t^P(logR_{t,t+1}) = \sigma_{t+1}^2$ and $Var_t^Q(logR_{t,t+1}) = \sigma_{t+1}^{*2}$) are

$$\sigma_{t+1}^2 = (c + gX_t)^2 \tag{2.44}$$

$$\sigma_{t+1}^{*2} = (c + gX_t)^2\Omega^* \tag{2.45}$$

When $\Omega^* > 1$, the instantaneous volatility of return under risk neutral measure is higher than that under physical measure. This is equivalent to say that the instantaneous physical volatility is lower than the instantaneous implied volatility when the logarithm of pricing kernel is a quadratic function of risk factors ($A > 0$).

By comparing the expectation of instantaneous volatility of return under physical measure and risk neutral measure, we can understand their relationship on average. The variance of return at time $t+1$ under real probability measure and risk neutral probability measure can also be expressed as,

$$\sigma_{t+1}^2 = (c + g\phi X_{t-1} + g\varepsilon_t^P)^2 \tag{2.46}$$

$$\sigma_{t+1}^{*2} = (c + \omega^* g + g\phi^* X_{t-1} + g\Sigma^*\varepsilon_t^Q)^2\Omega^* \tag{2.47}$$

Conditional on time t-1, the variance of return under P and Q is,

$$E_{t-1}^P\sigma_{t+1}^2 = (c + g\phi X_{t-1})^2 + g^2 \tag{2.48}$$

$$E^Q_{t-1}\sigma^2_{t+1} = (c+\omega^* g + g\phi^* X_{t-1})^2\Omega^* + g^2\Omega^{*2} \tag{2.49}$$

The conditional volatilities under the two different measure depend on risk factor X_{t-1} in our model, which implies that implied volatility is not always higher than physical volatility.

To explain that the physical volatility is lower than implied volatility on average, we need to calculate the unconditional expectation of conditional volatility under the two measures. The unconditional expectation of conditional volatility describes the conditional volatility of market index on average. By comparing the unconditional expectation of conditional volatility under the two different measure, we can get insights on the relationship between the physical volatility and implied volatility.

$$E^P\sigma^2_{t+1} = c^2 + \frac{g^2}{1-\phi^2} \tag{2.50}$$

$$E^Q\sigma^2_{t+1} = (c+\frac{g\omega^*}{1-\phi^*})^2\Omega^* + \frac{g^2\Omega^{*2}}{1-\phi^{*2}} \tag{2.51}$$

When $c>0$, g and ω^* has the same sign, $0<\phi^*<1$, $\Omega^*>1$ and ϕ is close to ϕ^*, the unconditional expectation of conditional variance of return under risk neutral measure is larger than that under physical measure. In our calibrations, $c>0, g<0, \omega^*<0, \Omega^*>1$, ϕ and ϕ^* are very close, this leads that the physical volatility is lower than the implied volatility on average.

To better illustrate the relationship between physical volatility and implied volatility in our model, we filtered the instantaneous physical volatility and implied volatility based on

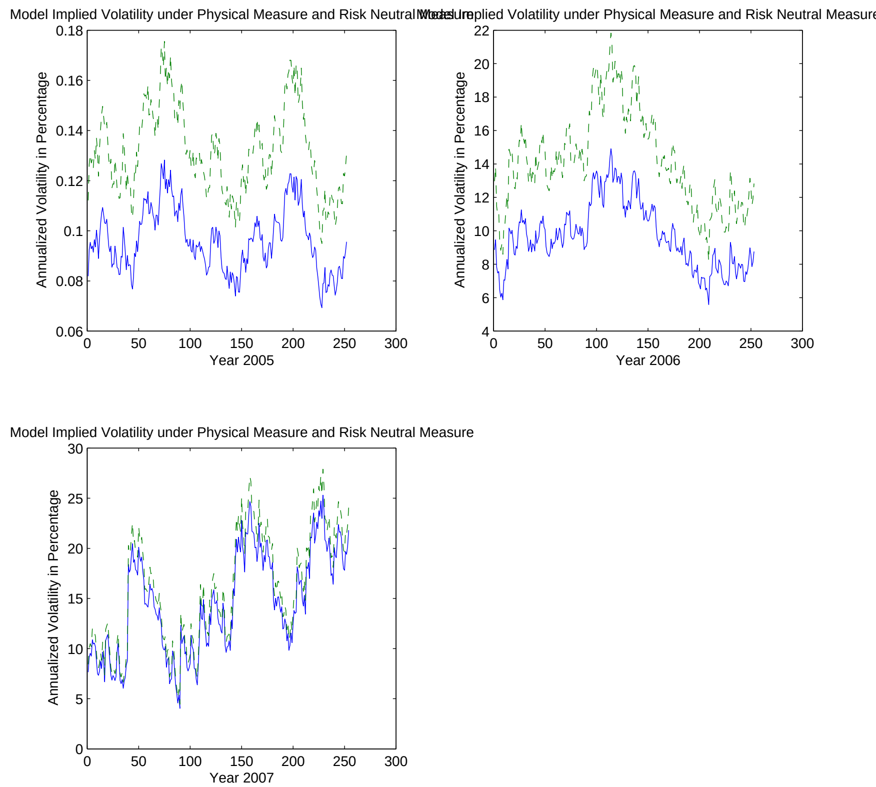

Figure 2.6: Implied volatility under physical measure and risk neutral measure
This figure describes the model implied volatility under physical measure and risk neutral measure from 2005 to 2007. '-' stands for the volatility under physical measure , '- -' stands for the volatility under risk neutral measure. All the volatility is calculated from the parameters in table 3.

the calibrations of our model from 2005 to 2007.

2.4.3 Positive Equity Risk Premium and Fatter Tails

In GARCH option pricing models, the equity risk premium is often assumed to be proportion to conditional variance. Empirical studies find that the equity risk premium is positive on average. Intuitively, investors need to be compensated by taking the risk (variance), thus the average return under risk neutral measure should be less than that under physical measure.

In our one single risk factor model, the return processes under physical measure and risk neutral measure are:

$$ln(R_{t,t+1}) = (b+g\phi)X_t^2 + (d+c\phi)X_t + f + (c+gX_t)\varepsilon_t^P$$
$$ln(R_{t,t+1}) = (b+g\phi^*)X_t^2 + (d+c\phi^* + g\omega^*)X_t + f + c\omega^* + (c+gX_t)\Delta^*\varepsilon_t^Q \quad (2.52)$$

and

The equity risk premium (ERP) in our model can be represented as:

$$ERP = -(c+gX_t)[(\phi^* - \phi)X_t + \omega^*] \quad (2.53)$$

When the unconditional expectation of EPR is positive, the indicates that the EPR is positive on average. The unconditional expectation of EPR in our model is:

$$E[EPR] = -\frac{g(\phi^* - \phi)}{(1-\phi)^2} \quad (2.54)$$

If $g < 0$ and $\phi < \phi^* < 1$, on average, the ERP will be positive. If $\phi = \phi^*$, there is no equity risk premium in our model. Jackwerth(1999) gives a literature review on option-implied risk neutral distributions. Even though there are different approaches to estimate option-implied risk neutral distributions, they are based on the same theory. In theory, in an complete market, Breeden and Litzenberger(1978) gives an exact formula for calibrating the option-implied risk neutral distributions. The second derivative of a European call price C taken with respect to its strike price K is the state-contingent price π on the future asset price ending up at exactly the strike price of the option.

$$\frac{\partial^2 C}{\partial K^2} = \pi_{S=K} \tag{2.55}$$

It is documented that the option-implied risk neutral return distribution has relative fat tails than the physical distribution (See Jackwerth (1999, 2000), Ronsenberg and Engle (2002) and others). This leads us to check the higher order information of market index return.

The conditional variance and kurtosis of return under P and Q are,

$$Var^P_{t-1} ln(R_{t,t+1}) = (c + gX_t)^2 Var^Q_{t-1} ln(R_{t,t+1}) = \Omega(c + gX_t)^2 \tag{2.56}$$

$$Kurtosis^P_{t-1} ln(R_{t,t+1}) = 3(c + gX_t)^4 \tag{2.57}$$

$$Kurtosis^Q_{t-1} ln(R_{t,t+1}) = 3\Omega^{*2}(c + gX_t)^4 \tag{2.58}$$

and

When $\Omega^* > 1$, the conditional Kurtosis of return under risk neutral measure is always larger than that under physical measure. This implies that the probability density function of return under risk neutral measure may have fatter tails than that under physical measure.

From above, Ω^* adjusts the difference between higher order moment expectation of return under risk neutral measure and physical measure. When $\Omega^* > 1$, the conditional variance and kurtosis of return under risk neutral measure is larger than that under physical measure, which implies that the distribution of log return under risk neutral measure has fatter tails than that under physical measure.

In Heston and Nandi (2000), the variance of conditional variance is a linear function of past conditional variance under physical measure and risk neutral measure can be both written as:

$$Var_{t-1}\sigma_{t+1}^2 = 2\alpha^2 + 4\alpha^2\gamma^2\sigma_t^2 \tag{2.59}$$

Thus the variance of conditional variance under physical measure and risk neutral measure are the same in Heston and Nandi (2000). However, Christoffersen, Heston and Jacob (2011) proposes a new Heston and Nandi model, in which the variance risk premium can differentiate the variance of conditional variance under physical measure and risk neutral measure. In our model, we have similar features.

In our model, the conditional variance of σ_{t+1}^2 under physical measure can be represented as a linear function of σ_t^2 and $X(t-1)$.

$$Var_{t-1}\sigma_{t+1}^2 = 3g^4 + 4g^2\phi^2\sigma_t^2 + 4g^2[c^2(1-\phi^2) + 2gc\phi(1-\phi)X_{t-1}] \tag{2.60}$$

However, under risk neutral measure,

$$Var_{t-1}^Q\sigma_{t+1}^2 = 3g^4\Omega^{*2} + 4g^2\phi^{*2}\sigma_t^{2Q} + 4g^2\Omega^*(c-\phi^*c+\omega^*g)(c+\phi^*c+\omega^*g+2\phi^*X_{t-1}) \tag{2.61}$$

When $c = 0$, the conditional variance of conditional variance in our model is similar to Christoffersen, Heston and Jacob (2011), when $\omega < 0, g > 0$ and $\Omega^* > 1$, the variance of conditional variance under risk neutral measure is always larger than that under physical measure. When $c \neq 0$, our model provides a more general setting to describe the relationship between the variances of conditional variance under physical and risk neutral measure.

2.5 Discussion

2.5.1 Relationship Between Return and Risk

The Merton (1973) seminar paper implies the following equilibrium relationship between risk and return:

$$\mu_i - r = \gamma\sigma_{im} + \gamma_z\sigma_{iz} \tag{2.62}$$

where σ_{im} and σ_{iz} are the covariance of individual stock with market return and with future investment opportunity. Many time series literatures and cross sectional literatures are based on the ICAPM model. However, researches in this two areas are often separate. In this section, we will discuss the linkage of our model with time series literature and cross sectional literatures on asset return in equity market.

Time Series Relationship

Empirical studies usually focus on the time series implication of Merton's ICAPM model in equilibrium and narrowly apply it to the market portfolio. When the hedging component $\sigma_{iz} = 0$, this leads to the following risk-return relation for the market portfolio:

$$\mu_m - r = A\sigma_m^2 \tag{2.63}$$

If the investment opportunity is stochastic, literatures often project σ_{iz} linearly to state variables X (risk factors). The risk-return relationship for the market portfolio becomes:

$$\mu_i - r = A\sigma_{im} + \gamma X \tag{2.64}$$

Popular GARCH option pricing models usually neglect the risk factor X and simply assume that expected excess return of market portfolio only depends on its variance. In most cases, the equity premium is identified as plosive in many empirical studies of GARCH option pricing models. While in our model, the return and conditional variance are both quadratic functions of risk factors. Two things make our model differ from GARCH pricing model in the implication of the time series relationship between return and risk.

Firstly, GARCH option pricing model assumes that return are only compensated by its conditional variance, while the risk and return relationship may be mixed in our model.

Although the linear relationship between expected excess return and variance is simple and intuitive, their relationship in literatures are mixed. For example, French, Schwert and Stambaugh (1987), Campbell and Hentchel (1992), Harrison and Zhang (1993), Goyal and Santa-Clara (2003) and Bollerslev and Zhou (2006) all find an insignificant relationship between return and variance by using different data and different measure of variance. Other studies even find that the intertemporal relationship between variance and return is negative (See Campbell (1987), Breen, Glosten and Jjagannathan (1989), Whitelaw (2000) and Brandt and Kang (2004)). Also there are some studies support the positive relationship between the return and risk (See Chou (1988), Bollerslev, Engle and Wooldridge (1988), Scruggs (1998) and Ghysels, Santa-Clara, and Valkanov (2005)).

In our model, given that risk factors follow OU processes, the conditional variance of market excess return is a quadratic function of risk factors $(c+gX_t)'(c+gX_t)$, while the expected excess return is also a quadratic function of risk factors, $X_t'(b+g\phi)X_t+(d+c\phi)X_t+f$. Because both the excess return of market return and its variance are quadratic function of risk factors, their relationship can be mixed. Here we use a simple one-factor model as an example. When $b+g\phi>0$ and $\frac{c}{g}=\frac{d+c\phi}{2(b+g\phi)}$, the return and variance is positively related. When $b+g\phi<0$ and $\frac{c}{g}=\frac{d+c\phi}{2(b+g\phi)}$, the return and variance is negatively related. When $\frac{c}{g}\neq\frac{d+c\phi}{2(b+g\phi)}$, there is no monotone relationship between the return and the risk.

Secondly, GARCH option pricing model neglect the changes in investment opportunities, while changes in investment opportunities can be proxied by risk factors in our multi-factor model. If the changes in investment opportunities are not significant, there is no need to justify that. However, Guo and Whitelaw (2006) identifies two components of expected returns-the risk component and the component due to the desire to hedge changes in vestment opportunities. They find that expected returns are driven primarily by the hedging component. Also many empirical studies show that risk factors proxy for the hedging need of stochastic investment opportunities are significant in explaining return. For example, Bali and Engle (2010) explores the time series relationship between expected returns and risk for a large cross section of industry and size/book-to-market portfolios. They find that the HML is a priced risk factor and can be viewed as a proxy for investment opportunities;

Allen, Bali and Tang (2012) find a bank-specific systemic risk priced for financial firms. A usual way to proxy the innovations of future investment opportunities is to project them to several risk factors. In a time series analysis, these risk factors are assumed to be followed by OU processes. For example, in Guo and Whitelaw (2006), they use a linear function of risk factors to proxy the condition variance of market variance, these risk factors follow OU processes. Besides that the hedging component in their paper is also proxied by a linear function of risk factors which follow OU processes. In our model, we also assume that risk factors follow OU processes. We can easily justify both variance of market return and hedging component.

Cross Sectional Relationship

Besides the time series relationship between risk and return, cross sectional relationship between risk and return is one of the most important topics in finance. Many literatures often study these two topics separately. For example, GARCH option pricing models only describe the time series process of market return, there is no further information about its implication on the cross sectional relationship between risk and return. For our model, we not only link time series literatures but also the cross sectional literatures.

In cross sectional literatures, researchers mainly focus on several strands to explain cross sectional anomalies in equity market. The first strand is the linear multi-factor model. This strand in literatures focus on adding more systematic risk factors or innovations from future investments based on constant ICAPM (Merton 1973) or APT. Examples are Fama-French's 3 factor model, Cahort's 4 factor model, etc. The second strand is conditional CAPM model, which assumes that the beta of individual stock or well diversified portfolio is time varying and depends on several instruments (See Campebell and Cochrane (1999),

Table 2.6: CCAPM and ICAPM

Model	Example	Risk factors
CCAPM	Petkova and Zhang (2005)	$[DIV_t, DEF_t, TERM_t, TB_t]$
Return and Risk	$E_t r_{i,t+1} = (\alpha_i + \beta_i X_t)'(\alpha_i + \beta_i X_t)$ $E_t r_{M,t+1} = \alpha_M + \beta_M X_t$	
ICAPM	Petkova (2006)	$[R_{M,t}, DIV_t, DEF_t, TERM_t, RF_t, R_{HML,t}, R_{SMB,t}]$
Return and Risk	$E_t r_{i,t+1} = \alpha_i + \beta_i X_t$	

Petkova and Zhang (2005)). Similar to Conditional CAPM model, Bali and Engle (2009) assume that the beta of risk factors in ICAPM models are also time varying and depends on several instruments.

These three models can be nested within our model. Table 6 gives out the examples how our model can accommodate these three different settings.

We can easily express our model in the form of CCAPM or ICAPM models. For instance, in Petkova and Zhang (2005), the expected excess return of individual stock or portfolio is a quadratic function of risk factors which is the same as our setting. We can also justify expected market return as a linear function of risk factors by setting A, B and G as zero matrix. In Petkova (2006), the market excess return is not specified. In our model, the excess return of market portfolio is a quadratic function of risk factors. Given that the expected return of individual stock or portfolio is also a quadratic function, we can always express the expected return of individual stock or portfolio as a linear function of market return and other risk factors.

Given the good performance of these two models in explaining cross sectional variation of stocks in equity market, we expect that our model can also have good explaining power in this topic. However, in literatures, What optimal risk factors should be included is still debatable. In this paper, we do not focus on analyzing the cross sectional variation across stocks in equity market. We only provide a framework for future studies.

2.5.2 Links to Bond Market

In term structure literatures, there two main stands, affine term structure models (ATSMs) and quadratic term structure models (QTSMs). However, ATSMs generates negative interest rates with positive probability. This raise concerns about arbitrage possibilities and their real-time applicability. In contrast, QTSMs guarantees a positive interest by structure. The quadratic relationship between interest and risk factors make the model flexible in capturing bond derivatives. Ahn, Dittmar and Gallant (2001) and Leippold and Wu (2000) suggest that QTSMs can potentially outperform ATSMs. Leippold an Wu (2002) provides a theoretic framework for QTSMs. They prove that the necessary and sufficient conditions for the QTSMs are:

1. interest rate $r(X_t) = X_t' A_r X_t + b_r' X_t + c_r$ with $A_r \in R^{n \times n}, b_r \in R^n$ and $c_r \in R^n$
2. $X_{t+1} = a^* + b^* X_t + \Sigma$, $a^* \in R^n$, $b^* \in R^n$ and Σ is a constant matrix.

If equity market and bond market are driven by the same risk factors, we can easily adjust the QTSMs in our frame work. By incorporating the interest as a quadratic function of risk factors, we can estimate the model by combining data in equity and bond market together.

2.6 Conclusion

In this paper, we propose a general multi-factor model with quadratic pricing kernel. Following the standard multi-factor models, the return of an underlying asset is a linear function of multi-risk factors, while the coefficients are time varying and proxied by past risk factors. It turns out that many standard ICAPMs and CCPMs are nested within our setting. By assuming all the risk factors follow OU processes, our model can capture the time series variation of an asset's return. We also show that the volatility process in our model are similar to standard GARCH option pricing models. By examining option data from 2005 to 2007, our simple one factor model can outperform both Christoffersen, Heston and Nandi (2011) and ad-hoc Black Scholes model in sample and our of sample.

This model also provides possible explanations for the U shape relationship between the pricing kernel and market index return, the implied volatility puzzle and fat tails of risk neutral return density function relative to the physical distribution. Most importantly, this is the first model which provides a unified setting to link cross sectional literatures, time series literatures, option pricing literatures and term structure literatures.

Our model provides a potential framework to research into several interesting questions. Whether bond market and equity market are driven by the same common risk factors? what risk factors are important in equity derivative market? Whether our model can fit the derivatives in equity and bond market simultaneously? All these are left for future works.

Proof of proposition 1

From non-arbitrage theory, the expected return of risky assets under risk neutral measure is risk free rate, we can get

$$E_t^Q[exp(\lambda'\varepsilon_{t+1})] = E_t[exp(r_{t,t+1} + \lambda'\varepsilon_{t+1})m_{t,t+1}]$$

We also know that $E[exp(r_{t,t+1})m(t,t+1)] = 1$. By expanding $E_t[exp(r_{t,t+1} + \lambda'\varepsilon_{t+1})m_{t,t+1}]$ and substitute the expression of $E[exp(r_{t,t+1})m(t,t+1)]$ in it, we can get

$$E_t[exp(r_{t,t+1} + \lambda'\varepsilon_{t+1})m_{t,t+1}] = exp(\frac{1}{2}\lambda'W\lambda + (((2\omega + 2\phi X_t)'A + C' + X_t'G))'\Sigma W\lambda)$$

where $W = (I - 2\Sigma' A\Sigma)^{-1}$.

This implies that there is a linear transformation for the innovation ε_{t+1} from physical measure to risk neutral measure.

$$\varepsilon_{t+1} = \mu_t + W^{\frac{1}{2}} \varepsilon^{Q}_{t+1}$$

where $W^{\frac{1}{2}}(W^{\frac{1}{2}})' = W$ $\mu_t = W'\Sigma(2A\omega + C + (2A\phi + G)X_t)$ and ε^{Q}_{t+1} is a standard normal distribution under risk neutral probability measure.

Proof of proposition 2

We assume that the logarithm of pricing kernel is a quadratic function risk factors:

$$ln[m(t,t+1)] = X'_{t+1}AX_{t+1} + X'_tBX_t + C'X_{t+1} + D'X_t + X'_tGX_{t+1} + F$$

Because $(I-2A)^{-1} = \Omega^*$, $\omega = 0$ and $\Omega = I$, by using the formula $E[exp(\varepsilon'\beta\varepsilon + \gamma\varepsilon)] = exp(-\frac{1}{2}log|I-2\beta| + \frac{1}{2}\gamma(I-2\beta)^{-1}\gamma')$, we can expand the expectation of pricing kernel as:

$$\begin{aligned} E_t[m(t,t+1)] &= exp[X'_t(\phi'A\phi + B + \phi'G + \frac{1}{2}(2\phi'A + G)\Omega^*(2\phi'A+G)')X_t \\ &\quad +(C'\phi + D' + C'\Omega^*(2A'\phi + G))X_t + F - \frac{1}{2}log|\Omega^{*-1}| + \frac{1}{2}C'\Omega^*C] \end{aligned}$$

To make $E_t[m(t,t+1)]$ always equal to $exp(-r_{t,t+1})$, $E_t[m(t,t+1)]$ should not depend on X_t. thus the coefficients of X_t and X'_tX_t must be zero and the constant term is equal to $-r_{t,t+1}$.Thus we have three constraints:

$$\phi'A\phi + B + \phi'G + \frac{1}{2}(2\phi'A + G)\Omega^*(2\phi'A+G)' = 0$$

$$C'\phi + D' + C'\Omega^{*}(2A'\phi + G) = 0$$

$$r_{t,t+1} + F - \frac{1}{2}log|\Omega^{*-1}| + \frac{1}{2}C'\Omega^{*}C = 0$$

Proof of proposition 3

Because the return of risky asset should be risk free rate under risk neutral measure, we have

$$\begin{aligned} E_t[R_{t,t+1}m(t)] &= E_t exp(X'_{t+1}AX_{t+1} + X'_tBX_t + (c' + C')X_{t+1} \\ &\quad + (D' + d')X_t + X'_{t+1}(G + g)X_t + F + f) \end{aligned}$$

By subtracting $E_t[m(t,t+1)]$ from $E_t[R_{t,t+1}m(t)]$, we can have $E_t[R_{t,t+1}m(t)] - E_t[m(t,t+1)] = 0$ independent of X_t. Expanding it and let the coefficients of X_t and $X_t'X_t$ and the constant term equal to 0, we have other three constraints:

$$b + g\phi^{*} + \frac{1}{2}g\Omega^{*}g = 0 \tag{2.65}$$

$$d + \phi^{*'}c + g\omega + g\Omega^{*}c = 0 \tag{2.66}$$

$$f + c'\omega^{*} + \frac{1}{2}c'\Omega^{*}c = 0 \tag{2.67}$$

From proposition 2 and 3, all the parameters in our model can be expressed by 6 free parameters $\Theta = \{\phi, \omega^*, \phi^*, \Omega^*, c, g\}$.

$$A = \frac{1}{2}(I - \Omega^{*-1}) \tag{2.68}$$

$$B = \frac{1}{2}\phi'\phi - \frac{1}{2}\phi^{*'}\Omega^{*-1}\phi^* \tag{2.69}$$

$$C = \Omega^{*-1}\omega^* \tag{2.70}$$

$$D = -\phi^*\Omega^{-1}\omega^* \tag{2.71}$$

$$G = \Omega^{*-1}\phi^* - \phi \tag{2.72}$$

$$F = 0.5ln(|\Omega^{*-1}|) - 0.5\omega^{*'}\Omega^{*-1}\omega^* \tag{2.73}$$

$$b = -0.5g'\Omega^* g - g\phi^* \tag{2.74}$$

$$d = -\phi^* c - g\omega^* - g\Omega^* c \tag{2.75}$$

$$f = -0.5c'\Omega^* c - c'\omega^* \tag{2.76}$$

Proof of proposition 4

Assume that $E_t^Q[(\frac{S_T}{S_t})^\varphi] = exp(\alpha_t + \beta_t' + X_t'\gamma_t X_t)$, We know that Under risk neutral probability measure, $X_{t+1} = \omega^* + \phi^* X_t + \Delta^* \varepsilon_{t+1}^Q$.

$$\begin{aligned}
E_t^Q[(\frac{S_T}{S_t})^\varphi] &= E_t^Q\left\{(\frac{S_{t+1}}{S_t})^\varphi E_{t+1}^Q[(\frac{S_T}{S_{t+1}})^\varphi]\right\} \\
&= E_t^Q[\varphi(X_t'bX_t + c'X_{t+1} + d'X_t + X_t'gX_{t+1} + f) + \alpha_{t+1} + \beta_{t+1}'X_{t+1} \\
&\quad +X_{t+1}'\gamma_{t+1}X_{t+1}] \\
&= E_t^Q[(\varphi(c'\phi^* + d') + \beta_{t+1}'\phi^*)X_t + \varphi(f + c'\omega^* + \beta\omega^*) + \omega^{*\prime}\gamma\omega^* + \alpha_{t+1} \\
&\quad +(\varphi c' + \varphi X_t' g + \beta_{t+1}' + 2X_t'(\phi^*)'\gamma_{t+1})\Delta^*\varepsilon^Q + (\varepsilon^Q)'\Delta^{*'}\gamma_{t+1}\Delta^*\varepsilon^Q] \\
&\quad +X_t'(\varphi g\phi^* + \varphi b + (\phi^*)'\gamma_{t+1}\phi^*)X_t \\
&= \varphi f + \alpha_{t+1} + \frac{1}{2}log|I - 2\Delta^{*'}\gamma_{t+1}\Delta^*| \\
&\quad +\frac{1}{2}(\varphi c' + \beta_{t+1}')(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi c + \beta_{t+1}) \\
&\quad +[\varphi(c'\phi^* + d') \\
&\quad +\beta_{t+1}'\phi^* + (\varphi c' + \beta_{t+1}')(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi g' + 2\gamma_{t+1}'\phi^*)]X_t \\
&\quad +X_t'(\varphi b + \varphi g\phi^* + (\phi^*)'\gamma_{t+1}\phi^* + (\varphi g \\
&\quad +2\gamma_{t+1})(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi g + 2\gamma_{t+1})')X_t
\end{aligned}$$

Thus by definition

$$\begin{aligned}\alpha_t &= r_{t,t+1} + \varphi f + \alpha_{t+1} - \frac{1}{2} log|I - 2\Delta^{*'}\gamma_{t+1}\Delta^*| + \omega^{*'}\gamma_{t+1}\omega^* + \omega^*(\varphi c + \beta_{t+1}) \\ &+ \frac{1}{2}(\varphi c' + \beta'_{t+1} + 2\omega^{*'}\gamma_{t+1})(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi c + \beta_{t+1} + 2\gamma_{t+1}\omega^*)\end{aligned}$$

$$\begin{aligned}\beta_t &= \varphi(\phi^{*'}c + d + g\omega^*) + \phi^{*'}\beta_{t+1} + 2\phi^{*'}\gamma_{t+1}\omega^* \\ &+ (2\phi^{*'}\gamma_{t+1} + \varphi g)(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(\varphi c + \beta_{t+1} + 2\gamma_{t+1}\omega^*)\end{aligned}$$

$$\gamma_t = \varphi b + \varphi g\phi^* + \phi^{*'}\gamma_{t+1}\phi^* + \frac{1}{2}(2\phi^*\gamma_{t+1} + \varphi g)(I - 2\Delta^{*'}\gamma_{t+1}\Delta^*)^{-1}(2\phi^*\gamma_{t+1} + \varphi g)'$$

Because $E_T^Q[(\frac{S_T}{S_t})^{\varphi}] = 1$, we can always get $\alpha_T = 0$, $\beta_T = 0$ and $\gamma_T = 0$. By recursive substitution, given the initial value of $\alpha_T = 0$, $\beta_T = 0$ and $\gamma_T = 0$, we can get α_t, β_t and γ_t.

CHAPTER 3: CATEGORIZED IDIOSYNCRATIC RISK PREMIUM

3.1 Introduction

In finance literature, whether idiosyncratic risk is priced has received considerable attention. Classic portfolio theory suggests that representative investors will hold well-diversified portfolios, in which idiosyncratic risk will be diversified away, thus there is no compensation for holding idiosyncratic risk. However, in reality not all investors hold well diversified portfolios. Especially for retail investors, their portfolios are extremely less diversified on average. Statman (1987) finds that a portfolio need to contain at least 30 stocks to well diversify idiosyncratic volatility. By studying a sample of more than 62,000 household investors from 1991 to 1996, Goetzmann and Kumar (2004) find that more than 25 percent of the investor portfolios contain only one stock; over half of the investor portfolios contain no more than three stocks; and only ten percent of the investor portfolios contain more than ten stocks. When investors hold

Some other theories imply that idiosyncratic risk should be positively priced when investors hold under-diversified portfolios. Merton (1987) suggests that in an information-segmented world, firms with larger idiosyncratic volatility require higher average return to compensate investors for holding under-diversified portfolios.

Based on different representative investors, these two theories have different answers to the question that whether idiosyncratic risk is priced. In reality, different investors are not evenly distributed in the market and are concentrated in different groups of stocks. In behavioral finance, investors are sorted into two groups, retail investors and institutional investors. These two groups often trade stocks in different ways and have different appetite for stocks. Kumar and Lee (2006) find that retail investors are highly concentrated in small-cap, low-priced and value stocks. Brandt, Brav, Graham and Kumar (2010) find that retail investors have a strong preference for low-priced stocks. Kumar (2009) and Han and Kumar (2010) find that retail investors prefer stocks with lottery features such as positive skewness, low price and high idiosyncratic volatility. In contrast, Gompers and metrick (2001) document that institutional investors increases demand for large firms and decrease demand for small firms from 1980 to 1996. Brandt, Brav, Graham and Kumar (2010) also find that institutional investors are concentrated in large, growth, high-priced stocks. As these two groups are heavily concentrated in different stocks, we expect that representative investors of these stocks are different.

However, in regardless of the representative investors, empirical studies often treat stocks with different representative investors the same and test the question for all stocks simultaneously. The empirical results so far are mixed. Malkiel and Xu (2002); Spiegel and Wang (2006); Chua, Goh, and Zhang (2010) and Fu (2009) all find that idiosyncratic volatility risk premium is positive and significant. Ang, Hodrick, Xing and Zhang (2006, 2009) find a negative relationship between low realized idiosyncratic volatility in the previous month and the portfolio return in the subsequent month. When the short-term return reverse is controlled, Huang, Liu, Rhee and Zhang (2009) and Fu (2009) both find that the negative relationship in Ang, Hodrick, Xing and Zhang (2006) disappears. Bali and Cakici (2008) find that idiosyncratic volatility risk premium is sensitive to the frequency of data and the approach to estimate conditional expected idiosyncratic volatility. Using equal weighted portfolio return instead of value weighted portfolio return, they find an insignificant relationship between equal-weighted portfolio return and realized idiosyncratic volatility in the previous month.

If the two theories are both correct, we expect that stocks with different representative investors may have different idiosyncratic volatility risk premium. It is highly possible that we will be misguided by tests based on the full sample. For example, there are two different groups of stocks in the market. One group has a positive risk premium while the other group has a negative risk premium. Tests based on the full sample may support an insignificant risk premium, which is far beyond the truth.

In this chapter, we attempt to fill the gap between theories and empirical studies on idiosyncratic risk by answering the following questions: (1) whether idiosyncratic volatility risk premium is different for stocks with different representative investors; (2) What is the relationship between idiosyncratic volatility risk premium and representative investors?

Many literatures show that stock price is an accurate proxy for retail ownership, we use stock historical moving average price as a proxy for retail ownership in this chapter. In Brandt, Brav and Graham and Kumar (2010), they find that low-priced stocks have less than 5.25 percent institution ownership on average while the average institution ownership of high-priced stocks is 56.44 percent in their sample. Green and Hwang (2009) show that stocks undergo splits experience an increase in co-movements with low-priced stocks and a decrease in their co-movements with high-priced stocks. They imply that investors categorize stocks based on price. Several reasons may explain why retail investors prefer low-priced stocks and institutional investors dislike low-priced stocks. For example, the return of low-priced stocks has lottery features, such as high idiosyncratic volatility and positive skewness (See Kumar(2009a), Kumar(2009b) and Bali, Cakici and Whitelaw (2011)); retail investors may overweight the low probability of extreme return (Barberis and Huang 2008); institutional investors avoid trading low-priced stocks due to illiquidity, high transaction or prudence reasons (See Lakonishok, shleifer, and Vishny (1992), Del Guercio (1996) and Brave and Heaton (1997)).

For each individual stock, we calculate its historical moving average price between month $t-n$ to $t-1$ as its persistent price level. All stocks are sorted into 5 groups by stock historical moving average price. We use a price dummy D_{price} to represent the five price groups. D_{price} is from 1 to 5, representing the lowest-priced group to the highest-priced group respectively.

For robustness, we use three different measures to estimate conditional expected idiosyncratic volatility. For the first measure, we assume that the idiosyncratic volatility is a martingale and use realized monthly idiosyncratic volatility in month t to predict idiosyncratic volatility in month $t+1$. For the second measure, we do exponential smoothing for idiosyncratic volatility and estimate conditional expected idiosyncratic volatility using Riskmetric variance model. For the third measure, following Fu (2009), we use EGARCH (1,1) to predict conditional expected idiosyncratic volatility for each individual stock based on monthly return in previous months.

To test the relationship between idiosyncratic volatility risk premium and price level, we reply on three different approaches.

Firstly, we add an interaction term between price dummy and conditional expected idiosyncratic volatility to the cross-sectional regressions and test the significance of the coefficient estimate for the full sample. It turns out that the coefficient estimate of the interaction term is negative and significant at 5% level for all three measures. This indicates that low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks. In other words, the more retail investors are concentrated in a stock, the higher idiosyncratic volatility the stock will have. However, the sign and scale of idiosyncratic volatility risk premium is sensitive to the measure of conditional expected idiosyncratic volatility. For the first two measures, only the lowest-price quintile has a positive idiosyncratic volatility risk premium while the other quintiles have negative idiosyncratic volatility risk premiums. When idiosyncratic volatility is measured by EGARCH (1,1) model, all the price quintiles have positive risk premiums.

Secondly, we run cross-sectional regressions for subsamples. In the first test, we run cross-sectional regressions for the full sample. However, this test mechanically adds two constraints: (1) other risk premiums are cross-sectional invariant (2) and the difference in idiosyncratic volatility risk premium between adjacent groups is the same. To reduce the concern on the two constraints, We can run cross-sectional regressions for each price group. Subsample tests still show that low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks, which is robust to all the three measures. Besides that, for all three measures, lowest-priced stocks have significantly positive risk premiums while highest-priced stocks have insignificant risk premium. Because lowest-priced stocks and highest-priced stocks are extremely held by retail investors and institutional investors respectively, this finding is consistent with classical portfolio theory and Merton (1987).

Thirdly, we seek evidence from time-series regressions to support that low-priced stocks have significantly higher idiosyncratic volatility risk premium than high-priced stocks. By comparing the returns between low-priced stocks and high-priced stocks, we uncover that a simple trading strategy, long lowest-priced stocks and short highest-priced stocks, can bring abnormal return which is unexplained by popular systematic factors. If high return of low-priced stocks can be explained by high idiosyncratic volatility and high idiosyncratic risk premium, we expect that the abnormal return will be largest in highest idiosyncratic volatility quintile. By comparing the returns between low-priced stocks and high-priced stocks within each idiosyncratic volatility quintile, we find that the abnormal return is still persistent and is largest in the highest idiosyncratic volatility quintile. All these findings suggest that low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks.

In summary, we show that idiosyncratic volatility risk premium is cross-sectional variant, especially low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks. Evidence from subsample suggests that lowest-priced stocks have a positive idiosyncratic volatility risk premium while highest-priced stocks have an insignificant idiosyncratic volatility. As retail investors and institutional investors are heavily concentrated in lowest-priced stocks and highest-priced stocks, this finding supports the theoretic predictions in Merton (1987) and classical portfolio theory simultaneously. It also implies that the mixture findings in literatures are driven by stocks with different representative investors.

The remainder of the chapter is organized as follows. In Section 2, we discuss the concept of idiosyncratic volatility and propose three different approaches to estimate conditional expected idiosyncratic volatility. In section 3, we discuss the cross-sectional relationship among expected return, conditional expected idiosyncratic volatility and historical moving average price which is a proxy for retail ownership. In section 4, we use three different approaches to empirically examine whether idiosyncratic volatility risk premium is cross-sectional variant. Section 5 concludes.

3.2 Idiosyncratic Volatility

In finance, idiosyncratic risk is defined as firm-specific risk. In contrast with systematic risk, idiosyncratic risk can be diversified away in a well-diversified portfolio. In this chapter, we assume that Fama-French three factors can capture all the systematic risk and the risk unexplained by Fama-French three factor model is idiosyncratic risk. Following standard literature, we use idiosyncratic volatility as a proxy for idiosyncratic risk. Idiosyncratic volatility is defined as the standard deviation of residuals from the Fama-French three factor model. To estimate realized idiosyncratic volatility of an individual stock, we regress its daily excess returns on daily Fama-French three factors every month by equation (1).

$$R_{i,d,t} = \alpha_{i,t} + \beta_{i,t,MKT} MKT_{d,t} + \beta_{i,t,SMB} SMB_{d,t} + \beta_{i,t,HML} HML_{d,t} + \varepsilon^{i}_{d,t}, \tag{3.1}$$

where $R^{i}_{i,d,t}$ is stock i's excess return. $MKT_{d,t}$ represents market excess return on day d in month t. $SMB_{d,t}$ denotes the size factor, which is the difference between the return on the portfolio of small stocks and the return on the portfolio of large stocks (SMB). $HML_{d,t}$ represents the book to market factor, which is the difference between the return on the portfolio of high book-to-market stocks and the return on the portfolio of low book-to-market stocks.

Daily stock returns are obtained from the Center for Research in Security Prices (CRSP). Our data include all common stocks traded in NYSE, NASDAQ and AMEX from January 1965 to Dec 2011. The daily information of Fama-French three factors is collected from

Kenneth R. French's Website.

Time-series regressions for each individual stock are implemented with a monthly rolling window. To avoid noisy estimations or over-fitting the data, we require that each individual firm in a month has a minimum of 15 trading days. Daily idiosyncratic volatility of individual stock i in month t is the standard deviation of $\varepsilon^i_{d,t}$. Monthly idiosyncratic volatility is adjusted as the square root of the number of trading days multiplied by the daily idiosyncratic volatility in that month.

The core principle in finance is risk required reward. Investors require higher expected return for bearing risk. In theory, the risk and return tradeoff should be contemporaneous. However neither expected return or expected risk can be observed directly in reality. In practice, a conventional way is to use realized return as expected return and use different model to estimate expected risk. In this chapter, we need to estimate conditional expected idiosyncratic volatility to test whether idiosyncratic volatility risk premium is sensitive to representative investors. We construct three different estimates based on realized idiosyncratic volatility in the previous months.

3.2.1 Estimating Idiosyncratic Volatility under Martingale Assumption

We assume that idiosyncratic volatility series follows a martingale. This assumption implies that stock $i's$ realized idiosyncratic volatility in month t is its conditional expected idiosyncratic volatility in month $t+1$ based on the information in month t. From now on, we denote this measure as IV1.

3.2.2 Estimating Idiosyncratic Volatility by Riskmetrics

Riskmetrics variance model, also known as an exponential smoother, is widely used to predict volatility in industry. This approach assigns different weights to past variances and predicts future variance as a weighted average of past variances by equation (2) and (3).

$$\sigma_{i,t}^2 = \sum_{j=1}^{n} w_{t-j}\sigma_{i,t-j}^2, \tag{3.2}$$

where

$$w_{t-j} = \frac{\lambda_i^j}{\sum_{j=1}^{n} \lambda_i^j}, \tag{3.3}$$

where $\sigma_{i,t}^2$ is stock $i's$ variance in month t; w_{t-j} is the weight of stock $i's$ variance in month $t-i$ and λ_i^j is the smoothing parameter for individual stock i. This approach is very simple to be implemented. For each individual stocks, only one smoothing parameter λ_i^j is required. For simplicity and also to reduce the concern of over-fitting, we use a common smoothing parameter for all stocks. It is well known that the first order correlation of variance is around 0.9. We set the smoothing parameter λ as 0.9 and use realized idiosyncratic variances in the past 12 months (n=12) to estimate conditional expected idiosyncratic volatility. The results are robust to different value of smoothing parameter and different length (n) of historical variance. [1]

Estimating Idiosyncratic Volatility by EGARCH (1,1) Model

In the first two measures, we predict monthly idiosyncratic volatility from daily data. For the third measure, we use EGARCH models to estimate the conditional expected idiosyncratic volatilities based on monthly returns. Compared with classic GARCH models, EGARCH model can capture the leverage effect that negative residual will have larger effect

[1] We use different λ (0.8, 0.85, 0.95) and n (n=3, 6 and 24) to estimate conditional idiosyncratic volatility. The results are similar for different estimations of conditional idiosyncratic volatility.

on the volatility than positive residual. Pagan and Schwert (1990) fit a number of different models to monthly U.S. stock returns and find that the EGARCH model is the best over-all. Fu (2009) uses EGARCH(p,q) model to estimate the conditional idiosyncratic volatility for individual stock i, where $p \leq 3$ and $q \leq 3$.

$$r_t^i = \alpha^i + \beta_{MKT}^i MKT_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \varepsilon_{i,t}^i,$$
$$\varepsilon_{it} \in N(0, \sigma_{it}^2),$$
$$ln\sigma_{it}^2 = \alpha_i + \sum_{l=1}^{p} b_{i,l} ln\sigma_{i,t-1}^2 + \sum_{k=1}^{q} c_{i,k} \left\{ \theta(\frac{\varepsilon_{i,t-k}}{\sigma_{i,t-k}}) + \gamma[|\frac{\varepsilon_{i,t-k}}{\sigma_{i,t-k}}| - (\frac{2}{\pi})^{\frac{1}{2}}] \right\} \tag{3.4}$$

where r_t^i is stock $i's$ monthly excess return, MKT_t, SMB_t, HML_t are the monthly Fama-French three factors. For simplicity, we use EGARCH(1,1) model for all stocks. Conditional expected idiosyncratic volatility for each individual stock is estimated by a monthly expanding window, which requires a minimum of 30 observations. All conditional expected idiosyncratic volatility is estimated by past information. In other words, conditional idiosyncratic volatility in month t is forecasted by monthly returns up to month $t-1$ for each individual stock.

3.3 Idiosyncratic Volatility and Expected Returns: Cross-Sectional Regressions

3.3.1 Idiosyncratic Volatility Risk Premium is Cross-sectional Invariant

Following Fama and MacBeth (1973), we attempt to test whether time-series average of the coefficient estimates of idiosyncratic volatility is significantly different from zero on a monthly basis. We run the following cross-sectional regressions each month for all stocks:

$$R_{i,t+1} = \beta_{0,t} + \beta_t X_{i,t} + \gamma_t E_t \sigma_{t+1} + \varepsilon_{i,t+1}, \tag{3.5}$$

where $R_{i,t+1}$ is stock $i's$ realized excess return in month $t+1$; $X_{i,t}$ is stock $i's$ firm-specific characteristics (etc., beta, size, book to market ratio) in month t. $E_t\sigma_{t+1}$ is stock $i's$ conditional expected idiosyncratic volatility in month $t+1$, $\varepsilon_{i,t+1}$ is the residual unexplained by firm-specific characteristics. γ_t is the idiosyncratic volatility risk premium in month t.

3.3.2 Idiosyncratic Volatility Risk Premium is Cross-sectional Variant

Modern financial economic theory assumes that representative investors are rational in two aspects, (1) they make optimal decisions based on the same axioms in expected utility theory (2) and they have the same expectation of future return. However, these two assumptions are too optimal and invalid in reality. In contrast, based on the irrefutable assumptions-investor sentiment and limitation of arbitrage, behavioral finance investigates whether irrational investors affect stock price. Literatures in this area find that retail investors prefer certain group of stocks, trade stocks together and affect stock price by creating a systematic risk. For example, Kumar and Lee (2006), Kumar (2009), Brandt, Brav, Graham and Kumar (2010) all find that retail investors are concentrated in certain group of stocks (i.e., small-cap, low-priced, value, positively skewed stocks). Feng and Seasholes (2004), Jackson (2003) and Barber, Odean, and Zhu (2003) find that trades of retail investors are positive correlated in different stock markets and in different country; Kumar and Lee (2006) find that systematic retail trading explains return co-movements for stocks with high retail concentration; Barber, Odean and Zhu (2008) and Han and Kumar (2010) both document that retail trading importantly affect stock prices using different data.

Shed lights by all these findings, in this chapter we conjecture that idiosyncratic volatility risk premium is clientele-based. The intuition is very simple. As retail investors and institutional investors concentrated in different group of stocks, these stocks will have different representative investors. By theory, retail investors require a positive idiosyncratic risk premium to compensate for holding under-diversified portfolios while institutional investors require nothing. When stocks have different representative investors and different representative investors require different compensation for holding idiosyncratic risk, it is naturally to guess that idiosyncratic volatility risk premium should be clientele-based. In this chapter, we try to test whether stocks with different representative investors have different idiosyncratic volatility risk premium.

To answer this question, we assume that representative investors of stocks with highest retail ownership are retail investors and representative investors of stocks with lowest retail ownership are institutional investors. 13(f) institutional holding data from Thomson Reuters is widely used as a measure of institutional ownership. However, we do not use this data in this chapter for several reasons. Firstly, the data only starts from 1980. To compare with other chapters, we need to an extensive data starting from 1960s. Secondly, the data is only quarterly updated. Volatility models are extensively explored for daily volatility or monthly volatility. It is not clear for us whether these models still work for quarterly volatility. Thirdly, only large institutions are required to report their holdings to SEC. A 1978 amendment to the Securities and Exchange Act of 1934 required all institutions with greater than 100 million of securities under discretionary management to report their holdings to the SEC, while mid-size institutions or wealthy individuals are exempt from this act. However, in U.S., a large proportion of stocks are owned by wealthy individuals. It is not clear whether the 13(f) institution holding data is an accurate proxy for institution ownership.

Instead of using 13(f) data, we use stock's price level (historical moving average price) as a proxy for stock's retail ownership. Green and Huang (2009) suggest investors categorize stocks based on price by studying the co-movements between stocks undergo splits and low-priced stocks. Especially, retail investors are concentrated in low-priced stocks and institutional investors prefer high-priced stocks. Brandt, Brav, Graham and Kumar (2009) find that low-priced stocks have less than 5.25 percent institution ownership on average, while institution ownership of high-priced stocks is 56.44 percent on average in their sample. The significantly negative relationship between retail ownership and price level are also documented in other literatures (i.e., Kumar and Lee (2006), Kumar (2009), Han and Kumar (2010) and Fu (2009)). In the other hand, institutional investors often avoid trading Low-priced stocks for prudence reason and high transaction cost (See Lakonishok, shleifer and Vishny (1992), Del Guercio (1996) and Brave and Heaton (1997)).

For an individual stock, we calculate its moving average price as a proxy for its price level.

$$\overline{P_{i,t}} = \frac{1}{n}\sum_{1}^{n} P_{i,t-n} \tag{3.6}$$

where $\overline{P_{i,t}}$ is the moving average price for individual stock i in month t and n is the length of prices we used to compute the moving average price. All the results reported in this chapter is based on $n = 3$ (we test our results with different n (n=1, 6, and 12), we get similar results as we set $n = 3$).

To test whether idiosyncratic volatility is cross-sectional variant, we add an interaction term between stock's price level and conditional expected idiosyncratic volatility to the cross-sectional regressions.

$$R_{i,t+1} = \beta_t X_{i,t} + \gamma_t E_t \sigma_{t+1} + \phi_t D_{price} E_t \sigma_{t+1} + \varepsilon_{i,t+1}, \tag{3.7}$$

D_{price} is a price dummy, a proxy to capture the price level of an individual stock. In every month, we sort all stocks into 5 groups by their historical moving average prices and assign 1 to 5 to D_{price} from lowest price level to highest price level. ϕ_t. the coefficient of the intersection term describes the difference in idiosyncratic volatility risk premium between stocks with different price level. For example, the idiosyncratic volatility risk premium of lowest-priced stocks is $\gamma_t + \phi_t$, while the idiosyncratic volatility risk premium of highest-priced stocks is $\gamma_t + 5\phi_t$. We test whether the time-series average coefficient estimate of the interaction term is significantly different from zero.

3.4 Empirical Results

3.4.1 Data and Variables

Our data include daily and monthly returns of all common stocks traded in NYSE, NASDAQ and AMEX from Jan 1964 to Dec 2011. Trading data are from the Center for Research in Security Prices (CRSP) and the book value of individual stocks are from Compustat. We use SP 500 index return as the market return and the one-month Treasury bill rate as the risk-free rate.

Following standard literature, we use firm's beta, size, book to market ratio and return in previous month to describe firm's characteristics. Firm's size and book to market ratio are adjusted by logarithm ($X_{i,t}$=[$Beta_{i,t}$, $Ln(Size)_{i,t}$, $ln(BE/ME)_{i,t}$, $Ret_{i,t-1}$]).

$Ln(Size)_{i,t}$ denotes the logarithm of firm i's capital capitalization in month t, which is calculated as the end-price in month $t-1$ multiplied by the number of stocks outstanding.

Because book value of a stock is updated yearly, Book to market ratio is calculated as the ratio between fiscal year-end book value of common equity and the calendar year-end market value. $Ret_{i,t-1}$ is stock i's return in month $t-1$ and is used to control for short-term return reversal.

For each individual stock, $Beta_{i,t}$ captures its systematic risk in month t. To reduce the possible effect of correlation between size and beta, we follow Fama-French (1992) and assign portfolio beta to individual stocks. In each month, stocks are sorted into ten groups by market capitalization. The cutoff is only determined by NYSE-listed stocks. Within each size group, we sort stocks into 10 portfolios by pre-ranking $beta$s. The beta of each individual stock is estimated from market model using the previous 24 to 60 months of returns, as available. We then regress the value-weighted return of each beta-size portfolio on the market return and the market return in previous month. The beta of each portfolio is the sum of the coefficient estimates of current market return and prior market return. At last, we assign the portfolio beta to each individual stock based on its size and beta rankings in each month.

To make sure all the information are available before they are used to explain the cross-section of stock returns. Following Fama-French (1992), we use size in month t-1 to explain the return in month t and use book to market ratio of fiscal year t to explain the returns for the months from Jul y in year t+1 to June in year t+2.

3.4.2 Main Results from Cross-Sectional Regressions

In this chapter, we run two different cross-sectional regressions for all stocks by equation (7) and (9), which can both be written as:

$$R_{i,t} = \gamma_{0,t} + \sum_{k=1}^{K} \gamma_{k,t} X_{k,i,t} + \varepsilon_{i,t}, i = 1,2,...,N_t, t = 1,2,...,T, \tag{3.8}$$

where $R_{i,t}$ is the realized excess return of stock i in month t. $X_{k,i,t}$ are the explanatory variables of cross-sectional expected returns. N_t denotes the number of stocks in month t, which may vary in each month. T is the total number of months in our data. We test whether the time-series average coefficients of conditional expected idiosyncratic volatility and the interaction term between price dummy and idiosyncratic volatility are significant different from zero. The final estimates of $\hat{\gamma}_k$ and its variance are:

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{kt} \tag{3.9}$$

$$var(\hat{\gamma}_k) = \frac{\sum_{t=1}^{T} (\hat{\gamma}_k - \hat{\gamma}_{kt})^2}{T(T-1)} \tag{3.10}$$

The t-statistic is the time-series mean of coefficient, $\hat{\gamma}_k$ divided by its standard deviation.

Cross-Sectional Regressions on the Full Sample

In this chapter, we use three different approaches to estimate conditional expected idiosyncratic volatility. Results based on the IV1, IV2 and IV3 are displayed in Panel A, B and C of Table 1 respectively. Model 1 is the typical model used to test whether idiosyncratic volatility risk premium is significantly different from zero on average. Using different measure of conditional expected idiosyncratic volatility risk premium, we find similar results as documented in literature.

Ang et all (2006) imply a significantly negative idiosyncratic volatility risk premium when it is estimated by IV1. This implies that lower idiosyncratic volatility in previous month predicts a higher return in the subsequent month. We also find that the idiosyncratic volatility risk premium is negative, -0.032 and is significant at 5 % level. When conditional expected idiosyncratic volatility is estimated by IV2, idiosyncratic risk premium is negative but not significant at 5% level. Using ARIMA(1,1) to estimate conditional expected idiosyncratic volatility, Huang, Liu, Rhee and Zhang (2009) also find similar result. Fu (2009) uses EGARCH (p,q) model to estimate conditional expected idiosyncratic volatility for each individual stock. He finds a significantly positive idiosyncratic volatility risk premium. Using a similar model to estimate idiosyncratic volatility, we find that idiosyncratic volatility risk premium is equal to 0.118 and is significant at 5% level.

Model (2) and Model (3) in Table 1 test whether idiosyncratic volatility risk premium is the same for stocks with different price levels. For all three measures, the coefficient estimates on idiosyncratic volatility are all become significantly positive, while the coefficient estimates on the interaction term are all negative and significant at 5% level. A negative coefficient estimate of the interaction term implies that low-priced stocks have a higher idiosyncratic volatility risk premium than high-priced stocks. If price level is a good proxy for retail ownership, this indicates the more retail investors are concentrated in a stock, the higher idiosyncratic volatility risk premium the stock has. The difference in idiosyncratic volatility risk premium for stocks with different price levels is large. When conditional expected idiosyncratic volatility is measured by IV1, the risk premium of lowest-priced s-

Table 3.1: Fama-Macbeth regressions on the full sample

This table reports the average coefficients in the Fama-MacBeth cross-sectional regressions for all common stocks in NYSE/AMEX/NASDAQ over the period from Jan 1964 to December 2011. IV1 is the realized idiosyncratic volatility in the previous month. IV2 is estimated by Riskmetrics based on realized idiosyncratic volatility in the previous 12 months. IV3 is estimated by EGARCH(1,1) model based on monthly return over the previous 24 to 60 months. Beta is estimated using the 10×10 size/beta double sorted portfolios following Fama and French (1992). Size and B/M are estimated as in Fama and French (1992) and adjusted by logarithm. Ret_{t-1} is stock return in previous month. D_{Price} is the price dummy, ranging from 1 (lowest-priced) to 5(highest-priced) based on stock historical moving average price in previous 3 months. $\bar{R}^2$ is the average cross-sectional adjusted R^2.

Models	Intercept	Beta	Size	B/M	Ret_{t-1}	IV	IV$\times D_{Price}$	$\bar{R}^2$
Panel A:	IV1: Realized Expected Idiosyncratic Volatility							
1	0.017	0.001	-0.001	0.002		-0.032		0.069
	(6.539)	(0.265)	(-3.897)	(3.695)		(-5.622)		
2	0.016	0.001	-0.001	0.001		0.014	-0.025	0.072
	(6.345)	(0.331)	(-2.116)	(2.886)		(2.058)	(-17.056)	
3	0.014	0.002	-0.000	0.001	-0.063	0.028	-0.025	0.079
	(5.415)	(0.069)	(-1.078)	(3.706)	(-16.947)	(4.028)	(-17.161)	
Panel B:	IV2: Expected Idiosyncratic Volatility by Riskmetrics							
1	0.03	-0.000	-0.001	0.002		-0.007		0.071
	(5.086)	(-0.086)	(-3.006)	(4.329)		(-0.590)		
2	0.013	0.000	-0.000	0.001		0.040	-0.029	0.074
	(5.211)	(0.074)	(-0.764)	(3.263)		(3.234)	(-16.088)	
3	0.013	0.000	-0.000	0.002	-0.065	0.041	-0.029	0.082
	(5.164)	(0.008)	(-0.512)	(3.804)	(-18.548)	(3.318)	(-16.540)	
Panel B:	IV3: Expected Idiosyncratic Volatility by EGARCH(1,1) Model							
1	-0.005	-0.004	0.001	0.004		0.116		0.074
	(-1.989)	(-1.768)	(2.668)	(8.313)		(11.498)		
2	-0.009	-0.004	0.002	0.003		0.189	-0.032	0.077
	(-3.358)	(-1.868)	(6.156)	(7.524)		(15.158)	(-16.700)	
3	-0.008	-0.004	-0.002	0.004	-0.064	0.188	-0.033	0.086
	(-3.076)	(-1.801)	(6.232)	(7.816)	(-18.923)	(15.488)	(-17.392)	

Table 3.2: Time-series average weight of stocks with different price level

Price Portfolio	1(lowest)	2	3	4	5(highest)
	1	2	3	4	5
mean	0.07	0.09	0.13	0.23	0.48
min	0.00	0.01	0.03	0.09	0.21
max	0.24	0.19	0.24	0.34	0.86

tocks is 0.03 while the risk premium of highest-priced stocks is -0.097; when conditional expected idiosyncratic volatility is measured by IV2, the risk premium of lowest-priced s-tocks is 0.012 while the risk premium of highest-priced stocks is -0.096. When conditional expected idiosyncratic volatility is measured by EGARCH(1,1) model, the risk premium of lowest-priced stocks is 0.155 while the risk premium of highest-priced stocks is 0.023.

In this section, we display the estimations of cross-sectional regressions for the full sample and find that idiosyncratic volatility risk premium is cross-sectional variant. Especially, low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks for all three measures. However, it is still puzzling that idiosyncratic volatility is negative for high-priced stocks when conditional expected idiosyncratic volatility is estimated by IV1 or IV2.

Cross-Sectional Regressions on the Sub-samples

In the last section, we add an interaction term to control the difference in idiosyncratic volatility between stocks within different price groups and run cross-sectional regressions for the full sample. However, this test is based on two important assumptions: (1) except for idiosyncratic volatility risk premium, other risk premiums are similar across stocks within different groups; (2) the difference in idiosyncratic volatility risk premium for stocks within

Table 3.3: Fama-Macbeth regressions on the subsamples

This table reports the average coefficients in the Fama-MacBeth cross-sectional regressions for all common stocks in NYSE/AMEX/NASDAQ over the period from Jan 1964 to December 2011. IV1 is the realized idiosyncratic volatility in the previous month. IV2 is estimated by Riskmetrics based on realized idiosyncratic volatility in the previous 12 months. IV3 is estimated by EGARCH(1,1) model based on monthly return over the previous 24 to 60 months. Beta is estimated using the 10×10 size/beta double sorted portfolios following Fama and French (1992) and are adjusted by logarithm. Size and B/M are estimated as in Fama and French (1992). Ret_{t-1} is stock return in previous month. Price portfolios 1 from (lowest-priced) to 5(highest-priced) are created by stock historical moving average price in previous 3 months. $\bar{R}^2$ is the average cross-sectional adjusted R^2.

Price Portfolios	Intercept	Size	B/M	IV	$\bar{R}^2$
Panel A:	IV1: Realized Expected Idiosyncratic Volatility				
1(lowest)	-0.042	0.017	0.007	0.040	0.045
	(-11.309)	(18.876)	(11.064)	(5.132)	
2	-0.035	0.011	0.005	0.014	0.046
	(-11.244)	(17.529)	(9.110)	(1.667)	
3	-0.024	0.007	0.004	-0.010	0.047
	(-8.476)	(15.653)	(6.559)	(-0.951)	
4	-0.023	0.006	0.004	-0.014	0.049
	(-8.463)	(15.324)	(6.445)	(-1.332)	
5(highest)	-0.026	0.005	0.005	-0.010	0.055
	(-9.599)	(14.853)	(7.106)	(-0.819)	
Panel B:	IV2: Expected Idiosyncratic Volatility by Riskmetrics				
1(lowest)	-0.072	0.020	0.008	0.145	0.049
	(-16.835)	(20.104)	(12.677)	(9.768)	
2	-0.054	0.012	0.007	0.103	0.053
	(-16.892)	(20.008)	(11.770)	(6.288)	
3	-0.036	0.008	0.005	0.059	0.054
	(-12.890)	(17.822)	(8.652)	(3.294)	
4	-0.031	0.006	0.005	0.037	0.057
	(-11.280)	(17.625)	(8.146)	(1.836)	
5(highest)	-0.030	0.005	0.005	0.020	0.064
	(-10.343)	(15.971)	(8.579)	(0.979)	
Panel C:	IV3: Expected Idiosyncratic Volatility by EGARCH(1,1) model				
1(lowest)	-0.086	0.021	0.012	0.271	0.065
	(-24.392)	(24.326)	(18.475)	(15.781)	
2	-0.064	0.013	0.008	0.189	0.058
	(-21.855)	(22.632)	(14.503)	(11.457)	
3	-0.046	0.009	0.006	0.131	0.056
	(-18.378)	(20.992)	(10.753)	(8.019)	
4	-0.039	0.007	0.006	0.095	0.057
	(-15.677)	(19.974)	(9.403)	(5.957)	
5(highest)	-0.038	0.006	0.006	0.067	0.062
	(-13.823)	(17.642)	(9.659)	(1.116)	

adjacent groups is the same.

To mitigate our concern on the two assumptions, we run cross-sectional regressions for stocks within each price group by model (1). The subsample tests allow stocks within different price group have different risk premiums and more importantly risk premiums for each price group can be estimated independently.

Table 3 yields striking evidence to support two theories in finance literature. Merton (1986) predicts that retail investors require a positive idiosyncratic volatility risk premium to compensate for holding under-diversified portfolios. If representative investors of certain stocks are retail investors, these stocks will have a positive idiosyncratic volatility risk premium. For all three different measures, lowest-priced stocks have a positive idiosyncratic volatility risk premium and it is significant at 5% level. To be addressed, lowest-priced stocks hold mainly by retail investors. This indicates that when retail investors are the representative investors of low-priced stocks, they require a positive idiosyncratic volatility, which is the same as predicted in Merton (1986). Classical portfolios theory shows that, when representative investors hold well diversified portfolios, idiosyncratic risk is diversified away and should not priced. In Table3, for all three different measures, highest-priced stocks have an insignificant idiosyncratic volatility risk premium. This implies that as representative investors of highest-priced stocks are institutional investors, these stocks do not have any idiosyncratic volatility risk premium.

In sum, results from cross-sectional regressions on the full sample and sub-samples both show that idiosyncratic volatility risk premium is cross-sectional variant, especially low-priced stocks earn a higher idiosyncratic volatility risk premium than high-priced stocks on average. Besides that, we get striking evidence from sub-sample tests to support that lowest-priced stocks earn a positive idiosyncratic volatility risk premium while highest-priced stocks earn an insignificant idiosyncratic volatility risk premium. If price is an accurate proxy for retail ownership, this indicates that retail investors as representative investors of lowest-priced stocks require a positive idiosyncratic volatility risk premium and institutional investors as representative investors of highest-priced stocks require nothing for idiosyncratic risk. This finding is consistent with the theoretic predictions in Merton (1987) and classical portfolio theory.

3.4.3 Main Results from Portfolio Analysis

Delong at all (1990) show that noise traders can earn a higher expected return than rational investors do for bearing a disproportionate amount of risk created by themselves. If price level is an accurate proxy for retail ownership, we can easily compare returns between low-priced stocks and high-priced stocks to test whether retail investors can earn a higher return than institutional investors on average.

In each month, we sort all stocks into five portfolios by historical moving average price and calculate value-weighted return for each portfolio. The second column in table 4 shows the simple value-weighted excess return of each portfolio, all the numbers in column 2 are in percentage and the number in bracket denotes the corresponding t statistics. The average returns from quintile 1 to 5 are 1.66%, 0.98%, 0.74%, 0.50% and 0.02% respectively. This indicates that Low-priced stocks earn a robust higher return than high-priced stock on average every month.

We also calculate risk-adjusted abnormal returns for all quintiles by different models. Column 3 to 5 in Table 4 reports CAPM-alpha, FF-alpha and Carhart-alpha for all quintiles. We can observe that low-priced portfolios have larger alpha than high-priced portfolios.

Table 3.4: Portfolios sorted by historical moving average price

We form value-weighted quintile portfolios every month by sorting stocks based on historical moving average price. Portfolios are formed every month, based on historical moving average price computed using monthly return in the previous 3 months. Portfolio 1(5) is the portfolio of stocks with the lowest (highest) moving average price. The statistics in the column labeled average return are measured in monthly percentage terms and apply to excess -returns, t statistics are displayed in the corresponding brackets. The row ``1-5" refers to the difference in monthly returns between portfolio 1 and portfolio 5. The Alpha columns report Jensen's alpha with respect to the CAPM, Fama-French (1993) three-factor model or Cahart (1997) four-factor model. Robust New-West (1987) t statistics are reported in the brackets. The sample period is January 1964 to December 2011.

Rank	Average Return	CAPM-Alpha	FF3-Alpha	Carhart-Alpha
1 (lowest)	1.66	1.16	0.85	1.18
	(4.57)	(5.13)	(5.11)	(6.09)
2	0.98	0.50	0.18	0.39
	(3.21)	(3.18)	(2.15)	(4.33)
3	0.74	0.29	0.00	0.12
	(2.74)	(2.27)	(0.00)	(2.13)
4	0.50	0.06	-0.15	-0.10
	(2.05)	(0.65)	(-3.15)	(-1.93)
5 (highest)	0.02	-0.42	-0.53	-0.54
	(0.08)	(-5.86)	(-13.36)	(-13.46)
1-5	1.64***	1.58***	1.37***	1.72***
	(7.84)	(8.04)	(7.85)	(8.56)

This pattern is robust to different models. We also compare the returns between low-priced stocks and high-priced stocks when size, book to market ratio, short-term reversal, liquidity or momentum effect is controlled. All these common factors cannot explain the abnormal return earned by low-priced stocks. All the tables of robustness check are available upon request.

To test whether low-priced quintiles have significant higher returns than high-priced quintiles, we can build up a portfolio by long lowest-priced stocks and short highest-priced stocks and test whether its alpha is significant from zero. The last row in Table 4 reports the performance of this portfolio formed on price quintiles. The 1-5 differences in average return, CAPM-alpha, FF-alpha and Carhart-alpha are 1.64%, 1.58%, 1.37% and 1.1.72% respectively and all the differences are significant at 5% level.

All the evidence in Table 4 supports that low-priced stocks earn higher return and higher risk adjusted return than high-priced on average and we can get positive abnormal return by long quintile 1 and short quintile 5. Together with the empirical findings that retail investors are mainly concentrated in low-priced stocks and institutional investors are concentrated in high-priced stocks, all above implies that retail investors can earn higher return than institutional investors on average.

As the difference in return between low-priced stocks and high-priced stocks cannot be explained by common systematic risk, we conjecture that the difference in return can be explained by idiosyncratic risk. Our story is very intuitive. As retail investors are concentrated in low-priced stocks, they can affect stock price by two ways. Firstly, trades of retail investors can increase idiosyncratic volatility. By studying a reform of the French stock market, Foucault et all (2011) find that retail trading activity has a positive effect on the volatility of stock return. This may explain why low-priced stocks have higher idiosyncratic volatility than high-priced stocks on average, which is widely documented in literatures (See Kumar and Lee (2006) and Brant et All (2009)). Secondly, because retail investors are concentrated in low-priced stocks while institutional investors are concentrated in high-priced stocks, by theory low-priced stocks should earn a higher idiosyncratic volatility risk premium than high price stocks. In other words, high return earned by low-priced stocks may be explained by high idiosyncratic volatility and high idiosyncratic volatility risk premium simultaneously.

In this chapter, our core goal is to prove that idiosyncratic volatility risk premium is cross-sectional variant. To find evidence from portfolio analysis, we can check whether low-priced stocks still earn higher returns than high-priced stocks when idiosyncratic volatility is controlled.

To control for idiosyncratic risk, every month we firstly sort all stocks into 5 groups by the three measures of conditional expected idiosyncratic volatility. Then within each group, we sort stocks into 5 groups by historical moving average price. Within each idiosyncratic volatility quintile, quintile 5 contains the highest-priced stocks while quintile 1 contains the lowest-priced stocks. We expect that high idiosyncratic volatility quintiles have larger difference in return between low-priced stocks and high-priced stocks than low idiosyncratic volatility quintiles, because the difference is enlarged by high-idiosyncratic volatility.

Table 5 reports Carhart-alpha of the 25 idiosyncratic volatility-price portfolios when conditional expected idiosyncratic volatility is measured by IV1. Within each idiosyncratic volatility quintile, low price quintile has a larger Carhart-alpha than high price quintile and the 1-5 Carhart-alpha are negative and significant at 5% level. This indicates that the abnormal return of low-priced stocks cannot explain by the story that low-priced stocks have high idiosyncratic volatility alone. As what we expected, the highest idiosyncratic volatility quintile has the largest 1-5 difference in Carhart-alpha 3.51%, while the lowest idiosyncratic volatility has the smallest 1-5 difference in Carhart-alpha 0.3%.

When idiosyncratic volatility is measured by IV2, We have similar findings in Table 6 as if idiosyncratic volatility is measured by IV1. The highest idiosyncratic volatility quintile has the largest 1-5 differences in Carhart-alpha, -3.66% and significant at 5% level, while the lowest idiosyncratic volatility quintile has the smallest 5-1 differences in Carhart-alpha, 0.27% and is also significant at 5% level. The phenomenon that low-priced stocks earn higher return than high-priced stocks is persistent in each idiosyncratic volatility quintile.

When idiosyncratic volatility is measured by IV3, in Table 7, the highest idiosyncratic volatility quintile has a significantly positive 1-5 difference in Carhart-alpha. This indicates that the anomaly that low-priced stocks earn higher return than high-priced stocks are mainly driven by stocks with high idiosyncratic volatility[2].

To sum up, in this section, we compare returns between low-priced stocks and high-priced stocks on average when idiosyncratic volatility is controlled. We find that low-priced stocks still earn higher return than high-priced stocks even when idiosyncratic volatility is controlled. Especially, the highest idiosyncratic volatility quintile has the largest 1-5 difference in Carhart-alpha while the lowest idiosyncratic volatility quintile has the smallest or insignificant 1-5 difference in Carhart-alpha. All the above is consistent with our story that low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks.

[2]Following Fu (2009), we estimate EGARCH (1,1) model for each individual stock every month. As attacked by Hui Guo and Michael Ferguson (2012), this approach may over-fit the data and estimations from which are too noisy.

Table 3.5: Alpha of price-idiosyncratic volatility portfolios

The table report Carhart (1997) alpha, with robust Newey-West (1987) t-statistics in the brackets. All the stocks are sorted into 5×5 Idiosyncratic volatility-Price portfolios. We firstly sort all stocks into five quintiles based on conditional expected idiosyncratic volatility measured by IV1 every month. IV1 is the realized idiosyncratic volatility relative to the FF-3 model in the previous month. Within each idiosyncratic volatility quintile, we sort all stocks into 5 price quintiles based on historical moving average price computed using monthly return in the previous 3 months. In the column ``IVOl'', Portfolio 1(5) is the portfolio of stocks with the lowest (highest) conditional idiosyncratic volatility. The column ``1-5'' refers to the difference in Carhart-alpha between portfolio 1 and portfolio 5 in each idiosyncratic volatility quintile. The sample period is January 1964 to December 2011. All the results are based on monthly value-weighted portfolio returns.

	Price Portfolio					
VOL portfolio	1	2	3	4	5	1-5
	Fama-French-Carhart Alpha					
1	0.35	0.22	0.15	0.11	0.06	0.30
	(4.66)	(3.20)	(2.19)	(1.55)	(0.78)	(3.73)
2	0.51	0.32	0.23	0.21	0.06	0.45
	(5.65)	(4.71)	(3.22)	(2.85)	(0.88)	(4.40)
3	0.73	0.29	0.30	0.27	0.16	0.58
	(6.49)	(3.85)	(4.03)	(3.35)	(2.38)	(4.46)
4	0.96	0.33	0.29	0.25	0.04	0.92
	(6.13)	(2.77)	(3.01)	(3.20)	(0.47)	(5.39)
5	3.39	0.36	0.20	0.06	-0.12	3.51
	(10.25)	(1.72)	(1.14)	(0.37)	(-0.75)	(12.43)

Table 3.6: Alpha of price-idiosyncratic volatility portfolios

The table reports Carhart (1997) alpha, with robust Newey-West (1987) t-statistics in the brackets. All the stocks are sorted into 5×5 Idiosyncratic volatility-Price portfolios. We firstly sort all stocks into five quintiles based on conditional expected idiosyncratic volatility measured by IV2 every month. IV2 is estimated from realized idiosyncratic volatility relative to the FF-3 model in the previous twelve months by Riskmetrics approach. Within each idiosyncratic volatility quintile, we sort all stocks into 5 price quintiles based on historical moving average price computed using monthly return in the previous 3 months. In the column ``IVOl", Portfolio 1(5) is the portfolio of stocks with the lowest (highest) conditional idiosyncratic volatility. The column ``1-5" refers to the difference in Carhart-alpha between portfolio 1 and portfolio 5 in each idiosyncratic volatility quintile. The sample period is January 1964 to December 2011. All the results are based on monthly value-weighted portfolio returns.

	Price Portfolio					
IVOL portfolio	1	2	3	4	5	1-5
	Fama-French-Carhart Alpha					
1	0.37	0.23	0.20	0.07	0.10	0.27
	(5.32)	(3.45)	(2.93)	(0.94)	(1.45)	(4.07)
2	0.43	0.22	0.16	0.12	0.05	0.38
	(5.14)	(3.11)	(2.24)	(1.60)	(0.68)	(4.44)
3	0.46	0.21	0.19	0.16	0.10	0.36
	(4.41)	(2.48)	(2.49)	(2.15)	(1.31)	(2.99)
4	0.69	0.17	0.15	0.25	0.04	0.65
	(4.64)	(1.46)	(1.50)	(2.74)	(0.37)	(3.64)
5	3.72	0.50	0.58	0.49	0.06	3.66
	(10.90)	(2.37)	(3.05)	(2.73)	(0.30)	(12.55)

Table 3.7: Alpha of price-idiosyncratic volatility portfolios

The table reports Carhart (1997) alpha, with robust Newey-West (1987) t-statistics in the brackets. All the stocks are sorted into 5×5 Idiosyncratic volatility-Price portfolios. We firstly sort all stocks into five quintiles based on conditional expected idiosyncratic volatility measured by IV3 every month. IV3 is estimated from monthly return in the previous 24 to 60 months by EGARCH (1,1) model every month. Within each idiosyncratic volatility quintile, we sort all stocks into 5 price quintile based on historical moving average price computed using monthly return in the previous 3 months. In the column ``IVOl", Portfolio 1(5) is the portfolio of stocks with the lowest (highest) conditional idiosyncratic volatility. The column ``1-5" refers to the difference in Carhart-alpha between portfolio 1 and portfolio 5 in each idiosyncratic volatility quintile. The sample period is January 1964 to December 2011. All the results are based on monthly value-weighted portfolio returns.

	Price Portfolio					
IVOL portfolio	1	2	3	4	5	1-5
	Fama-French-Cahart Alpha					
1	-0.02	0.06	0.06	-0.04	-0.09	0.07
	(-0.22)	(0.82)	(0.84)	(-0.60)	(-1.36)	(0.97)
2	0.05	0.02	0.02	0.03	0.02	0.03
	(0.65)	(0.23)	(0.21)	(0.41)	(0.31)	(0.36)
3	-0.01	-0.03	0.04	0.09	0.07	-0.08
	(-0.13)	(-0.35)	(0.50)	(1.30)	(1.17)	(-0.77)
4	0.22	-0.05	0.06	0.15	0.01	0.21
	(1.62)	(-0.45)	(0.67)	(1.91)	(0.12)	(1.30)
5	4.62	1.34	1.28	1.07	0.76	3.86
	(12.17)	(5.38)	(5.91)	(5.92)	(4.38)	(12.00)

3.5 Conclusion

Financial theories imply that idiosyncratic risk is priced differently by different representative investors. Merton (1987) implies that when representative investors hold under-diversified portfolios, idiosyncratic volatility is positively priced. In contrast, classical portfolio theory suggests that idiosyncratic risk should not be priced when representative investors hold well-diversified portfolios.

In reality, different stocks may have different representative investors. In finance, investors are often sorted into two groups-retail investors and institutional investors. Empirical studies find that these two groups have significantly different appetite for stocks and hold different portfolios. Retail investors prefer low-priced, small-cap, value stocks and stocks with lottery features and often hold under-diversified portfolios, while institutional investors prefer high-priced, large-cap and growth stocks and hold well-diversified portfolios.

In this chapter, we attempt to link idiosyncratic volatility risk premium with the type of representative investors and test whether idiosyncratic risk is priced differently in stocks with different representative investors. Shed lights from Brandt et all (2009) and other literatures, we use stock price level as a proxy for retail ownership and attempt to answer this question by three different approaches.

Firstly, we implement cross-sectional regressions for the full sample and test whether the coefficient estimate of an interaction term between price dummy and idiosyncratic volatility is significant from zero. For different measures of idiosyncratic volatility, the coefficient is negative and significant at 5% level. This indicates that low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks.

Secondly, we run cross-sectional regressions for subsamples. It turns out that lowest-priced stocks have the largest idiosyncratic volatility risk premium while highest-priced stocks have an insignificant idiosyncratic volatility risk premium. If representative investors of lowest-priced stocks and highest-priced stocks are retail investors and institutional investors respectively, these findings are consistent with theoretic predictions in Merton (1987) and classical portfolio theory.

Thirdly, we examine the returns of a set of portfolios that are sorted by idiosyncratic volatility and price level. We uncover a very robust result that low-priced stocks have high abnormal return. Especially the abnormal return is highest in the highest idiosyncratic volatility quintile and is the lowest or insignificant in the lowest idiosyncratic volatility quintile. This implies that, together with high idiosyncratic volatility, high idiosyncratic volatility risk premium can explain the high abnormal return of low-priced stocks.

In sum, we find evidence to support that idiosyncratic volatility risk premium is cross-sectional variant. Specifically speaking, low-priced stocks have higher idiosyncratic volatility risk premium than high-priced stocks. If price level is an accurate proxy for retail ownership, stocks with higher retail ownership has higher idiosyncratic volatility risk premium than stocks with lower retail ownership. Besides that, we discover striking evidence from subsample tests to support Merton (1987) and classical portfolio theory simultaneously. As stocks with different representative investors have different idiosyncratic volatility risk premiums, it requires us to carefully interpret the empirical findings of idiosyncratic volatility risk premium in and re-examine the way to test idiosyncratic volatility risk premium in literatures.

BIBLIOGRAPHY

Ahn, D., Dittmar, R.F. and Gallant A. R., 2002. Quadratic Term Structure Models: Theory and Evidence. *The Review of Financial Studies* 15(2), 243-288.

Ait-Sahalia, Y. and Lo, A. W., 1998. Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *The Journal of Finance* 53(2), 499-547.

Ang, A., G., Hodrisk, Y., Xing and X., Zhang, 2006. The Cross-section of Volatility and Expected Returns. *Journal of Finance* 61, 259-299.

Ang, A., G., Hodrisk, Y., Xing and X., Zhang, 2009. High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence. *Journal of Financial Economic* 91, 1-23.

Ang, A. and Kristensen, D., 2010. Testing Conditional Factor Models. *Working Paper*.

Avellaneda, Levy et all, 1995. Pricing And Hedging Derivative Securities In Markets With Uncertain Volatilities. *Applied Mathematical Finance* 2, 73-88.

Avellaneda, Levy et all, 1996. Managing The Volatility Risk Of Portfolios Of Derivative Securities: The Lagrangian Uncertain Volatility Model. *Applied Mathematical Finance* 3, 21-52.

Bakshi, G. and N. Kapadia, 2003. Delta-Hedged Gains and the Negative Market Volatility Risk Premium. *Review of Financial Studies* 16(2), 527-566.

Bakshi, G. and N. Kapadia, 2009. Idiosyncratic Risk and the Cross-section of Expected Stock Returns. *Journal of Financial Economics* 91, 24-37.

Bakshi, G., Madan, D. and G. Panayotov, 2010. Return of Claims on the Upside and the Viability of U-shaped Pricing Kernels. *Journal of Financial Economics* 97, 130-154.

Bali, T. G., 2008. The intertemporal Relationship between Expected Return and Risk. *Journal of Financial Economics* 87, 101-131.

Bali, T. G. and R. F. Engle, 2010. The intertemporal capital asset pricing model with dynamic conditional correlations. *Journal of Monetary Economics* 57(4), 377-390.

Bansal, R. and S. Viswanathan, 1993. No Arbitrage and Arbitrage Pricing: A New Approach. *The Journal of Fiance* 48(4), 1231-1262.

Barber, B.M., Odean, T. and N. Zhu, 2009. Do Retail Trades Move Markets. *Riview of Financial Studies* 22, 151-186.

Barone-Adesi, G., Engle, R. F. and L. Mancini, 2008. A GARCH Option Pricing Model with Filtered Historical Simulation. *Review of Financial Studies* 21(3), 1223-1258.

Bernardo, A. and O. Ledoit, 2000. Gain, Loss, and Asset Pricing. *Journal of Political Economy* 108, 144-172.

Bikbov, R. and M. Chernov, 2009. Unspanned Stochastic Volatility in Affine Model: Evidence from Eurodollar Futures and Options. *Management Science* 55, 1292-1305.

Black, F., 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics* 3, 167-179.

Bliss, R. R. and N. Panigirtzoglou, 2004. Option-Implied Risk Aversion Estimates. *Journal of Finance* 59(1), 407-446.

Bollerslev, T. and H. Zhou, 2009. Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22(11), 4463-4492.

Bondarenko, O., 2003. Statistical Arbitrage and Securities Prices. *Review of Financial Studies* 16, 875-919.

Boyle, P., Feng, S., Tian, W. and T. Wang, 2008. Robust Stochastic Discount Factors. *Review of Financial Studies* 21, 1077-1122.

Brandt, M. W., Brave, A., Graham, J. R. and A. Kumar, 2010. The Idioyncratic Volatility Puzzle: Time Trend or Speculative Episodes? *Riview of Financial Studies* 23, 863-899.

Brennan, M. J., Wang, A. W., and Y. Xia, 2004. Estimation and Test of a Simple Model of Intertemporal Capital Asset Pricing, *The Journal of Finance* 59(4), 1743-1776.

Broadie, M., Chernov, M. and M. Johannes, 2007. Model Specifications and Risk Premia: Evidence from Futures Options. *Journal of Finance* 62, 1453-1490.

Campbell, John Y., 1993. Intertemporal Asset Pricng without Consumption Data. *American Economic Review* 83, 487-512.

Campbell, John Y., 1996. Understanding Risk and Return. *Journal of Political Economy* 52, 57-82.

Carhart, Mark M., 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 104, 298-345.

Carr, P. and D. B. Madan, 2001. Pricing and Hedging in Incomplete Markets. *Journal of Financial Economics* 62, 131-167.

Carr, P. and L. Wu, 2004. Time-changed Levy Processes and Option Pricing. *Journal of*

Financial Economics 71 (3): 113-141.

Carr, P., and L. Wu, 2009. Variance Risk Premiums. *Review of Financial Studies* 22 (3): 1311-1341.

Casasus, J., and P. Collin-Dufresne, 2005. Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates. *Journal of Finance* 60, 2283-331.

Chabi-Yo, F., Garcia, R., and E., Renault, 209. State Dependence Can Explain the Risk Aversion Puzzle. *Review of Financial Studies* 21 (2), 973-1011.

Cheredito, P., D. Filipovic., and R. Kimmel, 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics* 83, 123-170.

Christoffersen, P., S. Heston, and K. Jacobs, 2012. Option Anomalies and the Pricing Kernel. *Working paper*.

Cochrane, J., and J. Saa-Requejo, 2000. Beyond-Arbitrage: Good Deal Asset Prices Bounds in Incomplete Markets. *Journal of Political Economy* 108, 79-119.

Collin-Dufresne, P., and R. Goldstein, 2002. Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility. *Journal of Finance* 57, 1685-1730.

Cont, R., 2006. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance* 16, 519-547.

Dai, Q. and K. Singleton, 2003. Term Structure Dynamics in Theory and Reality. *Review of Financial Studies* 16(3), 631-678.

Deaton, A., and G. Laroque, 1996. On the Behavior of Commodity Prices. *Review of Economic Studies* 69, 1-23.

Delong, J. B., Shleifer, A., Summer, L. H. and R. J. Waldmann, 1990. Noise Trader Risk in Financial Markets. *Journal of Political Economy* 98, 703-738.

Dittmar, R. F., 2002. Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross section of Equity Return. *The Journal of Finance* 57(1), 369-403.

Doran, James S. and E. Ronn, 2008. Computing The Market Price of Volatility Risk in The Energy Commodity Markets. *Journal of Banking and Finance* 32, 2541-2552.

Duffie, D., Pan, J. and K. Singleton, 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68, 1343-1376.

Fama, E.F. and K.R. French, 1992. The Cross-section of Expected Stock Returns. *Journal* *of Finance* 47, 427-465.

Fama, E.F. and K.R. French, 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 25, 23-49.

Fama, E., F. and K., R., French. 1996. The CAPM is Wanted, Dear or Alive. *The Journal of Finance* 51(5), 1947-1958.

Fama, E.F. and J.D. MacBeth, 1973. Risk Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 71, 607-636.

Foucault, T., Sraer, D. and D.J. Thesmar, 2011. Individual Investors and Volatility. *Journal of Finance* 66, 1369-1406.

Fu, E., 1992. The Cross-section of Expected Stock Returns. *Journal of Finance* 47, 427-465.

Garlappi, L., Uppal, R. and T. Wang, 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20, 41-81.

Ghysels, E., Santa-Clara P. and R. Valkanov, 2005. There is a Risk-return Trade-off After All. *Journal of Financial Economics* 76, 509-548.

Gibson, R. and E. S. Schwartz, 1990. Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *Journal of Finance* 45, 959-976.

Goetzmann, W., Ingersoll, J. and M. Spiegel, 2007. Portfolio Performance Manipulation and Manipulation-proof Performance Measures. *Review of Financial Studies* 20(5), 1503-1546.

Gompers, P. and A. Metrick, 2003. Institutional Investors and Equity Prices. *Quarterly Journal of Economics* 116, 229-259.

Goyal, A. and P. Santa-Clara, 2003. Idiosyncratic Risk Matters!. *Journal of Finance* 58, 975-1007.

Green, T.C. and B. Huang, 2000. Price-Based Return Comovement. *Journal of Finance* 93, 37-50.

Grinblatt, M. and M. Keloarju, 2000. The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set. *Journal of Financial Economics* 55, 43-67.

Guo, H. and F. Whitelaw, 2006. Unconvering the Risk-return Relation in the Stock Market. *Journal of Finance* 61, 1433-1464.

Hansen, L. and R. Jagannathan. 1997. Assessing Specification Errors in Stochastic Discount Factors Models. *Journal of Finance* 52, 557-90.

Hansen, L. and T. Sargent, 2001. Robust Control and Model Uncertainty. *American Economic Review* 91, 60-66.

Harrison, J. M. and D. M. Kreps. 1979. Martingale and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* 30, 381-408.

Harvey, C. R. and A. Siddique, 2000. Conditional Skewness in Asset Pricing Tests. *Journal of Finance* 55, 1263-1295.

Hentschel, L., 1995. All in the Family Nesting Symmetric and Asymmetric GARCH models. *Journal of Financial Economics* 39, 71-104.

Heston, S., 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies* 6, 327-443.

Heston, S. and S., Nandi, 2000. A Closed-Form GARCH Option Valuation Model. *Review of Financial Studies* 13(3), 585-625.

Hong, H. 2000. A Model of Returns and Trading in Futures Markets. *Journal of Finance* 959-988.

Huang, W., Liu, Q., Rhee, S.G. and L. Zhang, 2010. Return Reversals, Idiosyncratic Risk and Expected Returns. *Review of Financial Studies* 23p2006, 147-168.

Hughen, W. K., 2009. A Maximal Affine Stochastic Volatility Model of Oil Prices. *Journal of Future Matkets*, 30, 101-133.

Kogan, L., Livdan, D. and A. Yaron, 2008. Oil Futures Prices in a Production Economy with Investment Constraints. *Journal of Finance* 64, 1345-1375.

Jagannathan, R. and Z. Wang., 1996. The Conditional CAPM and the Cross-Section of Expected Returns. *Journal of Finance* 51, 3-53.

Jackwerth , J. C., 2000. Recovering Risk Aversion from Option Prices and Realized Returns. *Review of Financial Studies* 13(2), 433-451.

Jankunas, A., 2001. Optimal Contingent Claims. *Annals of Applied Probability* 11, 735-749.

Kaniel, R., Saar, G. and S. Titman, 2008. Individual Investor Trading and Stock Returns. *Journal of Finance* 63, 273-310.

Kumar, A. and C. Lee, 2006. Retail Investor Sentiment and Retail Comovements. *Journal* *of Finance* 61, 2451-2486.

Kumar, A., 2009. Who Games in the Stock Market. *Journal of Finance* 64, 1889-1933.

Kumar, A., 2009. Dynamic Style Preferences of Individual Investors and Stock Returns. *Journal of Financial and Quantitative Analysis* 44, 607-640.

Leippold, M. and L. Wu. 2002. Asset Pricing Under The Quadratic Class. *Journal of Financial and Quantitative Analysis* 37(2), 271-295.

Li, H. and F. Zhao. 2006. Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives. *Journal of Finance* 61, 341-378.

Liew, J. and M. Vassalou. 2000. Can Book-to-market, Size and Momentum be Risk Factors that Predict Economic Growth. *Journal of Financial Economics* 57, 221-245.

Lo, A. and M. Mueller, 2010. WARNING: Physics Envy May Be Hazardous To Your Wealth!. *Journal of Investment Management* 8, 13-63.

Maio, P. and P. Santa-Clara, 2012. Multifactor Models and Their Consistency with ICAPM. *Journal of Financial Economics*, Forthcoming.

Merton, R.C., 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41, 867-887.

Merton, R.C., 1987. Presidential Address: A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance* 42, 483-510.

Miltersen, K. and E.S. Schwartz. 1998. Pricing of Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates. *Journal of Financial and Quantitative Analysis* 33, 33-59.

Nagel, S. and K.J. Singleton, 2011. Estimation and Evaluation of Conditional Asset Pricing Models. *Journal of Finance* 66(3), 873-909.

Newest, W.K. and K.D. West, 1987. A Simple Positive-definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

Pontiff, J., 2006. Costly Arbitrage and the Myth of Idiosyncratic Risk. *Journal of Accounting and Economics* 42, 35-52.

Petkova, R., 2006. Do the Fama-French Factors Proxy for Innovations in Predictive Variable?. *Journal of Finance* 61(2), 581-612.

Petkova, R. and L. Zhang, 2005. Is Value Riskier than Growth?. *Journal of Financial Economics* 78, 187-202.

Rosenberg, J. V. and R. F. Engle, 2002. Empirical Pricing kernels. *Journal of Financial Economics* 64, 341-372.

Routledge, B. R., Seppi, D. J. and C. S. Spatt, 2000. Equilibrium Forward Curves for Commodities. *Journal of Finance* 55, 1297-338.

Romano, M. and N. Touzi, 1997. Contingent Claims and Market Completeness in a Stochastic Volatility Model. *Mathematical Finance* 7, 399-412.

Samuelson, P, 1965. Proof that properly anticipated price fluctuate randomly. *Industrial Management Review* 6, 41-50.

Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance* 19, 425-442.

Schwartz, E. S., 1997. The Stochastic Behavior of Commodity Prices: Implications for valuation and Hedging. *Journal of Finance* 52, 923-73.

Schweizer, M., 1995. Variance-Optimal Hedging in Discrete Time. *Mathematics of Operations Research* 20, 1-32.

Schweizer, M., 1996. Approximation Pricing and the Variance-Optimal Martingale Measure. *Annals of Probability* 24, 206-236.

Stambaugh, R. F., Yu, J. and Y. Yuan, Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *Working Paper*.

Tian, W., Wang, T. and H. Yan, 2010. Correlation Uncertainty and Equilibrium Asset Prices. *Working paper*.

Trolle, A. B. and E. S. Schwartz, 2009. Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. *Review of Financial Studies* 22, 4423-4461.

Xu, Y. and B. G. Malkiel, 1996. The Conditional CAPM and the Cross-Section of Expected Returns. *Journal of Business* 51 (1996), 3-53.