Essays on Stochastic Volatility and Jumps

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities.

2012

Ke Chen

Manchester Business School

Contents

Word count: 27904

List of Tables

List of Figures

The University of Manchester

Ke Chen

Doctor of Philosophy

Essays on Stochastic Volatility and Jumps

December 2012

ABSTRACT

This thesis studies a few different finance topics on the application and modelling of jump and stochastic volatility process. First, the thesis proposed a non-parametric method to estimate the impact of jump dependence, which is important for portfolio selection problem. Comparing with existing literature, the new approach requires much less restricted assumption on the jump process, and estimation results suggest that the economical significance of jumps is largely mis-estimated in portfolio optimization problem. Second, this thesis investigates the time varying variance risk premium, in a framework of stochastic volatility with stochastic jump intensity. The proposed model considers jump intensity as an extra factor which is driven by realized jumps, in addition to a stochastic volatility model. The results provide strong evidence of multiple factors in the market and show how they drive the variance risk premium. Thirdly, the thesis uses the proposed models to price options on equity and VIX consistently. Based on calibrated model parameters, the thesis shows how to calculate the unconditional correlation of VIX future between different maturities.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgment

First, I would like to take this opportunity to thank my supervisor, Professor Ser-Huang Poon, for her constructive and helpful support. I should also thank Dr. Simon Acomb for his generous help in my PhD works. And I owe a debt of gratitude to Professor Michael Brennan for pointing out the key mistakes in the early stage of my works. Last but not least, I should thank my wife for her encouragement and support for my study.

Chapter 1

Introduction

This thesis concerns jumps and stochastic volatility in the stock price process. Jumps in equity market prices are widely documented in the press and current literature. It is well recognised that jumps have a significant impact on portfolio management, risk premium, and asset pricing. The fundamental difference between jumps and diffusion processes makes many mathematical results and estimation procedures, developed for the diffusion process, obsolete. For example, Kalman filter can not be applied when there are jumps; asset dependence is much more complicated to classify for jump process than pure diffusion process, and jump dependence can not be modelled in any convenient way as the number of assets increases. Very often, strong restrictions and assumptions are imposed in the model in order to quantify the jump dependence. Unfortunately, such restrictions and assumptions might result in serious mis-specification of the model and bias in the estimation, which will lead to inappropriate financial decisions and huge financial loss. Therefore, appropriate mathematical approaches must be chosen carefully to fully appreciate the nature of the problem caused by jumps.

Another important feature in equity market is the time varying volatility. Stochastic volatility models are one of the most popular ways to handle such feature. Compared with discrete time models, such as GARCH, stochastic volatility models, such as Heston and the so-called 3/2 models, are mathematically more convenient for derivative pricing and could have semi-closed form solution. But many stochastic volatility models have been criticized for the failure to track the underlying dynamics, and poor calibration results with respect to the whole implied volatility surface. Due to the difficulty in fitting the whole surface, some studies allow model parameters

to change for different maturities and strike prices. Although such method can help to achieve perfect fit to the market data, it leads to inconsistent pricing and it is no more than an ad hoc interpolation method; the calibrated model does not reflect the true dynamics of the underlying process. The use of a volatility model with a wrong underlying price dynamics to price derivatives is dangerous, and even more so for hedging. The other way to achieve a better fit to market data without sacrificing the close form solution is to add jumps, e.g., compound Poisson jumps, to the stochastic volatility process, or as an extra stochastic factor in the model. However, many argue if such extra complexity is justified.

Given the challenges listed above, this thesis aims to make a contribution to the modelling of jump and stochastic volatility processes in portfolio selection, estimation of variance risk premium, and derivative pricing and hedging.

First of all, I propose a non-parametric framework to handle jump processes across international equity markets in portfolio selection. In portfolio optimisation with the presence of jumps, some researches only consider systemic co-jumps due to the difficulty of modelling jump dependence, while other researches only take into account independent idiosyncratic jumps, which only affect one particular asset. My new approach allows both systemic and idiosyncratic jumps, and captures the jump dependence in a non-parametric way. The impact of jumps and their dependence on optimal portfolio weights is easy to estimate in this setup.

Second, I study the time varying risk premium embedded in variance swaps, especially that due to a realized jump. I propose a stochastic volatility model with an extra factor in the form of self exciting jump to capture the impact of a realized jump on the risk premium. The results show a strong support for a second factor for short term maturity options and variance swap. The estimation results show a good fit to the volatility surface dynamics, even during the sub-prime crisis, and is able to extract the time varying risk premium due to stochastic volatility and jump components.

Third, I show how to price options on SPX and VIX consistently using the two-factor models. The calibration results are shown to be satisfactory. Furthermore, I prove that single factor models imply perfect correlation among VIX futures, contradicting market observation. Multi-factor models are needed to produce more realistic VIX future correlation term structure. Finally, I show that multi-factor models which produce equally good fit to the two volatility surfaces could lead to a different VIX future correlation term structure, which has important implication for hedging.

In Chapter 2, I provide a concise and comprehensive review on the current literature on the three aspects of jumps and stochastic volatility addressed in this thesis. In the review, I start with the modelling of jump processes in portfolio selection. I then extend the review to modelling jumps with time varying jump intensity and time varying risk premium. Finally, I provide an overview of pricing methods on equity and volatility derivatives.

Chapter 3 proposes a general framework for taking jump into consideration in portfolio selection. In particular, by assuming that the information of jumps is fully contained in the univariate return series, the estimation can be separated into two parts; the estimation of jump process for each univariate time series and the estimation of jump dependence. I estimate the jump diffusion model for weekly MSCI index returns of eleven developed and emerging markets, and study the impact of jumps on portfolio weights. Comparing with the previous literature, I reach the same conclusion that, for the developed markets, the omission of systematic jumps has little impact on portfolio selection. However, for emerging markets, the loss in certainty equivalence due to the wrong jump dependence assumptions is economically significant.

Chapter 4 introduces the pricing methods for affine models used in Chapters 5 and 6. I use some well known affine models in finance as examples to show how the fundamental partial differential equation (PDE) for the characteristic function is derived and present the option price as an integral, and I compare two different schemes for calculating this integral numerically.

In Chapter 5, I introduce a self-exciting jump process to capture the market dynamics after a big realised negative jump, and investigate how risk premium responds afterward. The dynamics of stochastic variance and jump intensity, together with their premium, are estimated jointly from both stock return series and option prices. Instead of using option prices directly, I use cumulants constructed from option prices under the risk neutral measure. I use particle filter nested in an EM (Expectation Maximisation) framework for the estimation. The observed high skewness in the implied volatility surface during the sub-prime crisis suggests that time varying jump intensity plays a very important role in explaining the risk premium during the sub-prime crisis, while stochastic variance dominates the risk premium in other calm periods.

Chapter 6 shows how to price contingent claims on both equity and volatility derivatives consistently using the joint characteristic function of equity price and

state variables. Based on the linear approximation of jump size, I prove that one factor models imply perfect correlation across the entire term structure of VIX future, which is not the case empirically. I then demonstrate how to use multi-factor models to calculate the optimal hedge ratio for hedging VIX option with VIX future. I further derive the unconditional correlation term structure of VIX future implied by the model based on the stationary distribution of the state variables. By comparing the VIX future correlation implied by different models, I show that the two implied volatility surfaces are not sufficient for distinguishing the right multi-factor model for hedging VIX future; multi-factor models that provide equally good fit to the two volatility surfaces produce very different VIX future correlation term structure.

Finally, Chapter 7 concludes with a summary of the main contributions of this thesis.

Chapter 2

Literature Review

In this thesis, I model and estimate jumps and stochastic volatility processes for portfolio selection, risk premium estimation, and derivatives pricing and hedging. This chapter provides a concise and comprehensive review on the related literature. First, I start with Merton (1976), which is among the first to introduce the jump process in modelling underlying dynamics, and to examine the impact of jumps on portfolio selection. For the topic of time varying risk premium, a few important works in the literature, e.g., Pan (2002) and Eraker (2004) are discussed. In relation to my work in consistent pricing of equity and volatility derivatives, I present a short comparison of different approaches adopted in Sepp (2008), Bergomi (2007), and Psychoyios et al. (2010). For each part, I also identify the gaps in the existing literature and explain how this thesis contributes to the literature.

Since some advanced mathematical tools, the integral transform methods in particular, are important for my research, I also provide a short review on the technique and its applications in finance research.

2.1 Estimation of Jumps

In the topics studied in this thesis, the role of jumps is one of the most important subjects. Before moving onto the main content, I need to clarify what kind of jumps is concerned in my research. The characteristic of jumps is not self-explained and could be different depending on the context.

In mathematical sense, jumps, as opposed to diffusion processes, are used to describe the discontinuous movement. Based on this property, Barndorff-Nielsen and

Shephard (2006) propose a non-parametric method to isolate jumps by comparing the quadratic variation and the bi-power variation of the time series. For a return series, r_t, composed of a drift, a diffusion process and a jump process,

$$r_t = a_t + \sigma_t W_t + J_t \tag{2.1}$$

the quadratic variation and the bi-power variation from time t to $t + \Delta t$ are defined as,

$$\begin{aligned} \text{QV} &= \text{QV}^c + \text{QV}^J \\ &= \text{p} - \lim_{\delta \to 0} \sum_{j=1}^{\Delta t/\delta} r_j^2 \\ \text{BPV} &= \text{p} - \lim_{\delta \to 0} \sum_{j=2}^{\Delta t/\delta} |r_j| \, |r_{j-1}| \end{aligned}$$

where QV^c is the quadratic variation caused by the diffusion process, QV^J is the quadratic variation caused by jumps, and $\text{p} - \lim$ denotes the limit in probability. Barndorff-Nielsen and Shephard (2004) shows that the bi-power variation is independent from the discontinuous component and

$$\text{BPV} = \frac{2}{\pi} \text{QV}^c$$

Therefore, the contribution to the quadratic variation by the discontinuous component, i.e., jumps, is

$$\text{QV}^J = \text{QV} - \frac{\pi}{2} \text{BPV}$$

which measures the significance of jumps during the time period of t to $t + \Delta t$. Such method is very appealing as it is model and distribution independent and easy to implement. However, it is subject to one limitation that the above equations only hold asymptotically as the time interval between observations, $\delta \to 0$, which is only possible with high frequency data.

For low frequency, e.g., daily or weekly, observations, such condition makes it impossible to achieve a stable estimation. We have to use other implications in equation (2.1) to do the estimation. In general, the idea is to match the distribution of r_t implied by the model to that in the market. The estimation method could

be MLE, GMM, or other more sophisticated methods, which depend either on the likelihood or moments. Apparently, this approach is model dependent. For example, if we use different assumption on the volatility process σ_t, i.e., constant volatility, CIR process, or log-normal process, the implied distribution of jumps will also be affected, since the observed distribution of return series remains the same.

Furthermore, the estimation by following this idea could be under either physical measure or risk neutral measure. If it is under physical measure, the observation frequency could change the estimation results, since the data statistics change. While under risk neutral measure, jumps are used to provide better fit for the implied volatility surface (see Bates (2000) for example), which could change dramatically from day to day calibration.

It is not difficult to see that only when the proposed model is the true model that drives market movements, these estimations of jumps are consistent with each other. But it is impossible to achieve, since there is no such model (or no one knows it). The best we can do is to mimic the market under certain conditions by some proposed models which can explain the market feature we are interested in. Since I only work on low frequency data in this thesis, the estimation of jumps will be different under different topics, since model assumptions varies.

2.2 Portfolio Optimisation with Jumps

Merton (1976) introduces a jump process to model dynamics as follows,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + \sum_{i=1}^{N_t} Y_i \tag{2.2}$$

where μ and σ are the drift and volatility, respectively, W_t is a standard Wiener process, N_t is a Poisson process with jump intensity λ, and $Y_i \sim N(\theta, \delta^2)$ is the jump size.[1] Merton (1976) shows how to price European options with the proposed model. Aase (1984) is the first to solve the portfolio optimisation problem using this jump diffusion model, however, no jump dependence between assets is specified. In another words, the jumps are all idiosyncratic jumps. In such a setup, Aase (1984) shows

[1]There is another convention to present jump process by using random measure. The convention here follows that in Merton (1976).

that the optimisation problem for a portfolio consisting of N assets is to solve,

$$\begin{aligned} 0 \;=\; \max_{\omega} \Bigg\{ & \frac{\partial V(W_t,t)}{\partial t} + \frac{\partial V(W_t,t)}{\partial W_t} W_t \sum_{i=1}^{N} \omega_i \mu_i \\ & + \frac{1}{2} \frac{\partial^2 V(W_t,t)}{\partial W_t^2} W_t^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{i,j} \omega_i \omega_j \sigma_i \sigma_j \\ & + \sum_{i=1}^{N} \lambda_i \mathbb{E}\left[V(W_t + W_t \omega_i J_i, t) - V(W_t,t) \right] \Bigg\} \end{aligned} \tag{2.3}$$

where W_t is the wealth or value of the portfolio, ω_i is the investment weight for ith asset, $\rho_{i,j}$ is the correlation between the diffusion parts of the ith and jth assets, and $V(W_t,t)$ is defined as,

$$V(W_t,t) \equiv \max_{\omega} \mathbb{E}\left[U(W_T)\right]$$

and $U(W_T)$ is the utility function. Since there is no dependence between jumps in different assets, the impact of jumps on portfolio selection is the sum of each individual jump effect as shown in equation (2.3), i.e., the summation is outside of the expectation in the last term. Similar specification can be found in Jeanblanc-Picque and Pontier (1990).

In 70's and early 80's, such a framework worked fine. In the international markets, the big movements were often uncorrelated prior to 1990. However, this is no longer the case now. The Asian financial crisis in 1997, the Dot-com bubble in 2001, and the sub-prime crisis in 2008 are clear evidence that the stock markets can jump together and jumps can cause serious damage when they occur at the same time. To examine the impact of systemic co-jump, Das and Uppal (2004) use the Merton jump diffusion process in equation (2.2) to model the univariate stock price process, but assume that all the markets share the same jump timing N_t, and that the jump size Y for each market are perfectly correlated. Under these assumptions, Das and Uppal (2004)

show that the optimisation problem becomes

$$\begin{aligned} 0 \;=\; \max_{\omega} \Bigg\{ & \frac{\partial V(W_t,t)}{\partial t} + \frac{\partial V(W_t,t)}{\partial W_t} W_t \sum_{i=1}^{N} \omega_i \mu_i \\ & + \frac{1}{2} \frac{\partial^2 V(W_t,t)}{\partial W_t^2} W_t^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{i,j} \omega_i \omega_j \sigma_i \sigma_j \\ & + \lambda \mathbb{E}\Big[V(W_t + W_t \sum_{i=1}^{N} \omega_i J_i, t) - V(W_t,t) \Big] \Bigg\} \end{aligned} \tag{2.4}$$

Comparing equations (2.3) and (2.4), the only difference is the last term, which is due to the jump component. Unlike Aase (1984), the impact of jump is the effect of the sum of individual jumps as shown in equation (2.4), i.e., the summation is inside of the expectation in the last term. Das and Uppal (2004) show that systemic jump risk reduces the gains from diversification, and makes leveraged portfolios more susceptible to large losses. In a more recent paper, Aït-Sahalia et al. (2009) provide the close form solution of the optimal portfolio weights under the same restrictive co-jump assumptions as in Das and Uppal (2004).

The two sets of studies described above consider either systemic or idiosyncratic jumps, but not both. However, international markets are subject to both types of jumps, so the jump risk could be under- or over-estimated, and the optimal portfolio weight might be biased due to the mis-specification of jumps. The restrictive assumptions used in previous studies are most likely due to the difficulty in estimating jump dependence. In Chapter 3, I attempt to solve this problem by a non-parametric approach. My method permits both systemic and idiosyncratic jumps, and also works for the case when only a subset of the portfolio constituent markets jump together.

2.3 Variance Risk Premium

The market risk premium and the pricing of derivatives in the setting of constant volatility is well established, for example CAPM and Black-Scholes model. However, it is well recognised that volatility in the financial market is stochastic. Carr and Wu (2009) calculate the profit and loss produced from the difference between the realized variance with the expectation of variance under the risk neutral measure. They claim the statistically significant loss is a clear evidence of variance risk premium.

Among the studies of risk premium embedded in option prices, Guo (1998) examine the Heston model, Pan (2002) considers the model of Bates (2000), which has jumps in return series in addition to stochastic volatility, whereas Eraker (2004) further allows jumps in the stochastic volatility process. In the more recent literature, researchers started to add more factors as sources of stochastic volatility, which help to provide better fit to the implied volatility surface and the underlying dynamics. For example, Bergomi (2007) uses two factors with different timescale dynamics for the variance process in order to fit the VIX smile. By using high frequency data, Todorov (2010) isolates jumps and demonstrates that variance risk premium is exceptionally high after realized jumps. His result suggests there might be an extra risk factor associated with realized jumps contributing to the variance risk premium.

In most of the previous studies, jumps are modelled with a constant jump intensity. Since jumps contribute to the total variance, if the jump intensity is stochastic, the jump process will generate stochastic variance. An interesting research question concerns whether the jump intensity is driven by realized jumps. Such a jump process is called self exciting jump, first studied in Hawkes (1971) and seen application in research on crimes, disease spreading, and earthquakes. This idea of self-exciting jump is appealing in finance research, especially for its explanatory power for the extreme events, such as the sub-prime crisis in 2008. The first notable study of similar dynamics is the GARJI (Generalised Autoregressive Conditional Jump Intensity) model proposed by Maheu and McCurdy (2004), in which the jump intensity follows an autoregressive model and is driven by the likelihood of jump in the previous day. Specifically, the process for the jump intensity is,

$$\begin{aligned} \lambda_t &= \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1} \\ \xi_{t-1} &= \mathbb{E}[n_{t-1}|r_{t-1}] - \lambda_{t-1} \end{aligned} \tag{2.5}$$

where $\mathbb{E}[n_{t-1}|r_{t-1}]$ is the ex-post expectation of the number of jumps given the previous day return r_{t-1}, ξ_{t-1} is a measurement of the strength of the unexpected movement in the jump intensity, and ρ controls the speed of jump intensity decaying back to baseline level λ_0.[2] Although, the GARJI model captures the desired jump feature in the return series well, it is difficult to extend the discrete time specification for pricing options, and hence limits its potential for explaining the risk premium

[2]The notations here follow those in Maheu and McCurdy (2004), and are not necessary consistent with other parts of this thesis.

embedded in option prices. A continuous time version appears in Ait-Sahalia et al. (2010), but their focus is on modelling the return series only.

Option prices are often used to estimate the variance risk premium. Ever since the seminal work by Demeterfi et al. (1999), which provides a model free replication strategy for the expectation of variance based on European options, volatility derivatives, such as variance swap and VIX, have become increasingly popular in the last decade. So it is now common to use the market prices of vanilla options and volatility derivatives to estimate the risk premium (see Carr and Wu (2009) and Duan and Yeh (2011)).

In Chapter 5, I propose an affine model with a self-exciting jump feature. The expectations of short term variance and skewness of log returns under the risk neutral measure are used in the estimation since they are easy to derive given the model is affine. These two measures are approximations of the level and skewness of the ATM implied volatility at short term maturities. By tracking the stock return series and expectation of variance and skewness replicated from option prices, I estimate the model parameters together with the time varying variance risk premium and explain how investors respond in the stock crashes marked by big jumps as in the 2008 subprime crisis.

2.4 VIX Option Pricing

The Volatility Index (VIX) was first introduced by the Chicago Board Options Exchange (CBOE) in 1993. In 2003, CBOE adopted the model free approach by Demeterfi et al. (1999) to revised the VIX definition as follows,

$$\mathrm{VIX}_t = \sqrt{\frac{1}{T-t}\left(\int_0^{F_t} \frac{2}{K^2}\mathrm{Put}(K)dK + \int_{F_t}^{\infty} \frac{2}{K^2}\mathrm{Call}(K)dK\right)} \tag{2.6}$$

where K is the strike price, T is the time of maturity, and F_t is the forward price. The details are described in CBOE (2009). Since its first launch in 1993, VIX, as a measurement for the market volatility, was known as the 'fear' index. After the revision in 2003, futures and options written on VIX became available for the use in hedging and speculation. The trading volume of these derivatives on VIX has increased by 20 folds since 2006, and the pricing of VIX derivatives has become a new hot topic in finance.

In the literature, there are two different approaches for pricing VIX derivatives. One is to treat VIX as an independent security and model it as the underlying asset (see, for example, Grunbichler and Longstaff (1996), Bergomi (2007), and Psychoyios et al. (2010)). The advantage of this approach is the tractability, since the closed form solution for the specified dynamics is usually available. The other pricing approach is to specify the dynamics for the equity price process first, and then derive the VIX dynamics and the price of VIX derivatives (See Zhang and Zhu (2006) and Sepp (2008)). The latter approach becomes more popular in recent research, since such method can price derivatives on SPX and VIX consistently, albeit with extra complexity when compared with the former one.

However, the extra constrains introduced in the consistent pricing approach means it is harder to fit market data well. The implied volatility of SPX and VIX options observed in the market usually has negative and positive skewness respectively. For the popular choice of stochastic volatility model, such as Heston model, the CIR process does not generate enough positive skewness to match the VIX smile. There are a few ways to fix this problem. One way is to add jumps into the CIR process (See Sepp (2008)). To further improve the fitness of the model, Lu and Zhu (2009) use multi-factor model to fit the term structure of VIX futures.

In Chapter 6, I follow the consistent pricing approach and test two affine models with multi-factor to price SPX and VIX options. The main contribution of this work is the analysis of the dynamics of VIX future term structure. Specifically, I study the correlation between VIX futures of different maturities, and the exposure of VIX options to VIX futures.

2.5 Pricing with Characteristic Function

The price of a contingent claim can be written as the discounted expectation of future payoff under the risk neutral measure,

$$\begin{aligned} f(X_t,t) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[f(X_T,T)|X_t] \\ &= e^{-r(T-t)}\int f(X_T,T)p^{\mathbb{Q}}(X_T|X_t)dX_T \end{aligned} \tag{2.7}$$

where $p^{\mathbb{Q}}(X_T|X_t)$ is the transition probability and $f(X_T,T)$ is the final payoff. If this probability function exists in close form, the derivative price can be evaluated as

an integral, as in the case of the Black-Scholes model. For other more complicated models, the transition probability function might not have close form, but equivalent alternative, i.e, characteristic function of the transition probability might have. Then the pricing formula becomes,

$$f(X_t, t) = \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)\phi(-\omega, X_t, T)d\omega \tag{2.8}$$

where $\phi(\omega, X_t, T)$ and $\hat{f}(\omega)$ are the characteristic function of $p^{\mathbb{Q}}(X_T|X_t)$ and $f(X_T, T)$ respectively.

Heston (1993) is the first to show how to price the option based on the characteristic function. Bates (1996) extends Heston model to incorporate jumps in the underlying price. A general framework of derivative pricing with affine model is introduced by Duffie et al. (2000), who show how the characteristic function for an affine model is derived. The general approach is well presented in Sepp (2003).

Sometimes, close form of characteristic function also exists for other non-affine models. For example, Carr and Sun (2007) derive the characteristic function for the 3/2 model. However, the derivation can not be easily extended to other non-affine models. For models which do not admit close form for either transition probability or characteristic function, approximation can be made (see Fouque et al. (2011) and Josep Perelló and Masoliver (2008)).

Whether the solution is presented in equations (2.7) or (2.8), the option price is represented as an integral, which has to be solved numerically. Such an integral can be evaluated by using direct methods such as Gauss Laguerre. Carr and Madan (1999) express equation (2.8) as a Fourier transform and use Fast Fourier Transform (FFT) technique to speed up the integration. Fang and Oosterlee (2008) apply a Cosine Fourier transform on the transition probability $p^{\mathbb{Q}}(X_T|X_t)$, and convert the integral into a summation of just a few key terms, which is much faster than FFT in most cases.

2.6 Conclusion

This chapter reviews the literature related to the three studies in this thesis and provides research questions in the next few chapters. I summarised a list of research

questions which I aim to address in this thesis:

1. How to specify the jump dependence between assets and what is the impact of jumps on portfolio selection? Unlike diffusion processes, it is difficult to specify the dependence between discontinuous jump processes, especially in the multivariate setting. Current research impose strong assumptions to make the estimation of the model feasible. Hence, the estimated risk might be under- or over-estimated, and the portfolio weights derived under such assumptions could be sub-optimal. The question is how much damage will such mis-specification cost? In Chapter 3, I solve such a problem in a non-parametric way.

2. Empirical evidence shows that the variance risk premium is abnormally high after big unexpected shocks. The question is how does one identify such events in the market and how to model the corresponding underlying dynamics? Another challenging question that follows is how to measure the risk premium associated with such big events? Is the framework good enough to explain asset pricing in rare events such as sub-prime crisis? Chapter 5 provides an answer to these questions.

3. The pricing methods for VIX future and option are discussed extensively in the literature. However, not many models fit the market prices well, especially for the entire volatility surface. Even for the models which fit the market prices well, they may not provide a correct dynamics for VIX. For the purpose of hedging, is the model reliable? In particular, how does one calculate the correlation between VIX futures with different maturities, and how does one match the model implied correlation to the market observations? These questions are addressed in Chapter 6.

Chapter 3

Jump Mis-specification and Portfolio Selection

3.1 Introduction

In this chapter we make a distinction between systemic co-jumps and independent idiosyncratic jumps, and examine the impact of their mis-specification on asset allocation. Specifically, we develop a framework that allows securities to have three distinct components: a diffusion part, a co-jump component that is related to systemic events, and an idiosyncratic jump component. In our model, the intensity and the timing of the systemic jump is the same across all assets whilst the timing and intensity of the idiosyncratic jump is, obviously, asset specific. It is to be noted that this chapter is about how jumps mis-specification may lead to jumps mis-estimation and suboptimal portfolios. However, jumps, by definition, are rare events and difficult to forecast over short horizons. Long horizon jump estimates are not efficient for use in short horizon portfolio rebalancing decisions. The recognition that stock market returns do jump together as well as separately is nevertheless crucial in cross markets asset allocation policy.

In order to empirically estimate jumps, we apply the Markov Chain Monte Carlo (MCMC) methodology to MSCI stock index returns for six developed markets and five emerging markets covering over 20 years of data, which includes several well known events that were widely covered by the media and had an impact on several markets. Our results show the mis-specification of jumps and co-jump has an economically significant impact, especially in cases where emerging markets are part of

the portfolio constituents. Our MCMC methodology is particularly interesting in this case as, when compared to the parametric approach, it does not impose a prior on the dependence structure of jumps.

The impact of jumps and high moments in portfolio selection has been demonstrated in many recent papers (see Ang and Bekaert (2002), Guidolin and Timmermann (2008), and Jondeau and Rockinger (2006)). Stock return jumps (or drastic volatility regime shifts) are a response to news arrival and political events (see Aggarwal et al. (1999) and Maheu and McCurdy (2004)). The existence of jumps in stock returns plays a crucial role in the context of international markets, as effective portfolio diversification may be compromised because it is difficult to hedge jump risk. Although there is significant empirical evidence of jumps in international markets (see Asgharian and Bengtsson (2006), Asgharian and Nossman (2011), and Pukthuanthong and Roll (2012) amongst others), it seems that greater attention has been given to the impact of systemic jumps in international portfolio allocation than to idiosyncratic jumps (see Ang and Chen (2002), Das and Uppal (2004), and Ait-Sahalia et al. (2010)).

The remainder of this chapter is organised as follows: Section 2 introduces the theoretical framework and derives the solution to the optimal portfolio weights under systemic and idiosyncratic jumps. Section 3 presents the data. Section 4 introduces the MCMC methodology. The results for the optimal portfolio weight are presented in Section 5. Finally, Section 6 concludes.

3.2 Asset Price Dynamics and Portfolio Selection

This section introduces the dynamics of stock prices and wealth, and shows the optimum portfolio weights in the presence of correlated systemic jumps as well as independent idiosyncratic jumps.

3.2.1 Asset Price Dynamics

Here, we start by describing the price dynamics for nth asset $S_n(t)$, for $n = 1, \ldots, N$,

$$d \ln S_n(t) = (\alpha_n - \frac{\sigma_n^2}{2})dt + \sigma_n dB_n(t) + \int_{\mathbb{R}^+ \times \mathbb{R}} J_n \mu(dt, dJ_n) \tag{3.1}$$

where α_n and σ_n are the drift and volatility respectively, $B_n(t)$ is the standard Brownian Motion. The variance-covariance of the diffusion part is Ω, with a typical element $cov(\sigma_n dB_n, \sigma_m dB_m) = \sigma_{nm} dt$. $\mu(dt, dJ_n)$ is the random measure for jump J_n in the log return. The jump is compound Poisson jump with compensator $\nu(dt, dJ_n) = f(J_n; \theta, \delta)\lambda_n dJ_n dt$, where λ_n is the jump intensity and $f(J_n; \theta, \delta^2)$ is the normal distribution with mean θ and variance δ^2. The stock price specification in equation (3.1)is the same as that in Das and Uppal (2004) except that they assume all the return series share the same jump timing and jump sizes are perfect correlated, while we leave the jump dependence unspecified in the model.

Let $\omega = [\omega_1, \ldots, \omega_N]$ be the $N \times 1$ vector of wealth invested in N risky assets and, without loss of generality, let the initial wealth be equal to unity. We further assume that the investor can choose freely between riskless and risky assets. Therefore, for any $t = 1, \ldots, T$, investor's wealth is given by

$$d \ln W(t) = (\omega' \boldsymbol{R} + r_f - \frac{\omega' \boldsymbol{\Omega} \omega}{2})dt + \omega' \boldsymbol{\sigma} \cdot d\boldsymbol{B}(t) + \int_{\mathbb{R}^+ \times \mathbb{R}^N} \omega' \boldsymbol{J} \mu(dt, d\boldsymbol{J}), \quad (3.2)$$

where $W(t) = \omega' \boldsymbol{S}(t)$, $\mathbf{S}(t) = [S_1(t), \ldots, S_N(t)]$, r_f is the risk free interest rate, $\boldsymbol{R} = [\alpha_1 - r_f, \ldots, \alpha_N - r_f]$, $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_N]$, $\boldsymbol{B}(t) = [B_1(t), \ldots, B_N(t)]$, $(\boldsymbol{\sigma} \cdot d\boldsymbol{B}(t))$ denotes the element-by-element multiplication of σ_n and $dB_n(t)$. $\mu(dt, d\boldsymbol{J})$ is the random measure for N dimensional jump process $\boldsymbol{J} = [J_1, \ldots, J_N]$. Since the jump dependence is left unspecified, we do not have explicit form for the compensator of this random measure.

3.2.2 Optimal Portfolio Weight

Considering the investor with power utility who wishes to maximise her expected utility over her terminal wealth, W_T, the portfolio allocation problem at time t becomes,

$$V(W_t, t) \equiv \max_{\omega} \mathbb{E}\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right], \quad (3.3)$$

where $\gamma > 0$ is the constant relative risk aversion. This problem yields the following Bellman equation:

$$\begin{aligned} 0 \quad &= \max_{\omega} \left\{ \tfrac{\partial V(W_t,t)}{\partial t} + \tfrac{\partial V(W_t,t)}{\partial W_t} W_t [\omega' \boldsymbol{R} + r_f] + \tfrac{1}{2} \tfrac{\partial^2 V(W_t,t)}{\partial W_t^2} W_t^2 \omega' \boldsymbol{\Omega} \omega \right. \\ &\quad \left. + \lambda \mathbb{E}[V(W_t + W_t \omega' \boldsymbol{J}_t, t) - V(W_t, t)] \right\} \end{aligned} \quad (3.4)$$

As shown in Das and Uppal (2004), $V(W_t, t)$ has the functional form,

$$V(W_t, t) = A(t)\frac{W_t^{1-\gamma}}{1-\gamma}$$

So substituting such solution form, the optimal portfolio weights can be derived by maximizing,

$$\frac{1}{A(t)}\frac{dA(t)}{dt} + (1-\gamma)\left[\omega'\boldsymbol{R} + r_f\right] - \frac{1}{2}\gamma(1-\gamma)\omega'\boldsymbol{\Omega}\omega + \lambda\mathbb{E}\left[V(1+\omega'\boldsymbol{J}_t, t) - 1\right] \quad (3.5)$$

Differentiating the expression above with respect to ω leads to the first order condition,

$$\boldsymbol{R} + \gamma\boldsymbol{\Omega}\omega + \lambda\mathbb{E}\left(\boldsymbol{J}\left(1+\omega'\boldsymbol{J}\right)^{\gamma}\right) = 0. \quad (3.6)$$

The equation is exactly the same as that in Das and Uppal (2004), because the dependence of jumps does not affect the derivation above. In fact, the jump dependence only affects the last term in equation (3.6), hence changes the optimal portfolio weights.

3.3 Jump, Co-Jump & Parameter Estimation

3.3.1 Data and Descriptive Statistics

Our data set consists of Wednesday to Wednesday weekly continuously compounded returns of eleven MSCI stock market indices downloaded from Datastream.[1] The sample period covers January 13, 1988 to October 7, 2009 with 1135 weekly log returns. Following Das and Uppal (2004), we split the sample in two groups: one containing developed markets (the United States (US), the United Kingdom (UK), Switzerland (SW), Germany (GE), France (FR) and Japan (JP)) and one containing emerging markets (Argentina (ARG), Hong Kong (HKG), Mexico (MEX), Singapore (SNG), and Thailand (THA)).

Table 3.1 reports the summary statistics for the weekly returns. We will use these descriptive statistics to explain the test results later. In the mean time, it is worth noting that emerging markets have, in general, higher mean returns and standard deviations. Argentina, in particular, has very high positive skewness and kurtosis.

[1] These weekly returns are measured in US$ for the convenience of the asset allocation tests later.

Table 3.1: Descriptive Statistics

Markets	Mean(%)	StDev(%)	Min(%)	Max(%)	Skew	Ex Kurtosis
Panel A: Developed markets						
France	0.1466	3.0118	-17.5809	12.8290	-0.6121	3.1732
Germany	0.1337	3.2413	-17.5035	13.9773	-0.7471	3.2504
Japan	-0.0267	3.1029	-15.0566	14.6184	0.1070	1.5683
Swiss	0.1756	2.5515	-13.6478	10.3072	-0.4325	2.0416
UK	0.0802	2.5824	-15.2200	11.5638	-0.4346	3.4238
US	0.1261	2.2800	-16.7475	10.3441	-0.6014	4.6401
Panel B: Emerging markets						
Argentina	0.2554	7.2176	-36.2730	80.5246	0.9951	16.5782
HK	0.1521	3.4953	-22.5505	13.5248	-0.7023	3.4432
Mexico	0.3471	4.5676	-37.3137	18.8525	-0.6544	6.0793
Singapore	0.1233	3.2106	-17.0474	18.3912	-0.2466	3.4104
Thailand	0.0645	5.0325	-22.8786	24.7848	-0.0466	2.5501

This table reports the first four moments of the weekly return for 11 MSCI stock indices, which are separated into two groups; developed markets and emerging markets. Sample period covers January 13, 1988 to October 07, 2009, with 1135 observations in total.

In Table 3.2 the cross correlation and covariance amongst the eleven stock markets are presented. The highest correlation coefficients are observed among the European markets, in particular between France and Germany, France and the UK, France and Switzerland and Germany and the UK. The lowest correlation is between Thailand and the US. The pairwise correlation is a lot lower among the emerging markets, especially those for Argentina, Mexico, and Thailand.

3.3.2 Estimating Jump and Co-Jump

One typical way of estimating a system of several jump diffusion time series is to specify the dependence of co-jumps, e.g., multivariate normal distribution or some copula function, and then estimate the parameters for the assumed dependence structure. Such parametric approach, however, imposes a strong prior in the estimation, which will bias the estimation results. For example, Das and Uppal (2004) assume all jumps are co-jumps. As a result, the jump estimation for the US stock index changes dramatically depending on which group it is placed in the joint estimation. When estimated together with developed stock markets, the jump intensity of the US is 0.0501 per month. However, when it is estimated together with the group of

Table 3.2: Cross Correlation and Covariance

Panel A: Developed markets

	France	Germany	Japan	Swiss	UK	US
France	0.0009	0.8479	0.4294	0.7515	0.7691	0.6324
Germany	0.0008	0.0011	0.4263	0.7559	0.7192	0.6333
Japan	0.0004	0.0004	0.001	0.4014	0.4202	0.3325
Swiss	0.0006	0.0006	0.0003	0.0007	0.6937	0.5568
UK	0.0006	0.0006	0.0003	0.0005	0.0007	0.6285
US	0.0004	0.0005	0.0002	0.0003	0.0004	0.0005

Panel B: Emerging markets

	US	Argentina	HK	Mexico	Singapore	Thailand
US	0.0005	0.2373	0.4443	0.5091	0.4492	0.2929
Argentina	0.0004	0.0052	0.2249	0.3232	0.2040	0.2007
HK	0.0004	0.0006	0.0012	0.3788	0.6412	0.4640
Mexico	0.0005	0.0011	0.0006	0.0021	0.3567	0.2896
Singapore	0.0003	0.0005	0.0007	0.0005	0.001	0.5607
Thailand	0.0003	0.0007	0.0008	0.0007	0.0009	0.0025

This table reports the covariance (bottom left) and correlation (top right) between developed markets in Panel A and between emerging markets in Panel B. Sample period covers from January 13, 1988 to October 07, 2009, with 1135 observations in total.

emerging markets, the jump intensity is just 0.0138, but the mean of the jump size doubled. It is clear that the (US stock returns) jump characteristics vary depending on the group members' multivariate jump pattern when the estimation is done jointly. But from an economic perspective, it is not clear why the (US stock returns) jump estimation should be conditional on there being a jump in another market.

Another problem with the parametric multivariate approach is the complex dependence structure of jumps. Even for the simplest multinormal distribution, the number of correlation parameters to be estimated is $\frac{1}{2}(N^2 - N)$ for N markets. To apply a method like moment matching, one would require many high moment conditions in order to identify the model's parameters. Unless one imposes strong assumptions on the dependence structure, e.g. some parameters are known or the same across series, the estimation will be highly unstable.

To overcome these problems, we assume that the information of jumps for any return series given by the return series itself is the same as that from N stock markets jointly, i.e.,

$$f(J_n|r_1, r_2, \cdots r_N) \quad = \quad f(J_n|r_n)$$

Based on this assumption, we perform the jump estimation into two steps. the first step is to estimate the jumps for each univariate return series, since the estimation of jumps does not rely on the information of other return series. The second step is to estimate the cross sectional dependence of the jump processes given the exact timing and size of the jump estimated from the univariate estimation. The method of Markov Chain Monte Carlo (MCMC) is well suited for solving this problem, since it gives the jump estimates, as well as the sample draws from the jump distribution, the latter of which can be used to estimate the cross sectional jump dependence between two assets via a bootstrap method. In the next two subsections, we first introduce the MCMC method and report the univariate estimation, and then explain how to use the sample of jump processes to calculate the relevant terms in equation (3.6).

In the first step of the two-step estimation, we estimate one single marginal distribution for jumps of a particular return series. This jump estimate encompasses both the systemic and the idiosyncratic jumps since is is not possible to distinguish the two in the univariate setting. Hence, the jump intensity from the univariate estimation is the sum of the systemic and the idiosyncratic jump intensities.

3.3.3 Univariate Estimation with MCMC

The MCMC is a simulation methodology for sampling from a desired probability distribution by constructing a Markov Chain of conditional distributions. It was introduced into finance by Bradley P. Carlin (1992), and successfully applied in the estimation of stochastic volatility models with jumps by Eraker et al. (2003) and Eraker (2004).

The key idea in the MCMC estimation is to construct the posterior distribution of the parameters Θ, $p(\Theta|Y)$, from the Bayes' rule and given the observation Y using the likelihood, $p(Y|\Theta)$, and the prior distribution, $p(\Theta)$. One advantage of the MCMC over other estimation methods is that it can handle latent process in the data, such as the jump process in stock returns. Indeed, one of the key advantages of the MCMC method is the extraction of the latent variables as the byproduct of the estimation.

Since the MCMC procedure here is applied to univariate series only, we drop

the subscript n for ease of exposition. Considering equation (3.1) and the set of parameters $\Theta = \{\alpha, \sigma, \theta, \delta, \lambda\}$, it is possible to construct a Markov Chain for the jump process and for the parameters by sampling from their posterior distributions. If the posterior and prior distributions are of the same family then they are called conjugate distributions, and the prior is called a conjugate prior. For the model proposed here all the parameters have corresponding conjugate priors. The details of conjugate prior are in Appendix.

Following Eraker et al. (2003), we ignore the probability of multiple jumps in the estimation, since the jump intensity is small. At each time t, we calculate the ex-post likelihood of no jump and one jump, $L(N_t^i = m|\Theta^{i-1}, Y_t^{i-1})$, $m = \{0, 1\}$. Therefore,

$$L(N_t^i = 1) = p(N_t = 1; \lambda^{i-1}) \int_{-\infty}^{\infty} f(r_t - \alpha^{i-1}; \xi, \sigma^{i-1}) f(\xi; \theta^{i-1}, \delta^{i-1}) d\xi \quad (3.7)$$

$$L(N_t^i = 0) = p(N_t = 0; \lambda^{i-1}) f(r_t - \alpha^{i-1}; 0, \sigma^{i-1}) \quad (3.8)$$

where r_t is the log return at time t, $p(m; \lambda)$ is the probability density function (pdf) of a Poisson distribution with mean value λ, $f(x; a, b)$ is the pdf of a normal distribution with mean value a and variance b^2. Since we consider the case of maximum one jump at any point in time, the ex-post probability distribution of jump number $P(N_t^i = m|\Theta^{i-1}, Y_t^{i-1})$ is a Bernoulli distribution,

$$P(N_t^i = m) = \frac{L(N_t^i = m)}{\sum_{i=0}^{1} L(N_t^i = i)}, \quad m = \{0, 1\} \quad (3.9)$$

Next, we draw from a uniform distribution and if the draw is greater than $P(N_t^i = 0)$, we set $N_t = 1$. The corresponding jump size is sampled using the Griddy Gibbs method from the distribution

$$P^J(J_t^i = \xi) \propto f(r_t - \alpha^{i-1}; \xi, \sigma^{i-1}) f(\xi; \theta^{i-1}, \delta^{i-1}), \quad t = 1 \cdots T \quad (3.10)$$

by the inverse CDF method, in which ξ is calculated at 200 grid points for the range of $(-5\text{abs}(r_t), 5\text{abs}(r_t))$. This process is repeated for $t = 1 \cdots T$ till we get $J_{1:T}^i$.

Given the sampled jump size and timing, we can sample model parameters for next iteration step,

$$P(\Theta^i) \propto f(r_{1:T}, J_{1:T}^i; \Theta^{i-1}) f(\Theta^{i-1})$$

where $f(\Theta)$ is the prior distribution for parameters. For the parameter set $\Theta =$

$\{\alpha, \sigma, \theta, \delta, \lambda\}$, we can use the conjugate prior to speed up the sampling. Given the prior distribution, the posterior distribution is known in close form in some cases. Otherwise a difficult numerical integration may be necessary as we did for the jump number and size.[2]

By sampling the parameters, Θ^i, and the jump process, $J^i_{1:T}$, iteratively, we construct a Markov Chain for each of them till they converge to a stationary distribution. The length of sample path is 6000. To eliminate the effect of initial value of parameters, Θ^0, we burn in the first 1000 samples.[3]

Table 3.3 reports the mean and standard error for all the parameters, estimated as the mean and the standard error of the stationary distributions obtained from the Markov Chain.

3.3.4 Probability of Co-Jump

Given the sample of jump processes from MCMC, we use a bootstrap method to estimate the probability of jump and co-jump between each pair of markets. By randomly picking the jump sample paths of any two return series, we can evaluate the dependence of jumps. Repeating this sampling process enough times (5000 in our estimation) and taking average, we have the estimates of co-jump statistics between any two return series. Table 3.4 presents the co-jump intensities of the two-asset portfolios, and the univariate jump intensity for each asset in the diagonal.

Compared with the jump intensity reported in Table 3.3, there is slight difference in the univariate jump intensity. This is because the jump intensity reported in Table 3.4 is calculated by bootstrap method using samples of jump process drawn from the MCMC method.

Table 3.5 reports the probability that a jump in asset i is co-jump with asset j. The conditional jump probability is defined as,

$$Jump\,(i,j) = \frac{\lambda_{i,j}}{\lambda_i} \tag{3.11}$$

where $\lambda_{i,j}$ is the co-jump intensity between markets i and j and λ_i is the jump inten-

[2]For a normal distribution, the conjugate prior and posterior distributions for the mean and variance are Normal-inverse-gamma distributions. For a Bernoulli distribution, the conjugate prior and posterior distributions for the probability are beta distributions. The prior distributions used in our estimation are $N - \Gamma^{-1}\{0, 10, 10, 20\}$ for $\{\alpha, \sigma^2\}$, $N - \Gamma^{-1}\{0, 10, 10, 30\}$ for $\{\theta, \delta^2\}$, and $Beta\{2, 20\}$ for $\{\lambda\}$.

[3]A Geweke test is done to ensure the convergence of distributions.

Table 3.3: Univariate Jump Diffusion Model Parameter Estimates

	$\alpha(\%)$	$\sigma(\%)$	$\theta(\%)$	$\delta(\%)$	λ
France	0.3511	6.1599	-2.6556	30.4200	0.0762
	(0.0877)	(0.3828)	(0.8275)	(6.1703)	(0.0171)
Germany	0.4126	6.8339	-3.6080	33.9471	0.0776
	(0.092)	(0.4254)	(0.9008)	(6.5991)	(0.0171)
HK	0.4228	7.8940	-3.1372	39.4002	0.0865
	(0.0989)	(0.4997)	(0.8814)	(7.4159)	(0.0173)
Japan	-0.1007	7.7952	1.2430	25.2938	0.0616
	(0.0942)	(0.4702)	(0.9601)	(6.0989)	(0.0192)
Singapore	0.2768	6.1476	-1.5259	38.5254	0.1007
	(0.0888)	(0.4161)	(0.7281)	(6.5693)	(0.0185)
Swiss	0.3026	4.8379	-1.8245	20.3501	0.0699
	(0.0749)	(0.2894)	(0.6919)	(4.37)	(0.0177)
UK	0.1970	4.4971	-1.5136	25.2636	0.0772
	(0.0737)	(0.277)	(0.7135)	(5.0679)	(0.0176)
US	0.2481	3.2960	-1.4072	19.9060	0.0853
	(0.0638)	(0.2177)	(0.5678)	(3.6704)	(0.0182)
Argentina	0.2661	23.5286	-0.0572	278.5364	0.1004
	(0.1641)	(1.5132)	(1.7134)	(45.2857)	(0.0165)
Mexico	0.4821	12.0227	-1.4936	92.3373	0.0909
	(0.1196)	(0.7482)	(1.1015)	(15.9921)	(0.0166)
Thailand	0.2002	16.3472	-1.4259	88.6337	0.0948
	(0.1428)	(1.0424)	(1.1708)	(16.1407)	(0.0184)

Estimated parameters and standard errors (in brackets) of 11 markets. Sample period covers from January 13, 1988 to October 07, 2009, with 1135 observations in total. The jump intensity is expressed as number of jumps per week.

Table 3.4: Jump and Co-Jump Intensities

Panel A: Developed Markets						
	France	Germany	Japan	Swiss	UK	US
France	0.0760					
Germany	0.0249	0.0779				
Japan	0.0072	0.0079	0.0611			
Swiss	0.0187	0.0189	0.0066	0.0695		
UK	0.0209	0.0202	0.0073	0.0164	0.0771	
US	0.0175	0.0176	0.0084	0.0161	0.0176	0.0851
Panel B: Emerging Markets						
	US	Argentina	HK	Mexico	Singapore	Thailand
US	0.0851					
Argentina	0.0105	0.0998				
HK	0.0129	0.0135	0.0863			
Mexico	0.0142	0.0146	0.0151	0.0907		
Singapore	0.0175	0.014	0.0251	0.0157	0.1005	
Thailand	0.0115	0.0116	0.0175	0.0119	0.0226	0.0945

This table reports the co-jump intensity for two-asset portfolios based on the bootstrap method using jump process samples drawn from the MCMC method. The diagonal element is the jump intensity for univariate return series.

sity of market i. It should be noted that the matrix of conditional jump probability is not symmetric, since the conditioning variable is different. For example, if there is a jump in the US, there is 20.64% chance that there is a jump in UK in the same week, but 22.78% vice versa.

There are a few interesting observations from Tables 3.4 and 3.5. First, the co-jump probability is significantly higher between developed markets than between emerging markets. This is not surprising since stock markets in developed markets are more harmonised and integrated, while the emerging markets tend to be driven by idiosyncratic events. Japan appears to be an exception in the developed markets group, perhaps because of its isolation due to geographical and culture reasons. Among the group of emerging markets, HK and Singapore are more likely to jump together than other pairs in the same group, perhaps they share the same time zone, culture, and are the most developed among the emerging markets group.

Table 3.5: Proportion of Co-Jump to Total Jumps

Panel A: Developed markets						
	France	Germany	Japan	Swiss	UK	US
France		0.3273	0.0944	0.2461	0.2748	0.2297
Germany	0.3193		0.1011	0.2427	0.2599	0.2263
Japan	0.1175	0.1289		0.1081	0.1193	0.1368
Swiss	0.2691	0.2720	0.0950		0.2358	0.2318
UK	0.2707	0.2625	0.0945	0.2125		0.2278
US	0.2050	0.2070	0.0981	0.1893	0.2064	
Panel B: Emerging markets						
	US	Argentina	HK	Mexico	Singapore	Thailand
US		0.1239	0.1515	0.1672	0.2054	0.1346
Argentina	0.1057		0.1357	0.1465	0.1405	0.1159
HK	0.1495	0.1569		0.1753	0.2905	0.2030
Mexico	0.1569	0.1611	0.1668		0.1733	0.1308
Singapore	0.1740	0.1395	0.2495	0.1565		0.2249
Thailand	0.1213	0.1223	0.1853	0.1255	0.2390	

The table reports the proportion of two markets' co-jump to the total jump of the market in the row.

3.3.5 Impact of Jump and Co-Jump

Given the sample of univariate jump processes, there are two ways to calculate the last two terms in equation (3.6). First, we can specify and estimate some copula function based on the sample of the jump process drawn from MCMC procedures. However, since co-jumps are typically rare, such an estimation will be very unstable and different copula function will produce very different results.

The second, and nonparametric, approach is to calculate the jump terms in equation (3.6) directly using the samples drawn from the MCMC,

$$\begin{aligned} \lambda \mathbb{E}\left[\boldsymbol{J}\left(1+\omega'\boldsymbol{J}\right)^{\gamma}\right] &= \lambda \mathbb{E}\left[\int \boldsymbol{J}\left(1+\omega'\boldsymbol{J}\right)^{\gamma}\mu(d\boldsymbol{J})\right] \\ &\approx \frac{1}{TM}\sum_{t=1}^{T}\sum_{i=1}^{M} J_{i,t}\left(1+\omega' J_{i,t}\right)^{\gamma} \end{aligned} \tag{3.12}$$

where the second approximation is by using the Monte Carlo method and each J_i is a draw from the MCMC method. Instead of assuming some specific form of cross sectional dependence of jumps, we use the sample from MCMC to represent the joint

dependence in a nonparametric way as described in Section 3.3.5. Compared with the parametric approach, this approach is much more stable. A second advantage of this approach is that it can be easily extended to higher dimensions or more complex dependence structure, for example, it can still be applied when only a subset of the portfolio constituent markets experience co-jumps, which is very difficult to model and solve using the parametric method.

3.4 Portfolio Selection

To evaluate the potential impact of jump estimation and mis-specification, we examine two-asset portfolios based on the eleven MSCI stock index return series. To provide a direct comparison with the results reported in Das and Uppal (2004), we group the eleven index returns into developed markets and emerging markets. The developed market group consists of MSCI stock index returns for France, Germany, Japan, Switzerland, the UK and the US, while the emerging market group includes Hong Kong, Singapore, Argentina, Mexico, and Thailand. We add US to the second group taking the perspective from a US investor.

3.4.1 Only Systemic Co-jump

To highlight the importance of distinguishing systemic jumps and independent idiosyncratic jumps in jump estimation, we study portfolios consist of only two assets. We replicate the estimation method for a model with systemic jumps only as in Das and Uppal (2004) for the two-asset portfolio case using the same moment conditions. However, for the two-asset portfolio, we have five parameters to estimate; one for the jump intensity λ and 2×2 for the jump distributions, while we only have six moment conditions, 2×1 kurtosis and 2×2 co-skewness. To improve the stability

of the estimation, we use an extra co-kurtosis condition as follows,

$$\begin{aligned}
\text{CoSkewness} &= \frac{\frac{\partial^3 \ln \Phi(\omega_n,\omega_m)}{\partial \omega_n^2 \partial \omega_m}}{\text{variance}_n \sqrt{\text{variance}_m}} \\
&= \frac{\lambda\left[2\mu_n\nu_n\nu_m + \mu_m(\mu_n^2+\nu_n^2)\right]}{\text{variance}_n \sqrt{\text{variance}_m}} \\
\text{Excess CoKurtosis} &= \frac{\frac{\partial^4 \ln \Phi(\omega_n,\omega_m)}{\partial \omega_n^2 \partial \omega_m^2}}{\text{variance}_n \text{variance}_m} \\
&= \frac{\lambda\left[4\mu_n\mu_m\nu_n\nu_m + 2\nu_n^2\nu_m^2 + (\mu_n^2+\nu_n^2)(\mu_m^2+\nu_m^2)\right]}{\text{variance}_n \text{variance}_m}
\end{aligned}$$

where $\Phi(\omega_n, \omega_m)$ is the joint characteristic function of return series and is reported in Das and Uppal (2003).[4]

The estimated co-jump intensity is reported in Table 3.6 while the mean and the standard deviation of jump distributions are reported in Table 3.7. Since the assumption there is all markets jump together, the jump intensity is for co-jump only, and no idiosyncratic jumps in both markets. In Table 3.7, the element in ith row and jth column is the mean and variance of jumps in ith market when it is in a portfolio with jth market. For example, for the portfolio of US and Argentina, the mean jump size is -0.1311 for the US market and 0.5623 for the Argentinian market.

The first observation from Tables 3.6 and 3.7 is that the jump intensity for one market depends on which other market it is grouped with. For example, the jump intensity for France is 0.0344 per week when it is in the portfolio with Germany, whereas it is just 0.0042 per week if it is paired with the US. If we refer to Table 3.7, we see that magnitude of jump size for French stock market is much higher when it is paired with US (-0.1553) than with Germany (-0.0403). This is because only larger movements in France are characterised as jumps when it is grouped with US, while more frequent and relatively smaller jumps are also identified as jumps when France is grouped with Germany. We can observe similar patterns in other pairs. Under the strict co-jump assumption of Das and Uppal (2004), jumps that take place in only one market will not be recoginised as jumps.

Another problem with the co-jump assumption in Das and Uppal (2004) is that it may bias the estimation of jump distribution. According to Table 3.1, only Japan

[4]It should be noted that there are other two asymmetric co-kurtosis defined as $\frac{\partial^4 \ln \Phi(\omega_n,\omega_m)}{\partial \omega_n \partial \omega_m^3}$ and $\frac{\partial^4 \ln \Phi(\omega_n,\omega_m)}{\partial \omega_n^3 \partial \omega_m}$, but these were not used in the estimation here.

Table 3.6: Estimates of Co-Jump Intensity

Panel A: Developed markets						
Market	France	Germany	Japan	Swiss	UK	US
France		0.0334	0.0163	0.0322	0.0105	0.0042
Germany			0.0285	0.0096	0.0230	0.0079
Japan				0.0145	0.0045	0.0045
Swiss					0.0063	0.0033
UK						0.0016
Panel B: Emerging markets						
Market	US	Argentina	HK	Mexico	Singapore	Thailand
US		0.0008	0.0004	0.0020	0.0030	0.0019
Argentina			0.0041	0.0008	0.0005	0.0005
HK				0.0006	0.0046	0.0008
Mexico					0.0016	0.0016
Singapore						0.0001

This table reports the estimates of the co-jump intensity for two-asset portfolios by assuming, as in Das and Uppal (2004), only systemic jumps. The moment conditions are the same as in Das and Uppal (2004) but with an extra co-kurtosis moment.

and Argentina have positive skewness, which suggests jumps are, in general, negative except these two markets.[5] Indeed, we do observe positive mean jump size for Japan and Argentina when they are paired up with other markets. However, we also noted many positive mean jump sizes in Table 3.7 for other markets (e.g. HK when paired with US or Mexico, Singapore when paired with US or Mexico, and Thailand when paired with US, HK, or Mexico.). Since idiosyncratic jumps are omitted in the co-jump model, the strong assumption of co-jump biased the jump estimates. The bias is less severe for developed markets since their jumps are more likely to be co-jumps.

These bias in jump estimates will certainly cause problem in portfolio selection. If less jumps are identified or negative jump is treated as positive, risk will be underestimated. We will further analyze this impact by comparing this approach with our approach in the next section.

3.4.2 Optimal Portfolio and CEQ

To fully appreciate the importance of distinguishing systematic and idiosyncratic jumps, we calculate optimal portfolio weights by solving equation (3.6). First we

[5]Since the distribution of jump is normal, only when the mean of jump size is negative, it will lead to negative skewness for the return series.

Table 3.7: Mean and Standard Deviation of Jump Size Distribution

Panel A: Developed markets						
Market	France	Germany	Japan	Swiss	UK	US
France		-0.0403	-0.0803	-0.0407	-0.0943	-0.1553
		0.0604	0.0503	0.0607	0.0524	0.0154
Germany	-0.0601		-0.0781	-0.1389	-0.0877	-0.1453
	0.0553		0.0454	0.0011	0.0431	0.0093
Japan	0.0169	0.0127		0.0214	0.0244	0.0097
	0.0717	0.0628		0.0727	0.0986	0.1012
Swiss	-0.0294	-0.0836	-0.0642		-0.1033	-0.1236
	0.0467	0.0302	0.0380		0.0189	0.0360
UK	-0.0389	-0.0249	-0.0995	-0.0573		-0.1278
	0.0745	0.0641	0.0582	0.0781		0.0766
US	-0.0823	-0.0573	-0.0939	-0.0925	-0.1666	
	0.0677	0.0669	0.0563	0.0645	0.0064	
Panel B: Emerging markets						
Market	US	Argentina	HK	Mexico	Singapore	Thailand
US		-0.1311	0.0687	-0.0977	-0.1129	-0.1140
		0.0988	0.1660	0.0821	0.0545	0.0724
Argentina	0.5623		0.2042	0.5585	0.6202	0.6365
	0.4419		0.3901	0.4472	0.4779	0.4794
HK	-0.2749	-0.1476		0.0725	-0.1827	-0.1235
	0.1192	0.0727		0.2208	0.0051	0.0625
Mexico	-0.3681	-0.2819	-0.3682		-0.2932	-0.2631
	0.1736	0.2139	0.1736		0.1261	0.1117
Singapore	0.0118	-0.1852	-0.0593	0.0245		-0.3552
	0.1408	0.1415	0.1136	0.1631		0.1041
Thailand	0.1082	-0.2685	0.0367	0.1575	-0.0420	
	0.2058	0.2098	0.1566	0.2171	0.4360	

The table reports the estimates of mean and standard deviation of jump size estimated for two-asset portfolios by assuming, as in Das and Uppal (2004), only systemic jumps. The moment conditions are the same as in Das and Uppal (2004) but with an extra co-kurtosis moment.

Table 3.8: Optimal Weight for Developed markets

	France	Germany	Japan	Swiss	UK	US
		0.4321	0.6732	0.0071	0.7576	0.3393
France		0.4286	0.6338	-0.0255	0.7227	0.3130
		0.4228	0.6086	-0.0334	0.6811	0.2960
	0.0673		0.5609	-0.1181	0.5134	0.2304
Germany	0.0260		0.5139	-0.1656	0.4654	0.1961
	0.0130		0.4824	-0.1685	0.4209	0.1681
	-0.4061	-0.3753		-0.4527	-0.2709	-0.3075
Japan	-0.4039	-0.3730		-0.4521	-0.2721	-0.3099
	-0.4083	-0.3665		-0.4532	-0.2812	-0.3278
	0.7639	0.8836	0.9912		1.1071	0.6797
Swiss	0.7639	0.8971	0.9530		1.0667	0.6593
	0.7343	0.8638	0.9200		1.0589	0.6326
	-0.4005	-0.1845	0.4158	-0.4798		-0.0943
UK	-0.3941	-0.1741	0.3984	-0.4710		-0.0962
	-0.3843	-0.1529	0.3705	-0.5027		-0.1140
	0.3221	0.3981	0.7446	0.1819	0.6726	
US	0.3134	0.3956	0.7174	0.1618	0.6484	
	0.2928	0.3845	0.6866	0.1416	0.6246	

This table reports the portfolio weights for portfolio in the developed market group. The entry in ith row and jth column represent the weight in ith market, if in portfolio with jth market. The three numbers in each cell are, respectively, the portfolio weights assuming no jumps, Das and Uppal (2004) approach, and our approach.

assume there is no jump, then the optimal weight is just the solution of mean variance analysis. Second, optimal weight is calculated by following Das and Uppal (2004), assuming no idiosyncratic jump and use the parameters estimated in Section 3.4.1. Finally, equation (3.6) is solved using numerical method introduced in Section 3.3.5. For all the three cases, we set risk aversion parameter γ to 3 , a constant risk free interest rate at 3% per annum and set the time horizon to one year.

The portfolio weights for developed and emerging markets are reported in Tables 3.8 and 3.9 respectively. Each cell contains three numbers, which correspond to the optimal weights obtained by mean-variance analysis, Das and Uppal (2004), and our approach respectively. The entry in ith row and jth column represents the weight in ith market, in the two-asset portfolio. First, we note that the investment in risky markets decreases when jump is considered. Such finding is consistent with Das and Uppal (2004). When the co-jump assumption is relaxed and the idiosyncratic jump is

Table 3.9: Optimal Weight for Emerging markets

	US	Argentina	HK	Mexico	Singapore	Thailand
		0.4082	0.3944	-0.0486	0.4584	0.5381
US		0.3996	0.3932	-0.0424	0.4356	0.5207
		0.4126	0.3710	-0.0285	0.4291	0.5155
	0.2626		0.2602	0.1837	0.2701	0.2799
Argentina	0.2561		0.2374	0.1765	0.2654	0.2763
	0.1730		0.1724	0.0989	0.1839	0.1981
	0.3099	0.3034		0.1317	0.3416	0.3911
HK	0.3004	0.2820		0.1283	0.3116	0.3681
	0.2535	0.2772		0.1026	0.2956	0.3634
	0.6414	0.5352	0.5909		0.6120	0.6255
Mexico	0.5857	0.5037	0.5496		0.5657	0.5807
	0.5434	0.4847	0.5060		0.5278	0.5496
	0.2325	0.2548	0.1402	0.0682		0.3272
Singapore	0.2248	0.2488	0.1434	0.0625		0.3219
	0.1799	0.2343	0.1116	0.0374		0.3224
	0.1043	0.0951	0.0496	0.0113	0.0586	
Thailand	0.0986	0.0909	0.0411	0.0092	0.0575	
	0.0439	0.0568	-0.0046	-0.0375	-0.0024	

This table reports the portfolio weights for portfolio in the emerging market group. The entry in ith row and jth column represent the weight in ith market, if in portfolio with jth market. The three number in each cell are, respectively, the portfolio weights assuming no jumps, Das and Uppal (2004) approach, and our approach.

taken into account, the investment in risky market is further reduced. As mentioned before, the strong assumption of co-jumps in Das and Uppal (2004) under-estimates the risk of jumps by omitting the idiosyncratic jumps. Risk aversed investors dislike jumps, and respond by reducing risky investment in order to reduce risk.

To understand the economic significance of jump mis-estimation and the omission of idiosyncratic jumps, we calculate the Certainty Equivalent measure (CEQ), which is, by definition, the additional wealth required to increase the expected utility of terminal wealth under the sub-optimal portfolio strategy to that achieved by the optimal strategy.

$$V\left(\left(1+\mathrm{CEQ}\right)W_t, t; \hat{\omega}\right) = V\left(W_t, t; \omega\right) \tag{3.13}$$

For an investor with power utility, the CEQ is given by

$$\mathrm{CEQ} = \left[\frac{A(t,\mathbf{w})^{\frac{1}{1-\gamma}}}{A(t,\hat{\mathbf{w}})} - 1\right] \tag{3.14}$$

where

$$\begin{aligned} A(t,\omega_i) = \; & \exp\Big(\big((1-\gamma)\left[\omega_i'\boldsymbol{R} + r_f\right] - \tfrac{1}{2}\gamma\left(1-\gamma\right)\omega_i'\Omega\omega_i \\ & + \lambda E\left[\left(1+\omega_i'\boldsymbol{J}\right)^{1-\gamma} - 1\right]\big)\left(T-t\right)\Big), \quad \omega_i = \{\omega, \hat{\omega}\} \end{aligned}$$

In Table 3.10, we report two CEQ numbers for each portfolio, the first is the Das and Uppal (2004) approach against the mean-variance analysis and the second is our approach against the mean-variance analysis. We did not compare the Das and Uppal (2004) approach against our approach as these two approaches can nest each other. But by comparing CEQ of Das and Uppal (2004) against that of our approach, we can still gauge the effect of jump assumptions on portfolio selection.

First, all CEQs in Panel A are relatively small, suggesting the negligible impact of omitting idiosyncratic jumps in the developed markets group. Nevertheless, the CEQ by our approach is consistently 2 to 8 times higher than that by Das and Uppal (2004) approach. However, for the emerging market group in panel B, the increase of CEQ is much more pronounced, and the CEQ gain by our approach is economically significant. The case for Argentina is the most extreme. The results confirm that portfolio performance is seriously affected if the idiosyncratic jumps are not estimated correctly.

Other collaborative evidence can be found in the original return series. Argentina,

Table 3.10: Certainty Equivalent(CEQ)

Panel A: Developed markets						
$\times 10^{-3}$	France	Germany	Japan	Swiss	UK	US
France		0.1671	0.1095	0.1432	0.0697	0.0703
		0.3255	0.3211	0.2787	0.3302	0.2643
Germany			0.1823	0.1369	0.1535	0.1073
			0.4960	0.3359	0.5069	0.4030
Japan				0.0768	0.0171	0.0343
				0.2735	0.1415	0.2005
Swiss					0.0608	0.0618
					0.2314	0.2869
UK						0.0277
						0.1741
Panel B: Emerging markets						
$\times 10^{-3}$	US	Argentina	HK	Mexico	Singapore	Thailand
US		0.0236	0.0096	0.5542	0.0365	0.0245
		3.1998	0.4157	1.5785	0.3472	0.8221
Argentina			0.3079	0.2621	0.0138	0.0105
			3.4298	4.2096	3.1496	3.3268
HK				0.3533	0.0794	0.0933
				1.6620	0.4316	0.8602
Mexico					0.4369	0.3941
					1.6138	1.9641
Singapore						0.0034
						0.7870

This table reports the CEQ of two-asset portfolios. The top and bottom CEQ values in each cell are, respectively, the CEQs by Das and Uppal (2004) approach and our approach with respect to the mean-variance analysis assuming no jumps.

for example, has a big dispersion and a large positive skewness, suggesting a strong jump process not related to other stock market return. However, since only co-jumps are identified in Das and Uppal (2004), the jump intensity for Argentina in Table 3.6 Panel B ranges from 0.0005 to 0.0041 per week; but the univariate estimate in Table 3.3 is 0.1004, at least 25 times larger. So the big idiosyncratic jumps in Argentina are omitted in the co-jump restriction. This is again confirmed in Table 3.9, in which the weight in Argentina is substantially reduced in our approach. Similar bias due to the omission of idiosyncratic jumps can be found for cases involves Mexico and Thailand, but less significant for the Hong Kong and Singapore pair. As we discussed before, Hong Kong and Singapore are more likely to co-jump, so the strong assumption of co-jump has done less damage.

3.4.3 Multiple Asset Portfolio

In the previous section, we see the impact of idiosyncratic jump is economically significant in two-asset portfolios. The next question that follows is if the impact of idiosyncratic jumps still so significant in portfolio with more than two assets? To answer this question, we construct two six-asset portfolios following Das and Uppal (2004). The CEQs of Das and Uppal (2004) and our approach are presented in Table 3.11. The results for developed market is consistent with two-asset portfolios case, i.e., omitting idiosyncratic jumps and jump mis-specification have little impact on portfolio selection as the portfolio weights are similar and CEQ gains are negligible. Nevertheless, the CEQ is about 3 times higher in our approach than in Das and Uppal (2004). However, for the group of emerging markets, we see a much smaller improvement in CEQ value in the six-asset portfolio compared with two-asset portfolios, suggesting that the omission of idiosyncratic jumps is much less important as the number of assets in the portfolio increases.

Unlike the two-asset portfolio, it is difficult to categorise jumps as idiosyncratic jumps or systemic jumps in a six-asset portfolio. For example, Argentina and Mexico are more likely to jump together, while US or other markets in the same portfolio have no jump. Therefore, the risk of idiosyncratic jumps is better diversified than in two-asset portfolio.

Another explanation is that when estimating co-jump intensity in Das and Uppal (2004), it is possible that more jumps are identified as co-jumps. Using Argentina and Mexico as example, if the co-jump between them is big enough, it will affect the

Table 3.11: Comparison of CEQ and Portfolio Weight for Das and Uppal (2004) and Our Approach

Developed Market	FR	GE	JP	SW	UK	US	Sum	CEQ
Diffusion Weight	0.4139	-0.1598	-0.4370	1.0224	-0.6748	0.4447	0.6093	
Das and Uppal (2004)	0.4098	-0.1978	-0.4429	1.0413	-0.6648	0.4247	0.5703	0.1265e-3
Our Approach	0.3947	-0.2264	-0.4307	1.0278	-0.6539	0.4215	0.5331	0.4956e-3
Emerging Market	HK	SNG	US	ARG	MEX	THA	Sum	CEQ
Diffusion Weight	0.1394	0.0149	-0.1642	0.1837	0.5487	-0.0499	0.6726	
Das and Uppal (2004)	0.0830	0.0409	-0.1594	0.1160	0.5229	-0.0386	0.5649	2.768e-3
Our Approach	0.1390	0.0071	-0.1090	0.0990	0.4955	-0.0809	0.5508	4.577e-3

This table reports the optimal portfolio weights calculated for each group using the base diffusion model, Das and Uppal (2004) model, and our jump diffusion model. CEQ is calculated as Das and Uppal (2004) and our jump diffusion model with respect to the base diffusion model.

GMM estimation in Das and Uppal (2004), and make the algorithm classify it as a co-jump shared by all six markets, and hence increase the co-jump intensity. Indeed, the co-jump intensity for the six-asset portfolio of emerging markets is 0.0069, much higher than the co-jump intensity of two-asset portfolio in Table 3.6. In this case, we see that the strong assumption of co-jump in Das and Uppal (2004) might also cause an overestimation of co-jump intensity.

Therefore, the CEQ by Das and Uppal (2004) approach will be more significant in cases of six-asset portfolios than in cases of two-asset portfolios, because the higher the co-jump intensity, the higher the CEQ. If we compare the CEQ of Das and Uppal (2004) reported for two-asset portfolio and six-asset portfolio, we notice a significant improvement, which verifies our arguments. But it is questionable whether we could trust CEQs reported by Das and Uppal (2004) approach, since the impact of jumps might be either under-estimated or over-estimated.

3.5 Conclusion

In this chapter, we provide a framework for modelling security returns with three distinct components: a diffusion, a systemic jump and an idiosyncratic jump. The jump dependence of different stock markets can be estimated in a nonparametric way. In our proposed approach, the jump estimation and jump dependence are estimated empirically in two steps. First, a Markov Chain Monte Carlo (MCMC) is applied to weekly market returns of eleven MSCI stock index return which are categorized in two groups; developed and emerging markets. The sampled paths of the eleven univariate jump components are then used to study their dependence structure.

Comparing with Das and Uppal (2004), our results suggest it is important to distinguish systematic and idiosyncratic jumps in stock market returns. Similar to Das and Uppal (2004), when we restrict the modelling assumptions to pure diffusion and diffusion plus systemic common jumps, we find no significant difference in portfolio choice and and CEQ against mean-variance analysis. The omission of co-jumps is adequately compensated by a large variance. By allowing idiosyncratic jumps, we find marginal improvement in portfolio selection among developed markets, but economically significant improvement in portfolios of emerging markets. The CEQ calculated by using our approach with unrestricted jumps is large for emerging markets, compared with that based on the Das and Uppal (2004) approach. As the

co-jump restriction imposed in Das and Uppal (2004) is too strong, the frequency of jumps are severely mis-estimated. The bias is the strongest for emerging markets, as they are prone to more idiosyncratic jumps. For developed markets, since they are more harmonised, the assumption of perfectly correlated jumps is less damaging. Adding idiosyncratic jumps has less impact on portfolio that consists of only developed markets.

This chapter is about how jump mis-specification may lead to jumps mis-estimation and suboptimal portfolios. It is important to note that jumps are rare and would require a long estimation sample period. When return and portfolio performance are evaluated over short horizon, the estimation of models with jump feature will be unstable as not enough jumps are observed during the period. But for a longer term asset allocation policy, the analysis and finding here will be useful.

Chapter 4

Derivative Pricing with Affine Models and Numerical Implementation

4.1 Introduction

Affine models are popular in finance, especially after the publication of Duffie et al. (2000). They represent a compromise between tractability and empirical complexity. The characteristic function of the underlying dynamics in affine models is an exponential affine function of the state variables. As a result, the price and characteristic function of the contingent claim can be expressed as an integral and evaluated numerically. In Section 4.2, we present some stochastic differential equations (SDE) that underlie well known affine models in finance. Section 4.3 links the SDE of the underlying dynamics to partial differential equation (PDE) of the characteristic function. Some pricing examples are presented in Section 4.4. Section 4.5 summarises the reasons why affine models are so attractive in option pricing. Section 4.6 introduces two different ways for calculating the option price numerically.

4.2 Candidate SDE

Not all stochastic differential equations (SDE) will lead to an affine model. We will present a more general specification of the affine model from Duffie et al. (2000) in the next section. In this section, we introduce a few well known finance affine models

and their associated SDEs. The most famous affine model is, perhaps, the Heston stochastic volatility (SV) model as shown below

$$\begin{aligned} d\ln S_t &= (r - \frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2 \\ [dW_t^1, dW_t^2] &= \rho dt \end{aligned}$$

The popularity of the Heston model is largely due to the resulting semi-closed form solution for European options as a direct consequence of the affine formulation.

In the Heston example above, there are two state variables, S_t and v_t. Affine models can have many state variables. For example, in Duffie et al. (2000), the long run mean of stochastic variance is driven by a second square root process,

$$\begin{aligned} d\ln S_t &= (r - \frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 \\ dv_t &= \kappa(\theta_t - v_t)dt + \sigma\sqrt{v_t}dW_t^2 \\ d\theta_t &= \kappa_\theta(\theta_\theta - \theta_t)dt + \sigma_\theta\sqrt{\theta_t}dW_t^3 \\ [dW_t^1, dW_t^2] &= \rho dt \end{aligned}$$

Here, only dW_t^1 and dW_t^2 are correlated; dW_t^3 is not correlated with dW_t^1 and dW_t^2. In the two-speed mean reversion SV model below, there are two independent Brownian motions for volatility each of which is correlated with the Brownian motion that drives the stock price. All the other correlation between dW_t^i are zero:

$$\begin{aligned} d\ln S_t &= (r - \frac{v_{1,t} + v_{2,t}}{2})dt + \sqrt{v_{1,t}}dW_t^1 + \sqrt{v_{2,t}}dW_t^2 \\ dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}dW_t^3 \\ dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}dW_t^4 \\ [dW_t^1, dW_t^3] &= \rho_1 dt \\ [dW_t^2, dW_t^4] &= \rho_2 dt \end{aligned}$$

As in Bates (1996), we can further extend the affine model to include discrete jump

process as shown below

$$\begin{aligned} d\ln S_t &= (r - \frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 + J_x dN(t) \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2 \\ [dW_t^1, dW_t^2] &= \rho dt \end{aligned}$$

where J_x is the jump in log return with a constant jump intensity. Since jump is normally modelled as a discrete process, it is a common practice to assume that it is not correlated with the continuous Brownian motions. The model remains affine whether J_x follows a normal distribution or a double exponential distribution. Furthermore, the model remains affine with a closed form characteristic function even if we introduce another jump process J_v into the SV process (see Duffie et al. (2000)) so long as J_v is linear in the state variable.

The constraint imposed by a constant jump intensity can be further relaxed. For example, Fang and Oosterlee (2008) introduce the stochastic jump intensity model,

$$\begin{aligned} d\ln S_t &= (r - \frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 + J_x dN(t) \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2 \\ d\lambda_t &= \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW_t^3 \\ [dW_t^1, dW_t^2] &= \rho dt \end{aligned}$$

It should be noted that dW_t^3 should have no correlation with dW_t^1 and dW_t^2. Otherwise, the model is no longer affine. The key issue with correlation restriction is the need to maintain linearity after applying Ito's lemma. This means all the variance-covariance terms must be linear in the state variables.

4.3 From SDE to PDE

Duffie et al. (2000) present a general affine model with n-state variables as

$$dX_t = u(X_t, t)dt + \sigma(X_t, t)dW_t + \int_{\mathbb{R}^n} h(J)\mu(dJ, dt) \tag{4.1}$$

where X and J are n-dimensional vectors; $h(J)$ is a function of jump size, $\mu(dJ, dt)$

is the random measure of the Poisson jump, and dW_t represents an n-dimensional independent Brownian motion.

4.3.1 Kolmogorov Backward Equation

Many finance problems involve estimating the probability distribution of future values of the state variable given the current values. Such a problem is called the "forward" problem. The Kolmogorov backward equation, on the other hand, is more useful when one is interested in the (hitting) probability that the state variable will reach a target region B at some future point in time. As we will soon see, Kolmogorov backward equation is very useful for derivative pricing because the price of the derivative is the expectation of the future payoff, which is a function of future values of the state variables.

For the diffusion case without jumps, the Kolmogorov backward equation is

$$-\frac{\partial}{\partial t}p(X_t,t) = \frac{\partial}{\partial X}p(X_t,t)\cdot u(X_t,t) + \frac{1}{2}\text{tr}\left(\Sigma(X_t,t)\frac{\partial^2}{\partial X^2}p(X_t,t)\right) \tag{4.2}$$

where $\Sigma(X_t,t) = \sigma(X_t,t)\sigma(X_t,t)^T$ is the covariance matrix of the diffusion part, tr is the sum of the diagonal of the matrix, $p(X_t,t)$ is the hitting probability with terminal condition,

$$p(X_T,T) = 1_{X(T)\in B} \tag{4.3}$$

and B is a predefined set. As an example, one could have a digital option that pays £1 if S_T is within the set B.

We can extend the backward equation (4.2) to include jumps. Assuming that jump time is independent from jump size, i.e. $\mu(dJ,dt) = \mu(dt)\mu(dJ)$, then equation (4.2) becomes

$$\begin{aligned} -\frac{\partial}{\partial t}p(X_t,t) &= u(X_t,t)\frac{\partial}{\partial X}p(X_t,t) + \frac{1}{2}\text{tr}\left(\Sigma(X_t,t)\frac{\partial^2}{\partial X^2}p(X_t,t)\right) \\ &+ \lambda_t\int_{\mathbb{R}^n}\left(p(X_t+h(J)) - p(X_t)\right)\mu(dJ) \end{aligned} \tag{4.4}$$

where λ_t, as one of the state variables, is the jump intensity such that $\lambda_t dt = \nu(dt)$, which is the compensator of the random measure $\mu(dt)$. To simplify the equation,

we can introduce the differential operator

$$\mathcal{L}(\cdot) = \frac{\partial}{\partial t}(\cdot) + u(X_t, t)\frac{\partial}{\partial X}(\cdot) + \frac{1}{2}\text{tr}\left(\Sigma(X_t, t)\frac{\partial^2}{\partial X^2}\right)(\cdot) + \lambda_t \int_{\mathbb{R}^n} \Delta(\cdot)\mu(dJ)$$

and now the backward equation (4.4) can be written as

$$\mathcal{L}p(X_t, t) = 0. \tag{4.5}$$

4.3.2 Green's Function

A Green's function is an integral kernel that can be used to solve a differential equation that is itself inhomogeneous or one with inhomogeneous boundary conditions. In the latter case, the solution involves changing the boundary condition in (4.3) into the delta function below

$$p(X_T, T) \quad = \quad \delta(X_T - x) \tag{4.6}$$

where x is a particular realisation of X_T. That is x is the only element in the set B in equation (4.3). Moreover,

$$p(X_T = x|X_t, t) = \mathbb{E}[p\,(X_T, T) \times 1_{X_T=x}|X_t, t] = \mathbb{E}[\delta(X_T - x)|X_t, t].$$

Using the Black-Scholes model as an example, we have in place of equation (4.1)

$$dS_t = rS_t dt + \sigma S_t dW_t \tag{4.7}$$

Here, S_t is the only state variable in the vector X_t. Let f denote the option whose payoff, $g\,(S)$, depends on S, and

$$\mathcal{L}f = \frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} \tag{4.8}$$

The value of the derivative at time T can now be expressed in terms of equation (4.6) as follows:

$$f_T = \int \delta(S_T - s)g(s)ds$$

The current option price as the discounted expectation of future payoff ,

$$f_t = e^{-r(T-t)} E\left[\int \delta(S_T - s) g(s) ds\right] = e^{-r(T-t)} \int p(S_T = s|S_t, t) g(s) ds$$

and by applying the differential operator

$$\begin{aligned} \mathcal{L}f &= re^{-r(T-t)} \int p(S_T = s|S_t, t) g(s) ds + \mathcal{L} \int p(S_T = s|S_t, t) g(s) ds \\ &= rf + \int \mathcal{L} p(S_T = s|S_t, t) g(s) ds \\ &= rf \end{aligned} \tag{4.9}$$

where the second equality is due to the fact that $\mathcal{L}$ is a linear operator. Since we know $\mathcal{L}p(S_T = s|S_t, t) = 0$ in Kolmogorov backward equation in equation (4.5), the last term $\int \mathcal{L}p(S_T = s|S_t, t) g(s) ds = 0$. Then the PDE for the current price of a contingent claim on the future payoff function $g(S)$ is, by combining equations (4.8) and (4.9),

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

which is exactly the same as the Black-Scholes fundamental PDE derived by hedging argument.

4.3.3 Feynman Kac Formula

The Feynman Kac formula provides a relationship between a solution of a PDE and the expectation of a stochastic process. For a one-dimensional diffusion process,

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW,$$

the Feynman Kac formula states that any function $f(X_t, t)$ defined as

$$f(X_t, t) = \mathbb{E}[f(X_T, T)|X(t)]$$

is the solution to the following PDE,

$$\frac{\partial f}{\partial t} + \mu(X_t, t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(X_t, t)\frac{\partial^2 f}{\partial x^2} = 0. \tag{4.10}$$

This result is closely connected with the results in the previous sections. Note that equation (4.10) can be applied to any functions $f(X_t, t)$, which is not necessary a derivative price.

Let us define $f(X_T, T) = \delta(X_T - x)$ and let us assume that the SDE has exactly the same form as equation 4.7. Then from equation (4.6),

$$\begin{aligned} f(X_T, T) &= p(X_T, T) \\ f(X_t, T) &= \mathbb{E}[p(X_T = x|X_t, t)] \end{aligned}$$

and we get the backward equation (4.5).

One of the most important results derived from the Feynman Kac formula is related to the characteristic function. Let

$$\begin{aligned} f(X_T, T) = \phi(\omega, X_T, T) = e^{i\omega X_T} \\ f(X_t, t) = \phi(\omega, X_t, T) = \mathbb{E}[e^{i\omega X_T}|X_t] \end{aligned}$$

Then the Feynman Kac formula implies

$$\mathcal{L}\phi(\omega, X_t, T) = 0 \tag{4.11}$$

Here, $\phi(\omega, X_t, T)$ is the characteristic function of X_T and is the key to many option pricing solutions as explained in the next sections.

4.4 Characteristic Function Examples

In this section, we show three examples on how the PDE was solved through the use of characteristic functions.

4.4.1 Jump Diffusion Model

The first example has only one state variable, and we have a jump diffusion process for stock price with SDE below

$$dx_t = (r - \frac{1}{2}\sigma^2 - \lambda m)dt + \sigma dW_t + JdN_t \tag{4.12}$$

where x is the log price, $\ln S_t$, and N_t is a Poisson Process with constant jump intensity λ, $m = \mathbb{E}[e^J - 1]$ is the mean value of jump size J, and the jump size J has a distribution π_J.

Based on the convolution property of characteristic function, the $\mathcal{F}_t$-conditional characteristic function of X_T is

$$\begin{aligned}
&\phi(\omega, x_t, \sigma^2, \lambda, t, T) \\
=&\phi_{Diffusion} \times \phi_{Jump} \\
=&e^{x_t i\omega + r(T-t)i\omega - \lambda m i\omega (T-t) - \frac{1}{2}(i\omega+\omega^2)(T-t)\sigma^2} e^{(\phi_{Jump}(\omega)-1)(T-t)\lambda}
\end{aligned} \tag{4.13}$$

where ϕ_{Jump} is the characteristic function of Jump size distribution π_J. The first equality holds because of the independence of the diffusion and the jump parts. In this simple example, the characteristic function can be easily derived in closed form by evaluating the mean of the diffusion part and the jump part in equation (4.12) separately. Note that the characteristic function in equation (4.13) is an exponent of affine function of σ^2, λ and the state variable x_t.

As we have a jump component in the stock price, the PDE includes an integral term for the jump as shown below

$$\frac{\partial f}{\partial t} + \mu(X_t, t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(X_t, t)\frac{\partial^2 f}{\partial x^2} + \lambda_t \int \left(f(X_t + h(J)) - f(X_t)\right) \mu(dJ) = 0$$

where μ is the drift term in equation (4.12). Such PDE is called PIDE (Partial Integral Differential equation). Invoking the Feynman-Kac formula in characteristic function form in equation (4.11) and use the fact that the moment generating function (MGF) $G(\omega) = \phi(-i\omega)$, we have

$$\begin{aligned}
G_t &+ (r - \frac{1}{2}\sigma^2 - \lambda m)G_x + \frac{1}{2}\sigma^2 G_{xx} \\
&+ \lambda \int_{\mathbb{R}} \left(G(x+J) - G(x)\right) f_J dJ \\
&= 0
\end{aligned} \tag{4.14}$$

with boundary condition $G_T = e^{x_T \omega}$. From the PIDE and the boundary condition, one may guess the solution to be of the form

$$G(\omega, x_t, v_t, \lambda_t, t, T) = e^{\omega x_t + r(T-t)\omega + A(\omega,t,T) + B(\omega,t,T)\sigma^2 + C(\omega,t,T)\lambda} \tag{4.15}$$

and for the jump component,

$$G(x + J) - G(x) = G(x)(e^{J\omega} - 1). \tag{4.16}$$

Substitute the solution forms, (4.15) and (4.16), into equation (4.14), we have

$$\begin{aligned} & (r - \frac{1}{2}\sigma^2 - \lambda m)\omega G + \frac{1}{2}\sigma^2\omega^2 G \\ & + \lambda G \int_{\mathbb{R}} (e^{J\omega} - 1) dJ \\ = & \left(r\omega - \frac{\partial A(\omega,t,T)}{\partial t} - \frac{\partial B(\omega,t,T)}{\partial t}\sigma^2 - \frac{\partial C(\omega,t,T)}{\partial t}\lambda \right) G \end{aligned}$$

Cancelling G on both sides, we get

$$\begin{aligned} & -\frac{1}{2}\sigma^2\omega + \frac{1}{2}\sigma^2\omega^2 \\ & + \lambda \int_{\mathbb{R}} (e^{J\omega} - 1) dJ - \lambda m\omega \\ = & -\frac{\partial A(\omega,t,T)}{\partial t} - \frac{\partial B(\omega,t,T)}{\partial t}\sigma^2 - \frac{\partial C(\omega,t,T)}{\partial t}\lambda \end{aligned}$$

Regroup the equation with respect to the parameters,

$$\begin{aligned} & (-\frac{1}{2}\omega + \frac{1}{2}\omega^2 + \frac{\partial B(\omega,t,T)}{\partial t})\sigma^2 \\ & + (\int_{\mathbb{R}} (e^{J\omega} - 1) dJ - m\omega + \frac{\partial C(\omega,t,T)}{\partial t})\lambda \\ & + \frac{\partial A(\omega,t,T)}{\partial t} \\ = & \ 0 \end{aligned}$$

which means that the following system of ODEs must have a solution

$$\begin{aligned}
\frac{\partial A(\omega,t,T)}{\partial t} &= 0 \\
\frac{\partial B(\omega,t,T)}{\partial t} &= \frac{1}{2}(\omega-\omega^2) \\
\frac{\partial C(\omega,t,T)}{\partial t} &= -\int_{\mathbb{R}}(e^{J\omega}-1)\pi_J dJ + m\omega
\end{aligned}$$

given the boundary condition, $G(\omega, x_T, V_T, T, T) = e^{x_T\omega}$, $A(\omega, T, T) = 0$, $B(\omega, T, T) = 0$, and $C(\omega, T, T) = 0$. Therefore we have the solutions,

$$\begin{aligned}
A &= 0 \\
B &= -\frac{1}{2}(\omega-\omega^2)(T-t) \\
C &= \left(\int_{\mathbb{R}}(e^{J\omega}-1)\pi_J dJ - m\omega\right)(T-t)
\end{aligned}$$

where $\int_{\mathbb{R}} e^{J\omega} f_J dJ = \phi_{Jump}(\omega)$ is the MGF of the jump size distribution. For normal distribution, $J \sim N(\epsilon, \delta^2)$,

$$\int_{\mathbb{R}}(e^{J\omega}-1)dJ = e^{\epsilon\omega+\delta^2\omega^2/2}.$$

Substituting A, B, and C into equation (4.15), we have $G(\omega) = \phi(-i\omega)$, which is exactly the relationship between MGF and the characteristic function as defined above.

4.4.2 Heston Model

The Heston SDE has two state variables as follows

$$\begin{aligned}
d\ln S_t &= (r-\frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 \\
dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2 \\
[dW_t^1, dW_t^2] &= \rho dt.
\end{aligned}$$

The characteristic function is given as

$$\phi(\omega, x_t, v_t, t, T) \;=\; \mathbb{E}[e^{i\omega x_T}|x_t, v_t]$$

where $x_t = \ln S_t$. As in the previous subsection, we invoke the Feynman-Kac formula in characteristic function form and use the moment generating function (MGF) of $\ln S_T$, $G(\omega) = \phi(-i\omega)$,

$$\begin{aligned} G_t &+ (r - \frac{1}{2}v_t)G_x + \frac{1}{2}v_t G_{xx} + \kappa(\theta - v_t)G_v \\ &+ \frac{1}{2}\sigma^2 v_t G_{vv} + \rho\sigma v_t G_{xv} \\ &= 0 \\ G_T &= e^{x_T\omega}. \end{aligned} \tag{4.17}$$

From the PDE and the boundary condition, one may guess that the solution of (4.17) is the exponential of an affine function in the state variables, x_t and v_t, and thus a possible solution is

$$G(\omega, x_t, v_t, \lambda_t, t, T) = e^{\omega x_t + r(T-t)\omega + A(\omega,t,T) + B(\omega,t,T)v_t} \tag{4.18}$$

Substituting the solution form into equation (4.17), we have

$$\begin{aligned} & -\frac{1}{2}v_t\omega + \frac{1}{2}v_t\omega^2 + \kappa(\theta - v_t)B(\omega, t, T) \\ + \;& \frac{1}{2}\sigma^2 v_t B(\omega, t, T)^2 + \rho\sigma v_t \omega B \\ = \;& -\frac{\partial A(\omega, t, T)}{\partial t} - \frac{\partial B(\omega, t, T)}{\partial t}v_t. \end{aligned}$$

Regrouping the equation with respect to the state variables v_t,

$$\begin{aligned} & (-\frac{1}{2}\omega + \frac{1}{2}\omega^2 - \kappa B(\omega, t, T) + \frac{1}{2}\sigma^2 B(\omega, t, T)^2 + \rho\sigma\omega B + \frac{\partial B(\omega, t, T)}{\partial t})v_t \\ + \;& \kappa\theta B(\omega, t, T) + \frac{\partial A(\omega, t, T)}{\partial t} \\ = \;& 0 \end{aligned}$$

Since v_t is a time varying processes, in order to get zero in the RHS, we have

$$\begin{aligned}
\frac{\partial A(\omega,t,T)}{\partial t} &= -\kappa\theta B(\omega,t,T) && (4.19)\\
\frac{\partial B(\omega,t,T)}{\partial t} &= \frac{(\omega-\omega^2)}{2} + (\kappa-\rho\sigma\omega)B(\omega,t,T) - \frac{\sigma^2 B(\omega,t,T)^2}{2} && (4.20)
\end{aligned}$$

Therefore, the PDE is now decomposed as a set of ODEs. The second nonlinear ODE is known as Riccati differential equation, which has an analytical solution which can be obtained by a change of variable. The solutions of (4.19) and (4.20) are

$$\begin{aligned}
A(\omega,t,T) &= -\frac{\kappa\theta}{\sigma^2}(\psi_+(T-t) + 2\ln\frac{\psi_- + \psi_+ e^{-\zeta(T-t)}}{2\zeta}\\
B(\omega,t,T) &= -(\omega-\omega^2)\frac{1-e^{-\zeta(T-t)}}{\psi_- + \psi_+ e^{-\zeta(T-t)}}\\
\zeta &= \sqrt{(\kappa-\rho\sigma\omega)^2 + \sigma^2(\omega-\omega^2)}\\
\psi_\pm &= \zeta \mp (\kappa-\rho\sigma\omega)
\end{aligned}$$

Finally, we verify that the terminal boundary condition is satisfied $G(\omega, x_T, V_T, T, T) = e^{x_T\omega}$ when $A(\omega,T,T) = 0$ and $B(\omega,T,T) = 0$.

4.4.3 Multiple Factors Affine Models

Here, we add stochastic jump intensity to Heston's stochastic volatility model;

$$\begin{aligned}
d\ln S_t &= (r - \frac{v_t}{2})dt + \sqrt{v_t}dW_t^1 + J_x dN(t)\\
dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2\\
d\lambda_t &= \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW_t^3\\
[dW_t^1, dW_t^2] &= \rho dt
\end{aligned}$$

Following the same procedure as in the previous section, we derive the PIDE for the moment generating function of $\ln S_T$,

$$
\begin{aligned}
G_t &+ (r - \frac{1}{2}v_t - \lambda_t m)G_x + \frac{1}{2}v_t G_{xx} \\
&+ \kappa(\theta - v_t)G_v + \frac{1}{2}\sigma^2 v_t G_{vv} + \rho\sigma v_t G_{xv} \\
&+ \kappa_\lambda(\theta_\lambda - \lambda_t)G_v + \frac{1}{2}\sigma_\lambda^2 \lambda_t G_{\lambda\lambda} \\
&+ \lambda_t \int_{\mathbb{R}} \left(G(x+J) - G(J)\right) f_J dJ \\
&= 0 \\
G_T &= e^{x_T \omega}
\end{aligned}
\tag{4.21}
$$

where $m = \int_{\mathbb{R}} e^J f_J dJ - 1$. We propose the solution of equation (4.21) as an exponential function of affine function of state variables, x_t, v_t and λ_t as follows

$$
G(\omega, x_t, v_t, \lambda_t, t, T) = e^{\omega x_t + r(T-t)\omega + A(\omega,t,T) + B(\omega,t,T)v_t + C(\omega,t,T)\lambda_t}
\tag{4.22}
$$

and when substituted into equation (4.21), we have

$$
\begin{aligned}
&\frac{1}{2}v_t\left(\omega^2 - \omega\right) - \lambda_t\left(m\omega - \int_{\mathbb{R}} e^{J\omega} f_J dJ - 1\right) \\
+ \;& \kappa(\theta - v_t)B(\omega,t,T) + \frac{1}{2}\sigma^2 v_t B(\omega,t,T)^2 + \rho\sigma v_t \omega B \\
+ \;& \kappa_\lambda(\theta_\lambda - \lambda_t)C(\omega,t,T) + \frac{1}{2}\sigma_\lambda^2 \lambda_t C(\omega,t,T)^2 \\
= \;& -\frac{\partial A(\omega,t,T)}{\partial t} - \frac{\partial B(\omega,t,T)}{\partial t}v_t - \frac{\partial C(\omega,t,T)}{\partial t}\lambda_t
\end{aligned}
$$

Regroup the equation with respect to the state variables v_t and λ_t,

$$
\begin{aligned}
&(-\frac{1}{2}\omega + \omega^2 - \kappa B(\omega,t,T) + \frac{1}{2}\sigma^2 B(\omega,t,T)^2 + \rho\sigma\omega B + \frac{\partial B(\omega,t,T)}{\partial t})v_t \\
+ \;& (-\kappa_\lambda C(\omega,t,T) + \frac{1}{2}\sigma_\lambda^2 C(\omega,t,T)^2 + \frac{\partial C(\omega,t,T)}{\partial t})\lambda_t \\
+ \;& \kappa\theta B(\omega,t,T) + \kappa_\lambda\theta_\lambda C(\omega,t,T) + \frac{\partial A(\omega,t,T)}{\partial t} \\
= \;& 0
\end{aligned}
$$

The corresponding set of ODEs is

$$\frac{\partial A(\omega,t,T)}{\partial t} = -\kappa\theta B(\omega,t,T) - \kappa_\lambda\theta_\lambda C(\omega,t,T) \tag{4.23}$$

$$\frac{\partial B(\omega,t,T)}{\partial t} = \frac{(\omega-\omega^2)}{2} + (\kappa-\rho\sigma\omega)B(\omega,t,T) - \frac{\sigma^2 B(\omega,t,T)^2}{2} \tag{4.24}$$

$$\frac{\partial C(\omega,t,T)}{\partial t} = \kappa_\lambda C(\omega,t,T) - \frac{1}{2}\sigma_\lambda^2 C(\omega,t,T)^2 - \Lambda \tag{4.25}$$

and the solution, as shown in Sepp (2003), is

$$\begin{aligned}
A(\omega,t,T) &= -\frac{\kappa\theta}{\sigma^2}\left(\psi_+(T-t) + 2\ln\frac{\psi_- + \psi_+ e^{-\zeta(T-t)}}{2\zeta}\right) \\
&- \frac{\kappa_\lambda\theta_\lambda}{\sigma_\lambda^2}\left(\chi_+(T-t) + 2\ln\frac{\chi_- + \chi_+ e^{-\xi(T-t)}}{2\xi}\right) \\
B(\omega,t,T) &= -(\omega-\omega^2)\frac{1-e^{-\zeta(T-t)}}{\psi_- + \psi_+ e^{-\zeta(T-t)}} \\
C(\omega,t,T) &= 2\Lambda\frac{1-e^{-\xi(T-t)}}{\chi_- + \chi_+ e^{-\xi(T-t)}} \\
\Lambda &= \int_{\mathbb{R}} e^{J\omega} f_J dJ - 1 - \omega\left(\int_{\mathbb{R}} e^{J} f_J dJ - 1\right) \\
\zeta &= \sqrt{(\kappa-\rho\sigma\omega)^2 + \sigma^2(\omega-\omega^2)} \\
\psi_\pm &= \zeta \mp (\kappa - \rho\sigma\omega) \\
\xi &= \sqrt{\kappa_\lambda^2 - 2\sigma_\lambda^2\Lambda} \\
\chi_\pm &= \xi \mp \kappa_\lambda
\end{aligned}$$

4.5 Why Affine

As shown in the three examples in the previous section, if the SDE is affine, the characteristic function is an exponential affine function of the state variables, $X_t = [X_1, \cdots, X_n]'$, e.g. $\phi(\omega, X_t, T) = e^{g(X_t)}$ and $g(X_t) = A_0 + A_1X_1 + \cdots + A_nX_n$. Then we can solve the PDE (4.11) in terms of the characteristic function as follows

$$
\begin{aligned}
\mathcal{L}\phi &= \frac{\partial \phi}{\partial t} + u(X_t,t)\frac{\partial \phi}{\partial X} + \frac{1}{2}\mathrm{tr}\left(\Sigma(X_t,t)\frac{\partial^2 \phi}{\partial X^2}\right) + \lambda_t(\int_{\mathbb{R}^n} \Delta\phi\,(\mu(dJ)) \\
&= [\dot{A}_1,\cdots,\dot{A}_t][X_1,\cdots,X_n]' e^{g(X(t))} + u(X_t,t)[A_1,\cdots,A_n]' e^{g(X_t)} \\
&+ \frac{1}{2}[A_1,\cdots,A_n]\Sigma(X_t,t)[A_1,\cdots,A_n]' e^{g(X_t)} \\
&+ e^{g(X_t)}\lambda_t \int_{\mathbb{R}^n}(e^{h(J)}-1)\mu(dJ) \\
&= 0
\end{aligned}
$$

where $\dot{A}$ represent the derivative of A with respect to time t. We first eliminate the term $e^{g(X_t)}$ to get

$$
\begin{aligned}
& \dot{A}_0 + [\dot{A}_1,\cdots,\dot{A}_t][X_1,\cdots,X_n]' + u(X_t,t)[A_1,\cdots,A_n]' \\
+ & \frac{1}{2}[A_1,\cdots,A_n]\Sigma(X_t,t)[A_1,\cdots,A_n]' \\
+ & \lambda_t \int_{\mathbb{R}^n}(e^{h(J)}-1)\mu(dJ) \\
= & \ 0
\end{aligned}
\tag{4.26}
$$

Recall that jump intensities are part of the state variables. We let the first m states to be from the diffusion part and the remaining $n-m$ states to be from the jump part, e.g. $X_q = \lambda_{q-m}$ for $q = m+1,\cdots,n$. Using the features of affine models that,

$$
\begin{aligned}
u(X,t)[A_1,\cdots,A_n]' &= \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n \\
[A_1,\cdots,A_n]\Sigma(X(t),t)[A_1,\cdots,A_n]' &= \beta_0 + \beta_1 X_1 + \cdots + \beta_m X_m \\
\lambda_t \int_{\mathbb{R}^n}(e^{h(J)}-1)\mu(dJ) &= \beta_{m+1}\lambda_{1,t} + \cdots + \beta_n \lambda_{n-m,t}
\end{aligned}
$$

Now, the PDE (4.26) can be regrouped as,

$$
\begin{bmatrix} \dot{A}_0 + \alpha_0 + \beta_0 \\ \dot{A}_1 + \alpha_1 + \beta_1 \\ \vdots \\ \dot{A}_n + \alpha_n + \beta_n \end{bmatrix}' \begin{bmatrix} 1 \\ X_1 \\ \vdots \\ X_n \end{bmatrix} = 0
$$

Since $X_{t,i}$ represents each individual dynamics of the underlying state variable, the coefficient must be zero to make the above equation holds all the time. Therefore, we have decomposed ODE set as follows,

$$\begin{bmatrix} \dot{A}_0 + \alpha_0 + \beta_0 \\ \dot{A}_1 + \alpha_1 + \beta_1 \\ \vdots \\ \dot{A}_n + \alpha_n + \beta_n \end{bmatrix} = 0$$

Compared with the original PDE, the ODE set is much easier to solve, either in closed form or numerically.

4.6 Numerical Integral

In the previous section, we show that it is relative easy to solve for characteristic function if the SDE is affine. In this section, we are going to introduce several ways of using the characteristic function to price options.

4.6.1 Fourier Transform

For any contingent claim, the current price is the discounted expectation of future payoff under the risk neutral measure,

$$\begin{aligned} f(X_t, t) &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[f(X_T, T)|X_t] \\ &= e^{-r(T-t)} \int_{\mathbb{R}^n} f(X_T, T) p(X_T|X_t) dX_T. \end{aligned} \tag{4.27}$$

Let us denote $\phi(-\omega, X_t, T)$ the characteristic function of the transition distribution $p(X_T|X_t)$, and $\hat{f}(\omega)$ the Fourier transform of the payoff function $f(X_T, T)$, or alternatively $f(X_T, T)$ is the inverse Fourier transform of $\hat{f}(\omega)$. Then

$$\begin{aligned} \phi(-\omega, X_t, T) &= \int_{\mathbb{R}^n} e^{-i\omega X_T} p(X_T|X_t) dX_T \\ \hat{f}(\omega) &= \int_{\mathbb{R}} f(x, T) e^{i\omega x} dx \\ f(X_T, T) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega X_T} d\omega \end{aligned}$$

Take the European call option for example,

$$\begin{aligned}
\hat{f}(\omega) &= \int_{\mathbb{R}} f(x,T)e^{i\omega x}dx \\
&= \int_{\ln K}^{\infty} (e^x - K)e^{i\omega x}dx \qquad (4.28) \\
&= \int_{\ln K}^{\infty} e^{(1+i\omega)x}dx - \int_{\ln K}^{\infty} e^{\ln K + i\omega x}dx \qquad (4.29) \\
&= \frac{e^{(1+i\omega)x}}{1+i\omega} \mid_{\ln K}^{\infty} - \frac{Ke^{i\omega x}}{1+i\omega} \mid_{\ln K}^{\infty} \qquad (4.30)
\end{aligned}$$

where $x = \ln S_T$. Given that $\left(\frac{xe^{(1+i\omega)x}}{1+i\omega}\right)$ or $\left(\frac{Ke^{i\omega x}}{1+i\omega}\right)$ does not necessary exist as $x \to \infty$, we have to put some constrains to ensure their existence. To make $e^{(1+i\omega)x} \to 0$ and $e^{i\omega x} \to 0$ when $x \to 0$, we need $\Im(\omega) > 1$, where $\Im(\cdot)$ denotes the imaginary part. The necessary constraint varies and depends on the option payoffs. For example, for put option, the constraint is $\Im(\omega) < 0$.

The integral can further be evaluated as

$$\begin{aligned}
\hat{f}(\omega) &= -\frac{Ke^{i\omega \ln K}}{1+i\omega} + \frac{Ke^{i\omega \ln K}}{i\omega} \\
&= \frac{e^{k(i\omega+1)}}{i\omega - \omega^2}
\end{aligned}$$

where $k = \ln K$.

Substitute the definition of inverse Fourier transform of $f(X_T, T)$ into equation (4.27), we have

$$\begin{aligned}
f(X_t,t) &= \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}^n} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{-i\omega X_T} d\omega p(X_T|X_t)dX_T \\
&= \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) \int_{\mathbb{R}^n} e^{-i\omega X_T} p(X_T|X_t)dX_T d\omega \qquad (4.31) \\
&= \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)\phi(-\omega, X_t, T)d\omega \qquad (4.32)
\end{aligned}$$

We rearranged the order of integration in equation (4.31) and applied the definition of characteristic function in (4.32).

It is not critical to have closed form solution in order to calculate the option price.

As long as the value inside the integral in (4.32) at each point of ω can be calculated, the integration can be evaluated numerically.

4.6.2 Fast Fourier Transform

Carr and Madan (1999) applied the method of FFT (Fast Fourier Transform) to speed up the calculation of the integral. However, equation (4.32) is not in the form of Fourier transform, whereas the FFT can only be used on integral of Fourier transform. To take advantage of the technique of FFT, the trick is to transform the target integral to a Fourier Transform.

First, according to Carr and Madan (1999), it is important that the condition in Plancherel's theorem is satisfied which requires the price function to be square integrable in order to apply (inversion) Fourier Transform.[1] Using the European call option as example, we can see that the $C(k) \to S$ as $k \to -\infty$, which is not square integrable. Carr and Madan (1999) suggest transforming the call price as follows

$$c_T(k) \equiv e^{\alpha k} C_T(k)$$

for some suitable α, which makes $c_T(k)$ square integrable. Now we can apply Fourier Transform on $c_T(k)$,

$$\psi_T(\omega) = \int_{-\infty}^{\infty} e^{i\omega} c_T(k) dk.$$

Substituting the inverse Fourier transform in equation (4.32) and the payoff function in (4.28), we have,

[1]Note that actually $\mathcal{L}^1$ is sufficient for Fourier transform and inverse Fourier transform. The discussion here follows Carr and Madan (1999) and use $\mathcal{L}^2$, the Plancherel condition.

$$
\begin{aligned}
\psi_T(\omega) &= \int_{-\infty}^{\infty} e^{i\omega} e^{\alpha k} C_T(k) dk \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} e^{i\omega k} \int_{k}^{\infty} e^{\alpha k} (e^x - e^k) p(x_T|x_t) dx dk \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} p(x_T|x_t) \int_{-\infty}^{x} e^{(i\omega+\alpha)k} (e^x - e^k) dk dx \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} (-\frac{e^{(i\omega+1+\alpha)x}}{\alpha+1+i\omega} + \frac{e^{(i\omega+1+\alpha)x}}{\alpha+i\omega}) p(x_T|x_t) dx \\
&= e^{-r(T-t)} \frac{\phi(\omega-(\alpha+1)i, S_t, T)}{\alpha^2+\alpha+i(2\alpha+1)\omega-\omega^2} \qquad (4.33)
\end{aligned}
$$

The last equality is through the application of Fourier transform of the pdf $p(x_T|x_t)$ with variable change $v = -i(i\omega + 1 + \alpha) = \omega - (\alpha + 1)i$.

Now we can back out the original option price $C_T(k)$ through inverse Fourier transform

$$
\begin{aligned}
C_T(k) &= e^{-\alpha k} c_T(k) \\
&= \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} \psi_T(\omega) e^{-i\omega k} d\omega \qquad (4.34) \\
&\approx \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N} \psi_T(j\Delta\omega) e^{-ij\Delta\omega k} \Delta\omega
\end{aligned}
$$

with $\Delta\omega = \frac{u}{N+1}$ and u is the upper bound for the truncation of characteristic function $\psi_T(\omega)$. We refer the reader to Carr and Madan (1999) for further discussion of the choice of u .

As mentioned before, we should choose α to make $c_T(k)$ square integrable. Since $C_T(k) \to S_T$ as $k \to -\infty$, the call option price is not square integrable. One necessary condition is $\alpha > 0$. Another condition we need, as suggested by Carr and Madan (1999), is $\psi(0)$ being finite, which is to ensure that $\frac{\phi(-(\alpha+1)i)}{\alpha^2+\alpha}$ is finite. Using the definition of characteristic function, this condition is equivalent to

$$\mathbb{E}[S_T^{\alpha+1}] < \infty$$

A popular choice for α is $\frac{1}{2}$. Then the integral (4.34) can be calculated efficiently by

FFT method.

The calculation of FFT requires $N \log_2 N$ operations which is much more efficient than the direct numerical integration in Section 4.6.1, which requires at least N^2 operations.

4.6.3 Fourier Cosine Transform

Recently, Fang and Oosterlee (2008) propose the use of Fourier Cosine transform (FCT) to calculate the integral in equation (4.32). The FCT method is normally more efficient than FFT method but not without other pitfalls. For any contingent claim with final payoff $f(x_T)$, the current value $f(x_t)$ is the discounted value of the expectation of the final payoff under risk neutral measure,

$$\begin{aligned} f(x_t) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[f(x_T)|\mathscr{F}_t] \\ &= e^{-r(T-t)}\int_{\mathbb{R}} f(x_T)p(x_T|x_t)dx_T \end{aligned} \tag{4.35}$$

The idea of FCT is to express the pdf, $p(x_T|x_t)$, as a series of cosine functions. As long as the pdf is smooth enough, the number of terms required for the integral approximation is much smaller than the FFT method to achieve the same level of accuracy. An additional requirement for the FCT method to work effectively is that the function to be decomposed must have a finite support so that the tail of the distribution can be truncated as follows

$$\begin{aligned} p(x_T|x_t) &= \sum_{j=0}^{\infty} a_j(x_t)\cos(j\pi\frac{x_T-l}{u-l}) \\ &\approx \sum_{j=0}^{M} a_j(x_t)\cos(j\pi\frac{x_T-l}{u-l}) \end{aligned}$$

with

$$a_j(x_t) = \frac{2}{u-l}\int_l^u p(x_T|x_t)\cos(j\pi\frac{x_T-l}{u-l})dx_T$$

where $[l, u]$ is the lower and upper boundaries for the truncated pdf $p(x_T|x_t)$.

The option price in equation (4.35) now becomes,

$$
\begin{aligned}
f(x_t) &= e^{-r(T-t)} \int_{\mathbb{R}} f(x_T) p(x_T|x_t) dx_T \\
&\approx e^{-r(T-t)} \int_l^u f(x_T) \sum_{j=0}^{M} a_j(x_t) \cos(j\pi \frac{x_T - l}{u-l}) dx_T \\
&= e^{-r(T-t)} \sum_{j=0}^{M} a_j(x_t) \int_l^u f(x_T) \cos(j\pi \frac{x_T - l}{u-l}) dx_T \\
&= e^{-r(T-t)} \sum_{j=0}^{M} a_j(x_t) V_i \qquad (4.36)
\end{aligned}
$$

where V_i is the coefficient for the cosine transform of payoff function. Using the European call and put options as examples, Fang and Oosterlee (2008) show that

$$
\begin{aligned}
V_j^{Call} &= \frac{2}{u-l} K(\chi_j(0,u) - \psi_j(0,u)) \\
V_j^{Put} &= \frac{2}{u-l} K(-\chi_j(l,0) + \psi_j(l,0))
\end{aligned}
$$

with

$$
\begin{aligned}
\chi_j(a,b) &= \frac{1}{1+(\frac{j\pi}{u-l})^2} \left[\cos(j\pi \frac{b-l}{u-l}) e^b - \cos(j\pi \frac{a-l}{u-l}) e^a \right. \\
&+ \left. \frac{j\pi}{u-l} \sin(j\pi \frac{b-l}{u-l}) e^b - \frac{j\pi}{u-l} \sin(j\pi \frac{a-l}{u-l}) e^a \right] \\
\psi_j(a,b) &= \begin{cases} \left[\sin(j\pi \frac{b-l}{u-l}) - \sin(j\pi \frac{a-l}{u-l}) \right] \frac{u-l}{j\pi} & j \neq 0 \\ b-a & j = 0 \end{cases}
\end{aligned}
$$

Now, we need to revisit the coefficients of cosine transform of transition distribution $p(x_T|x_t)$. Noting that[2]

$$
e^{i\omega x} = \cos(\omega x) + i \sin(\omega x)
$$

where $a_i(x_t)$ is just the real part of the Fourier Transform of $p(x_T|x_t)$,

[2]It should be noted that in Fang and Oosterlee (2008), $x = \ln(S/K)$.

$$
\begin{aligned}
a_j(x_t) &= \frac{2}{u-l}\int_l^u p(x_T|x_t)\cos(j\pi\frac{x_T-l}{u-l})dx_T \\
&\approx \Re(\frac{2}{u-l}\int_{\mathbb{R}} p(x_T|x_t)e^{ij\pi\frac{x_T-l}{u-l}}dx_T) \\
&= \frac{2}{u-l}\Re(\phi(\frac{j\pi}{u-l}, X(t), T)\times e^{-i\frac{jl\pi}{u-l}})
\end{aligned}
$$

and the option price is given by equation (4.36).

4.6.4 Fast Fourier Transform vs. Fourier Cosine Transform

Both FFT and FCT methods require the truncation of the target distribution; FFT in the ω domain while FCT in the x domain. Furthermore, FFT works by separating the integration interval into buckets. It is very inefficient for long intervals and it introduces discretisation errors. FCT works in a way similar to Taylor series expansion by ignoring the higher order term of the cosine expansion of the target distribution. If the distribution function is smooth, FCT is much more efficient, as shown in Fang and Oosterlee (2008). It is because FCT uses much fewer terms in the cosine expansion in (4.36) than the number of points needed to calculate the FFT integral in (4.34), i.e. $M \ll N$.

Table 4.1: Model Parameters

	Jump Diffusion	Heston	Multiple Factors
r	3%	3%	3%
σ^2	0.1		
κ		2	2
θ		0.05	0.05
ρ		-0.6	-0.6
σ		0.6	0.6
V_0		0.05	0.05
κ_λ			8
θ_λ			1
σ_λ			2
λ_0	3		3
μ	-5%		-5%
δ	10%		10%

The distribution of jump size is $N(\mu, \delta^2)$.

For illustration, we implemented the three examples used in Section 4.4 in Matlab.

The parameters values used in each case are reported in Table 4.1. Since the FFT algorithm is optimized in Matlab, we will not compare the computation speed of the two algorithms. Instead, we focus on the pricing accuracy when valuing ATM (at-the-money), deep ITM (in-the-money) and deep OTM (out-of-the-money) European call and put options, given the current stock price $S = 100$ and two time to maturity, viz. 1 year and 1 month. The pricing error is measured by put-call parity, $Error = C - P - S + e^{-r(T-t)}K$. The results for one year maturity options are reported in Table 4.2. The first number reported is from FFT and the second number from FCT. Here we have specifically chosen $M = 128$ which is much smaller than $N = 4096$. Table 4.2 shows that the pricing errors produced by FCT is in the same order as those produced by FFT, despite the small M. The speed of FCT becomes more relevant in empirical studies that require e.g. daily calibration (or optimisation) of option pricing model using a large amount of market data each day.

Table 4.2: Performance Comparison, FFT ($N = 4096$) / FCT ($M = 128$)

$T = 1$	K	Call	Put	Error
Jump Diffusion	20	80.59/80.591	0.00/0.00	0.00/0.00
	100	15.88/15.88	12.92/12.92	0.00/0.00
	200	0.64/0.64	94.73/94.73	0.00/0.00
Heston	20	80.59 /80.59	0.00/0.00	0.00/0.00
	100	9.69/9.68	6.74/6.72	0.00/0.00
	200	0.00/0.00	94.09/94.09	0.00/0.00
Multiple Factors	20	80.59/80.59	0.00/0.00	0.00/0.00
	100	11.03/11.02	8.07/8.06	0.00/0.00
	200	0.01/0.01	94.10/94.10	0.00/0.00

As shown in Table 4.3, at short maturity such as 1 month, the option pricing start to go wrong for FCT when the payoff strongly depends on the tail of the distribution (for deep OTM options). This problem can be easily fixed by re-adjusting the truncated domain, [u,l] , to include the full set with positive payoff. But this example serves to remind us the potential problem of using FCT. On the other hand, given a large N , the FFT result is much more stable, regardless the payoff functions and how they relate to the underlying distribution.

Table 4.3: Performance Comparison, FFT ($N = 4096$) / FCT ($M = 128$)

$T = 1/12$	K	Call	Put	Error
Jump Diffusion	20	80.05/11.69	0.00/0.00	0.00/0.00
	100	4.27/4.25	4.02/4.01	0.00/0.00
	200	0.00/0.00	99.50/99.47	0.00/0.03
Heston	20	80.05/0.00	0.00/5.46	0.00/-85.51
	100	2.68/2.65	2.43/2.40	0.00/0.00
	200	0.00/0.00	99.50/87.71	0.00/11.7902
Multiple Factors	20	80.05/1.06	0.00/0.29	0.00/-79.29
	100	3.21/3.18	2.96/2.93	0.00/0.00
	200	0.00/0.00	99.50/98.49	0.00/1.01

4.7 Conclusion

In this chapter, we reviewed the general approach to price options for affine models. The pricing method is based on Characteristic function. Based on Kolmogorov backward equation, we express Characteristic function as the solution of a PIDE for the given SDE. We examined a few affine models and show how to calculate Characteristic function from the PIDE, and even in close form in some cases. In the example of option pricing, we show how to derive the European style option price as an integral, which needs to be calculated numerically. Two popular numerical methods are described. Both methods give satisfactory results. But for short term maturities deep OTM options, FFT is more stable than FCT albeit with more computation.

Chapter 5

Variance Risk Premium under Stochastic Volatility and Self-Exciting Jumps

5.1 Introduction

It is well known in the finance literature that volatility of asset return is not constant over time. Models with time varying variance, e.g. ARCH/GARCH model and Heston model, show some success in fitting both the historical time series of stock returns and option prices (see Duan (1995), Eraker (1998), and Heston (1993)). Recent extensions add discontinuous jumps in the price and/or volatility dynamics and show significant improvement in the calibration (see Duffie et al. (2000) and Eraker (2004)). Given these extra risk factors, the question immediately followed is the amount of risk premium associated with each of the risk factors. A major concern of investors and traders of vanilla and volatility derivatives is how much in terms of risk premium they should pay or charge for the traded securities, e.g., variance swap.

Demeterfi et al. (1999) show the expected future variance or the fair price of the variance swap under the risk neutral measure can be replicated by a weighted sum of OTM European options across all strikes. In another word, the expected future variance is equal to a weighted integral of European option prices. Carr and Wu (2009) use a model free approach to calculate the difference between the realized variance and the expected variance under the risk neutral measure and show that variance risk premium is significant. If variance is stochastic, Pan (2002) shows that

the variance risk premium also changes over time.

More recently, Bergomi (2004) shows that a single volatility risk factor is not flexible enough to fit the volatility term structure implied by options, one needs at least a second volatility risk factor. Bergomi (2007) uses two volatility risk factors with different timescale dynamics to drive the underlying dynamics, which is assumed to be a pure diffusion. By using high frequency data, Todorov (2010) isolates jumps from diffusion process and demonstrates that variance risk premium is unusually high after realized jumps, suggesting that realized jumps in stock price might be a risk factor contributing to the risk premium.

Inspired by these previous studies, we keep the jump process in stock price, which seems to capture the extreme market reaction well under both physical and risk neutral measures. Unlike previous studies, we set the jump intensity to follow a self exciting dynamics as in Hawkes (1971) to allow direct feedback from realized jump to future jumps.

The idea of time varying jump intensity is not new. The GARJI (Generalised Autoregressive Conditional Jump Intensity) model proposed by Maheu and McCurdy (2004) has the jump intensity following an autoregressive process and driven by the likelihood of jump in the previous day as follows,

$$\begin{aligned} \lambda_t &= \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1} \\ \xi_{t-1} &= \mathbb{E}[n_{t-1}|r_{t-1}] - \lambda_{t-1} \end{aligned}$$

where $\mathbb{E}[n_{t-1}|r_{t-1}]$ is the ex-post expectation of the number of jumps given the information at time $t-1$, λ_t is the jump intensity conditional upon information at time $t-1$, and ξ_{t-1} is the jump intensity residual measuring the strength of the unexpected big movement in the return, which impacts on the jump intensity in the next period via parameter γ, and ρ controls the speed of decaying back to the baseline level λ_0. Maheu and McCurdy (2004) estimate a range of individual companies with different time horizons and show a much fast reverting speed in conditional jump intensity compared with the decaying speed in GARCH specification. A more recent paper by Ait-Sahalia et al. (2010) propose a model with both stochastic volatility and mutually exciting jumps to model the extreme dependency across international markets, and find strong evidence of self excitement and contagion effect.

Here, we focus on univariate feedback effect and slightly modify the Hawkes process to allow feedback from only realized *negative* jumps in our proposed Stochastic

Volatility Self-Exciting Jump model. Two similar models were studied in Carr and Wu (2008) and Fulop et al. (2012). The difference between them and our study lies in the definition of jumps and its impact on the market. The jump in our framework is defined as rare events that trigger a market crisis. In the calm period, jump intensity in our model is supposed to be small, and most of the return variations and risk premium will be explained well by stochastic variance. With an unexpected big loss, the market becomes chaotic and enters into crisis period. The jump intensity in our model is then excited and triggers further large jumps. Our empirical investigation indicates such a feature is important, especially for explaining the high skewness observed in the option market during the financial crisis. High jump intensity in the crisis period represents bigger uncertainty under physical measure, which attracts a higher risk premium for anticipation of further big losses under the risk neutral measure.

To estimate the model parameters and the associated risk premium, a joint estimation is carried out using information under both $\mathbb{P}$ and $\mathbb{Q}$ measures. Instead of using option prices directly, we use the implied cumulants under $\mathbb{Q}$ measure. Cumulants or moments are summary statistics of the implied distributions. Using cumulants of the implied distributions can distinguish the intrinsic difference of different factors easily. Furthermore, we derive the analytical price for the log contract and the variance swap, and estimate how the risk premium of variance swap changes after a realized jump. The biggest challenge in the estimation is that our model is strongly nonlinear and non-Gaussian; Kalman filter or its variants is not an appropriate tool for the estimation. So we use Particle filter in the filtering, but then faced with another problem that the likelihood is not continuous w.r.t parameters to be estimated. To overcome this last problem, we adopt the Expectation Maximisation (EM) algorithm since it requires least computation and converges quickly.

The rest of this chapter is organized as follows. We start with the introduction of our Stochastic Volatility with Self Exciting Jump model in Section 5.2. We explain the joint estimation method in Section 5.3. Section 5.4 discusses the empirical finding and the time varying risk premium of variance swap. Section 5.5 concludes. The details of cumulants derivation are included in the Appendix.

5.2 Stochastic Volatility with Self-Exciting Jump

Stochastic volatility models are well established in the literature, and strong empirical evidences suggest there are jumps in both volatility and stock return. Inspired by the findings in Todorov (2010) that the variance risk premium is unusually large after realised jumps, we introduce another factor associated with a large negative shocks in returns. Therefore, we propose a model as below:

$$\begin{aligned}
\frac{dS_t}{S_t} &= udt + \sqrt{v_t}dW_{1,t} \\
&+ \int_{\mathbb{R}\times\mathbb{R}^+} (e^{J_x}-1)(\mu(dJ_x, dJ_\lambda, dt) - \pi_x(J_x)dJ_x\lambda_t dt) \\
dv_t &= \kappa_v(\theta_v - v_t)dt + \sigma_v\sqrt{v_t}dW_{2,t} \\
d\lambda_t &= \kappa_\lambda(\theta_\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW_{3,t} \\
&+ \int_{\mathbb{R}\times\mathbb{R}^+} 1_{J_x<0} J_\lambda \mu(dJ_x, dJ_\lambda, dt) \\
[dW_{1,t}, dW_{2,t}] &= \rho dt
\end{aligned} \tag{5.1}$$

where $\mu(dJ_x, dJ_\lambda, dt)$ denotes the random measures of a two dimensional jump in price and jump intensity, and there is no correlation between $W_{2,t}$ and $W_{3,t}$, and between $W_{1,t}$ and $W_{3,t}$. We also assume that their jump size and jump timing are independent, therefore, the compensator of random measures can be written as $\pi(dJ_x, dJ_\lambda)dJ_x dJ_\lambda \lambda_t dt$, where λ_t denotes the jump intensity at time t. We assume that the jumps in equity price and jump intensity are compound Poisson jumps and the jump sizes in two these processes are independent, i.e., $\pi^{\mathbb{P}}(J_x, J_\lambda) = \pi_x(J_x)\pi_\lambda(J_\lambda)$, where $\pi_x(J_x)$ and $\pi_\lambda(J_\lambda)$ are Double Exponential distribution and Exponential distribution respectively,

$$\pi_x(J_x) = 1_{x>0}\frac{p}{\eta_u}e^{-J_x/\eta_u} + 1_{x<0}\frac{1-p}{\eta_d}e^{-J_x/\eta_d} \tag{5.2}$$

$$\pi_\lambda(J_\lambda) = \frac{1}{\eta}e^{-J_\lambda/\eta} \tag{5.3}$$

where $0 < p < 1$ and $\eta, \eta_u, \eta_d > 0$. The Double Exponential distribution is introduced by Kou (2002), and has a better fit for the heavy tail of the jump size.

It should be noted that κ_λ is not the true mean reverting speed, because the innovation term is not compensated and hence not a martingale. If we adjust the random measure by its compensator, we have $\varkappa_\lambda = \kappa_\lambda - (1-p)\eta$ and $\vartheta_\lambda = \frac{\kappa_\lambda\theta_\lambda}{\varkappa_\lambda}$,

which are the mean reversion speed and the long run mean of the jump intensity respectively.

5.2.1 Dynamics under $\mathbb{Q}$ Measure

5.2.1.1 Measure Change

Since our objective is to identify the risk premium, we need the dynamics under $\mathbb{Q}$ measure. The measure change from $\mathbb{P}$ to $\mathbb{Q}$ is,

$$\begin{aligned}\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} &= \mathcal{E}\left(-\int_0^t \gamma_v\sqrt{v_s}dW_{2,s}\right)\times\mathcal{E}\left(-\int_0^t \gamma_\lambda\sqrt{\lambda_s}dW_{3,s}\right)\\ &\times \mathcal{E}\left(\int_0^t \left(e^{-\gamma_{J_x}J_x-\gamma_{J_\lambda}J_\lambda}-1\right)\left(\mu(dJ_x,dJ_\lambda,dt)-\nu(dJ_x,dJ_\lambda)\lambda_t dt\right)\right)\end{aligned}$$

where $\mathcal{E}$ is the stochastic exponential and $\nu(dJ_x,dJ_\lambda)=\pi_x(J_x)\pi_\lambda(J_\lambda)dJ_xdJ_\lambda$.

For the Brownian motion term, we shift the mean and maintain the affine form of the model. Therefore, $dW^{\mathbb{Q}}_{2,t}=\gamma_v\sqrt{v_t}dt+dW^{\mathbb{P}}_{2,t}$ and $dW^{\mathbb{Q}}_{3,t}=\gamma_\lambda\sqrt{\lambda_t}dt+dW^{\mathbb{P}}_{3,t}$ are Brownian motions under $\mathbb{Q}$ measure. For the compound Poisson jump, we change the distribution of jump size by a constant shift. The distribution for jump size in log return becomes $\pi^{\mathbb{Q}}_x(J_x)=$ Double Exponential$(\eta^{\mathbb{Q}}_u,\eta^{\mathbb{Q}}_j,p^{\mathbb{Q}})$, where $\eta^{\mathbb{Q}}_u=\frac{1}{1+\gamma_{J_x}\eta_u}\eta^{\mathbb{P}}_u$, $\eta^{\mathbb{Q}}_d=\frac{1}{1-\gamma_{J_x}\eta_d}\eta^{\mathbb{P}}_d$, and $p^{\mathbb{Q}}=p^{\mathbb{P}}\left(p^{\mathbb{P}}+(1-p^{\mathbb{P}})\frac{1+\gamma_{J_x}\eta_u}{1-\gamma_{J_x}\eta_d}\right)^{-1}$. And the distribution of jump size in the jump intensity becomes $\pi^{\mathbb{Q}}_\lambda(J_\lambda)=$ Exponential$(\eta^{\mathbb{Q}})$, where $\eta^{\mathbb{Q}}=\frac{1}{1+\gamma_{J_\lambda}\eta}\eta^{\mathbb{P}}$. Hence the dynamics under $\mathbb{Q}$ measure is as follows,

$$\begin{aligned}\frac{dF_t}{F_t} &= \sqrt{v_t}dW^{\mathbb{Q}}_{1,t}\\ &+ \int_{\mathbb{R}\times\mathbb{R}^+}(e^{J_x}-1)(\mu^{\mathbb{Q}}(dJ_x,dJ_\lambda,dt)-\pi^{\mathbb{Q}}_x(J_x)dJ_x\lambda_t dt)\\ dv_t &= \kappa^{\mathbb{Q}}_v(\theta^{\mathbb{Q}}_v-v_t)dt+\sigma_v\sqrt{v_t}dW^{\mathbb{Q}}_{2,t}\\ d\lambda_t &= \kappa^{\mathbb{Q}}_\lambda(\theta^{\mathbb{Q}}_\lambda-\lambda_t)dt+\sigma_\lambda\sqrt{\lambda_t}dW^{\mathbb{Q}}_{3,t}\\ &+ \int_{\mathbb{R}\times\mathbb{R}^+}1_{J_x<0}J_\lambda\mu^{\mathbb{Q}}(dJ_x,dJ_\lambda,dt)\\ [dW_{1,t},dW_{2,t}] &= \rho dt\end{aligned}\tag{5.4}$$

where $\kappa^{\mathbb{Q}}_v=\kappa_v+\gamma_v\sigma_v$, $\theta^{\mathbb{Q}}_v=\frac{\kappa_v\theta_v}{\kappa^{\mathbb{Q}}_v}$, $\kappa^{Q}_\lambda=\kappa_\lambda+\gamma_\lambda\sigma_\lambda$, and $\theta^{\mathbb{Q}}_\lambda=\frac{\kappa_\lambda\theta_\lambda}{\kappa^{\mathbb{Q}}_\lambda}$. Because the random

measure in the jump intensity equation is not compensated, so the risk premium of jump size in jump intensity γ_{J_λ} does not appear in the equations under $\mathbb{Q}$ measure. But for the true mean reverting speed and the long run mean of jump intensity, we have $\varkappa_\lambda^{\mathbb{Q}} = \kappa_\lambda + \gamma_\lambda\sigma_\lambda - \gamma_{J_\lambda}(1-p)\eta$ and $\vartheta_\lambda^{\mathbb{Q}} = \frac{\kappa_\lambda\theta_\lambda}{\varkappa_\lambda^{\mathbb{Q}}}$, in which γ_{J_λ} has a direct impact on the model parameters under the risk neutral measure.

5.2.1.2 Characteristic Function of Log Price

Our proposed SVSEJ model is designed to be affine, so we are able to use the transform approach from Heston (1993) to price any claim contingent on the final state of the underlying. Therefore, we can calibrate our model to fit the European option prices listed in the market. First, the characteristic function is defined as follows,

$$\begin{aligned}\phi(\omega, x_t, v_t, \lambda_t, t, T) &= \mathbb{E}[e^{i\omega x_T}|x_t, v_t, \lambda_t] \\ &= \int_{-\infty}^{\infty} e^{i\omega x_T} f(x_T|x_t, v_t, \lambda_t)dx_T\end{aligned}$$

where $x_T = \ln F_T$ is the log price, which is driven by some stochastic processes. And then the moment generating function (MGF) is,

$$G(\omega, x_t, v_t, \lambda_t, t, T) = \phi(-i\omega, x_t, v_t, \lambda_t, t, T)$$

By Feynman-Kac theorem, the expectation of some function of x_T, can be determined by a PIDE (Partial Integral Differential Equation). Following the approach by Pan (2002) and Sepp (2003), we derive the PIDE for the moment generating function (MGF) of x_T for our SVSEJ model,

$$\begin{aligned}& G_t - (\frac{1}{2}v + \lambda_t\mathbb{E}[e^{J_x} - 1])G_x + \frac{1}{2}v_t G_{xx} + \kappa(\theta - v_t)G_v \\ + \;& \frac{1}{2}\sigma^2 v_t G_{vv} + \rho\sigma v_t G_{xv} + \kappa_\lambda(\theta_\lambda - \lambda_t)G_\lambda \\ + \;& \lambda_t \int [G(\omega, x + J_x, v_t, \lambda_t + 1_{J<0}J_\lambda, t, T) \\ - \;& G(\omega, x, v_t, \lambda_t, t, T)]\,\nu(dJ_x, dJ_\lambda) \\ = \;& 0\end{aligned} \tag{5.5}$$

with initial condition at time T,

$$G_T = e^{x_T \omega}$$

As shown in Appendix A, there is no analytical solution for the MGF G because of the self exciting jump part.

5.2.2 Model Discussion

Our proposed SVSEJ model is a multi-factor model. The single factor model is well studied in literature and some single factor models show good fit for both stock price movement and option volatility surface (see Pan (2002) and Eraker (2004)). However, there are a few fundamental problems with the single factor models. First, the volatility and skewness generated by the single factor models are one-to-one correspondence for the same maturity,[1] which is not the case in the market data. Figure 5.1 shows that for a given variance implied by option prices, and there can be different values of the skewness, especially when the variance is big. In another words, there are more than one factor controlling the market dynamics. Second, the term structure of variance implied by single factor models will be fixed. For example, given the price for variance swap maturing in one month, prices for variance swap of other maturities will be also known. The implication is that the correlation term structure of variance derivatives will always be one. But this is not true with the real market data.

Bergomi (2005) shows that extra factors are needed to fit the variance dynamics. Also, Cont and Da Fonseca (2002) claim that a multi-factor model is necessary to capture the volatility surface dynamics. Todorov (2010) shows that risk premium from a variance swap increases when jumps occurs, suggesting one of the risk factors will be strongly associated with realized jumps.

To accommodate all the empirically observed features, we construct the SVSEJ model and link the realized jump to future jump intensity in a self exciting manner. The self exciting jump process is first studied by Hawkes (1971), which is widely used in biology and seismology. In stock markets, negative jumps have a strong impact on the future stock dynamics, so we modify the original model to allow only negative

[1]The volatility surface is a function of the factor state. If the volatility at any point on the surface is known, we can back out the state of the factor, and then calculate other properties of the volatility surface, e.g., skewness.

jumps in equity price to trigger a jump in the jump intensity to capture such an asymmetric effect.

In our SVSEJ model, a jump in the jump intensity not only causes jumps in variance, but also introduce a big skewness in the implied distribution in the short term. The large implied skewness, which we can observe during the crisis, means that investors are willing to pay higher premium for deep OTM put option for protection from further big losses. Mathematically, such a large skewness is very difficult to be fully explained by continuous diffusion component in the underlying dynamics, because skewness generated by continuous part is in order of $(T-t)^2$. When time to maturity is small, diffusion component cannot provide enough skewness to match the observed skewness in the implied volatility surface. On the other hand, the skewness generated by discrete jumps is proportional to $T-t$. If we allow the jump intensity to jump, we effectively introduced jump in both variance and skewness.

Since jump in the jump intensity also causes a jump in total variance, the dynamics of stochastic variance in our model is kept simple. We introduce another Gaussian term, $\sigma_\lambda\sqrt{\lambda_t}dW_{3,t}$, into the dynamics of the jump intensity. The contribution of this term is small when the market is calm. The more important point is that when a jump occurs, jump intensity is excited; this term can help to explain the big fluctuation of the jump intensity afterward.

In similar work, Carr and Wu (2008) and Fulop et al. (2012) use Gamma process to model the jump process and restrict the jump size in return to be proportional to the jump size in the jump intensity. Compared with the compound Poisson process used in our model, Gamma process has jumps in any time interval, and the number of jumps is countable but infinite. So the jump intensity will jump proportionally to the jump size in stock market everyday in such framework. However, we believe that the market response to its past performance is strongly nonlinear depending on the size of the loss. Therefore, we choose compound Poisson process, which has a finite number of jumps, to model such rare events.

In our SVSEJ model, the jump size in stock return and jump intensity is independent. We believe that the movement in both market may be dependent, but not of any simple form due to many reasons, for example, liquidity, trading rules, etc. Especially when the observation frequency is low (daily or weekly), we would expect smaller correlation between the jump sizes in stock market and option market. When there is a big fluctuation in the option market, it is not necessary that we can observe a corresponding movement in the stock market at the daily or weekly basis. One in-

teresting example is the flash crash in May 6, 2010, when the Dow Jones Industrial Average suffered the biggest one day loss in history, but recovered within minutes. The option market responded to the crash and VIX increased by 27.5% on that day and continued to increase by another 22% on the next day. In contrast, the daily returns in the stock market was just -3.3% on that day and -1.3% on the next day.

Compared with the GARJI model in Maheu and McCurdy (2004), the jump intensity in our model is driven by realized jumps, and not the expectation of jump number. Moreover, it only responds to negative jumps and the size of the response is random. Here, both jump process and stochastic variance are stochastic, while in the GARJI model they are deterministic.

5.3 Joint Estimation under Two Measure

5.3.1 Cumulants of Log Return

A distribution can be fully represented by its cumulants (or moments). The result of Edgeworth expansion suggests that the lower cumulants are the most important in explaining the distribution. Here, we choose the second and third cumulants implied by option prices with maturities of two months and three months on each day.[2] To track the market cumulants in the estimation, we need to derive the cumulant implied by the model and the market option data. Since our mode is affine, the cumulant generating function is an affine function of the state variable,[3]

$$\begin{aligned} K(\omega, v_t, \lambda_t, t, T) &= \ln G(\omega, x_t, v_t, \lambda_t, t, T) - \ln e^{x_t \omega} \\ &= A(\omega, t, T) + B(\omega, t, T) v_t + C(\omega, t, T) \lambda_t \end{aligned}$$

So the ith cumulant of log return can be calculated using,

[2]The first cumulant, or the mean, under the risk neutral measure is ignored in the estimation, because the dynamics of it is very similar to the second cumulant, or the variance, and does not carry extra useful information. The fair price of the variance swap is actually a constant times the price of log contract, which is the first cumulant.

[3]The following equation is the cumulant generating function for the log return, not log price, so we need to subtract $\ln e^{x_t \omega}$.

$$\kappa_i = \frac{\partial^i K(\omega, v_t, \lambda_t, t, T)}{\partial \omega^i}|_{\omega=0}$$

which gives,

$$\kappa_i = \frac{\partial^i A(\omega, t, T)}{\partial \omega^i}|_{\omega=0} + \frac{\partial^i B(\omega, t, T)}{\partial \omega^i}|_{\omega=0} v_t + \frac{\partial^i C(\omega, t, T)}{\partial \omega^i}|_{\omega=0} \lambda_t \qquad (5.6)$$

We show in Appendix B how to derive the first cumulant for the log return implied by the model, which involves deriving the derivatives of A, B, and C w.r.t ω.[4] Following the similar procedure, we can take higher order derivatives w.r.t ω for higher order cumulants.

The above analysis describes model implied cumulants. On the other hand, cumulants implied by market data can be constructed by OTM option prices. Details can be found in Appendix C. The results in (5.22) to (5.27) in Appendix B are substituted into equation (5.6) to produce the second and third cumulants. By applying Taylor expansion, we get the approximation for the two cumulants of short term maturities as follows,

$$\begin{aligned} \kappa_2 &\approx v_t(T-t) + \mathbb{E}[J_x^2]\lambda_t(T-t) + O((T-t)^2) \\ \kappa_3 &\approx \mathbb{E}[J_x^3]\lambda_t(T-t) + O((T-t)^2) \end{aligned}$$

It is clear that the diffusion and the jump parts have the same order of contribution to the second cumulants κ_2, which is the variance. However, the third cumulant κ_3, or unnormalised skewness, is dominated by the jump component only. This is a clear indication that high skewness such as that observed during crisis can only be explained by jumps.

5.3.2 Estimating Hidden State Model

The dynamics of our SVSEJ model is driven by a few latent processes. We can rewrite the model into a state-space form and use filtering technique to back out the latent

[4] Although we do not use the first cumulant in the estimation. The first cumulant derived in the appendix is used later to determine the variance swap risk premium.

process. However, the popular Kalman filter is not applicable here as the dynamics is highly non-Gaussian and nonlinear. Therefore, we adopt a particle filter to do the filtering and parameter estimation.

For a time series of length T, the observation on each day is $y_t = \{r_t, \kappa_t^j(\tau_i)\}$, where r_t is the log return defined as $\ln \frac{S_{t+1}}{S_t}$ and $\kappa_t^i(\tau_j)$ is the jth cumulant of log future return of the ith maturity observed at time t, which is calculated from option prices as described in previous section. Therefore, the measurement equations can be written as,

$$\begin{aligned} r_t &= (u - \frac{v_t^2}{2} - \lambda_t(\frac{p}{1-\eta_u} + \frac{1-p}{1+\eta_d} - 1))\Delta t \qquad (5.7) \\ &+ \sqrt{v_t \Delta t} Z_{1,t} + \Sigma_{k=1}^{N_t} J_{k,x} \\ \kappa_t^j(\tau_i) &= A_{i,j} + B_{i,j} v_t + C_{i,j} \lambda_t + \sigma_{i,j} \epsilon_{i,j} \qquad (5.8) \end{aligned}$$

where $N_t \sim \text{Poisson}(\lambda_t \Delta t)$ and $J_{k,x} \sim \pi_x(J_x)$ has a double exponential distribution, $Z_{1,t}$ is a standard normal random variable, and ϵ_{ij} are mutually uncorrelated standard normal random noises.[5] The $A_{i,j}$, $B_{i,j}$, and $C_{i,j}$ can be derived in close form as shown in Appendix B. Based on the Euler approximation, the transition equation that drives the latent process $x_t = \{v_t, \lambda_t\}$ is given as,

$$\begin{aligned} v_t &= v_{t-1} + \kappa_v(\theta_v - v_{t-1})\Delta t + \sigma_v \sqrt{v_{t-1}\Delta t}(\rho Z_{1,t} + \sqrt{1-\rho^2} Z_{2,t}) \qquad (5.9) \\ \lambda_t &= \lambda_{t-1} + \kappa_\lambda(\theta_\lambda - \lambda_{t-1})\Delta t \qquad (5.10) \\ &+ \sigma_\lambda \sqrt{\lambda_{t-1}\Delta t} Z_{3,t} + \Sigma_{k=1}^{N_t} 1(J_{k,x} < 0) J_{k,\lambda} \end{aligned}$$

where $J_{k,\lambda} \sim \pi_\lambda(J_\lambda)$; $Z_{2,t}$ and $Z_{3,t}$ are uncorrelated standard normal random variables; and $1(J_{k,x} < 0)$ is the indicator function which takes value 1 when $J_{k,x} < 0$ and 0 otherwise.

[5]The integral for jump term is expressed as $\int_{\mathbb{R}\times\mathbb{R}^+}(e^{J_x} - 1)(\mu(dJ_x, dJ_\lambda, dt) - \pi_x(J_x)dJ_x\lambda_t dt) = \Sigma_{k=1}^{N_t} J_{k,x} - \lambda_t dt \int_{\mathbb{R}}(e^{J_x} - 1)\pi_x(J_x)dJ_x)$, and $\int_{\mathbb{R}}(e^{J_x} - 1)\pi_x(J_x)dJ_x = (\frac{p}{1-\eta_u} + \frac{1-p}{1+\eta_d} - 1)$ appears in the Δt term in equation (5.7).

5.3.3 Filtering Problem

The difficulty of filtering is due to the fact that the model is highly non-Gaussian and nonlinear, which means that Kalman filter and its variance are not applicable here. Instead, we adopt particle filter, which is basically a Monte Carlo approach to approximate the prediction and filtered distributions. There is an extensive literature on how to make particle filter efficient and accurate. We refer the reader to the book by Doucet et al. (2001) for more details.

Compared with other applications of particle filter, the biggest challenge in our model comes from the jumps in both observation and transition equations, which is supposed to be rare but very important for the dynamics. To make the filtering more efficient, in the standard Sequential Importance Resampling (SIR) framework (see Gordon et al. (1993)), we draw from the Bernoulli distribution for the jump timing at each time step.[6] The cumulants used in the observation equations is the second and the third cumulants for the implied distribution at two and three months. More detailed discussion regarding the choice of cumulants is in Section 5.4. The procedure of filtering is given below,

1. Initialisation, set $t = 0$ and $x_0^i = \{v_0, \lambda_0, J_{x,0}, J_{\lambda,0}\}$ for $i = 1 : M$.

2. Augmentation Step, let $t = t + 1$,

   (a) Prediction of latent process

   $$\{\tilde{v}_t^i, \tilde{\lambda}_t^i\} \quad \sim \quad p(v_t, \lambda_t | v_{t-1}^i, \lambda_{t-1}^i, J_{\lambda,t-1}^i)$$

   where $p(v_t, \lambda_t | v_{t-1}^i, \lambda_{x,t-1}^i, J_{\lambda,t-1}^i)$ is based on transition equations (5.9) and (5.10).

   (b) Using Bernoulli distribution with $p = 0.5$ to sample jumps timing N_t, and record the important weight for each particle as,

   $$\pi_t^i = \frac{\text{Poisson}(N_t^i; \tilde{\lambda}_t^i)}{\text{Bernoulli}(N_t^i; p)}$$
   $$\widetilde{J}_{x,t}^i \sim \pi_x(J_x) \quad \text{if } N_t^i = 1$$

[6] Random variance from Bernoulli distribution only takes value of 1 or 0. Here we ignore the probability of multiple jumps for each day. The result obtained later shows that the jump intensity is small enough to substantiate this assumption.

3. Updating Step,

   (a) Calculate the important weight of particles using,

$$w_t^i = \pi_t^i p(y_t|\tilde{v}_t^i, \tilde{\lambda}_t^i, \widetilde{J}_{x,t}^i)$$

   where $y_t = \{r_t, \kappa_t^2(2 \text{ Months}), \kappa_t^3(2 \text{ Months}), \kappa_t^2(3 \text{ Months}), \kappa_t^3(3 \text{ Months})\}$ and $p(y_t|\tilde{v}_t^i, \tilde{\lambda}_t^i, \widetilde{J}_{x,t}^i)$ is given by the observation equations (5.7) and (5.8).

   (b) Resample the new particles according to the normalized weight,

$$\{v_t^i, \lambda_t^i, J_{x,t}^i\} \sim \text{Multi-nomial}(\frac{w_t^i}{\sum_{i=1}^M w_t^i})$$

   (c) Predict jump size in jump intensity for particles with negative jump size,

$$J_{\lambda,t}^i \sim \pi_\lambda(J_\lambda) \quad \text{if } J_{x,t}^i < 0$$

4. Go to Step 2 until $t = T$.

The number of particles, M, used in the filtering is 2000.

5.3.4 Estimation Method

The estimation is done by the idea of Maximum Likelihood Estimation (MLE). Given the filtered samples, the likelihood is given by,

$$\begin{aligned} L[\Theta|y_{1:T}] &= \prod_{t=1}^{T} p(y_t|y_{1:t-1}, \Theta) \\ &= \prod_{t=1}^{T} \int p(y_t|x_t, \Theta) p(x_t|y_{1:t-1}, \Theta) dx_t \\ &\approx \prod_{t=1}^{T} \sum_{i=1}^{M} w_t^i \end{aligned}$$

where w_t^i is the importance weight from the previous subsection.

However, given the nature of Monte Carlo, the approximated likelihood function

based on Filtered process is not continuous.[7] Therefore, we cannot apply classical optimisation algorithm. Malik and Pitt (2011) proposed a method to approximate the likelihood for a state space model, which is continuous w.r.t to the change of unknown parameters. However, such a method is difficult to use for multidimensional problem, which is our case. Instead, we adopt the Expectation Maximisation (EM) for the estimation. Compared with direct MLE, EM method requires more computation, because it is an iterative method and furthermore it does not only need the filtered process, but also relies on the smoothed particle paths, which is much more time consuming to calculate. Despite these complications, other approaches would require even more computation efforts. Here we explained the EM procedure, which mainly follows Hürzeler and Künsch (1998), for applying on the particle filter and smoother.

1. Filtering: Using the technique explained in Section 5.3.3 to generate M particles according to $p(x_t|y_{1:t}, \Theta)$ at each time step.

2. Smoothing: Smoothing the particle path from $x_{1:T} \sim p(x_{1:T}|y_{1:T}, \Theta)$ backward based on the filtered samples. We explained the procedure briefly as follows,

   (a) Start from $t = T$.

   (b) Sample $x_t^i \sim p(x_t|y_{1:t}, \Theta)$, which we can simply use the smoothed particles from previous step.

   (c) Sample $x_{t-1}^i \sim p(x_{t-1}|x_t^i, y_{1:T}, \Theta) \propto p(x_{t-1}|y_{1:t-1}, \Theta)p(x_t^i|x_{t-1}, \Theta)$. We use the filtered particles at time $t-1$ to represent $p(x_{t-1}|y_{1:t-1}, \Theta)$. The probability $p(x_t^i|x_{t-1}, \Theta)$ can be easily calculated by transition equations (5.9) and (5.10).

   (d) Repeat previous three steps for all the particles $i = 1 : M$.

   (e) Set $t = t - 1$ and repeat from step (b) till $t = 1$.

3. Expectation: Calculate the expectation of log likelihood based on smoothed samples.

$$\begin{aligned} Q(\Theta^*|\Theta) &= \mathbb{E}_{x_{1:T}|y_{1:T},\Theta}\left[\ln p(x_{1:T}, y_{1:T}|\Theta^*)\right] \\ &\approx \frac{1}{M}\sum_{t=1}^{T}\sum_{i=1}^{M}\left(\ln p(x_t^i|x_{t-1}^i, \Theta^*) + \ln p(y_t|x_t^i, \Theta^*)\right) \end{aligned}$$

[7] The discontinuity mainly comes from the resampling step, in which one has to sample from the distribution represented by particles associated with different weights

where $\Theta^* = \Theta$ and $\{x_t^i\}_{t=1}^T$ are smoothed samples from smoothing step in 2.

4. Maximisation: Maximise $Q(\Theta^*|\Theta)$ w.r.t Θ^*, and let $\Theta = \Theta^*$. Then use the new Θ to repeat from step 1 until the parameters converge.

5.4 Empirical Results

5.4.1 Data Description

The data used in this study include S&P 500 daily closing price and the corresponding European options with maturities up to six months. The sample period is from Oct-2007 to Oct-2010. We choose this date range to investigate the interesting period of sub-prime crisis, when the market underwent a turbulent phase. We hope to explain how the market behaves in such a chaotic period. Vanilla option prices are used to construct the cumulants of log return. On the other hand, as shown in equation (5.8), to calculate cumulants, we need option prices with strike prices range from 0 to infinite, which we approximate by interpolation and extrapolation from available option prices using the SVI function in Gatheral (2004). As shown in Appendix C, the higher the moments, the more weights the deep OTM options have, and hence more extrapolation errors are introduced. Taking into account of these pitfalls, we choose to track only the second and the third cumulants, which correspond to the centered second and third moments.

The source of observation noise in equation (5.8) is the interpolation between maturities. In this study, we focus on the short term maturities, specifically 2 and 3 months. It is because the stock market movement has the most impact on the short term options which are also the most liquid. We drop the one month maturity options because the noise from interpolation is too big. In order to get the cumulants at the targeted maturities of interest, we linearly interpolate the two adjacent cumulants from the nearest two available maturities. For one month maturity, one of the cumulants used for interpolation is less than one month. It is well known that options expiring within two weeks are subject to large trading noise which will severely impact on the quality of the cumulants calculated from the option prices. In this study, we exclude all the options that mature within 2 weeks. We further filter out all options with trading volume less than 100 on each day.

In total, the data set consists of 770 daily observations, one market log return

Table 5.1: Data Statistics

	Mean	Standard Deviation
Log Return	-0.0003	0.0192
$\kappa_2(\tau_2)$	0.0170	0.0153
$\kappa_3(\tau_2)$	-0.0048	0.0072
$\kappa_2(\tau_3)$	0.0262	0.0213
$\kappa_3(\tau_3)$	-0.0090	0.0123

and four cumulants. We do not have any identification problem, since there are five measurement equations each day, which is more than the number of transition equations.[8] Table 5.1 reports the daily mean and standard deviation of the log return and cumulants. The mean value of the log return is negative due to the big loss during the sub-prime crisis. We can also see big standard deviation for all the cumulants in that period.

5.4.2 Empirical Findings and Analysis

Table 5.2 reports the parameters estimates and the associated standard errors. Most of the parameters show strong statistical significant because of the support of a large amount of data, except the mean return under the $\mathbb{P}$ measure. It is because mean return only relies on the log return series. Cumulants implied by options, which is under the $\mathbb{Q}$ measure, contains no information about the mean return.

Table 5.2: Parameters Estimates

	Mean(Std Deviation)		Mean(Std Deviation)
u	-0.1106(0.0574)	σ_λ	10.0698(0.0026)
κ_v	2.3910(0.0170)	η	5.5406(0.0002)
θ_v	0.0949(0.0003)	γ_v	-1.1549(0.0096)
σ_v	0.7311(0.0017)	γ_λ	3.7750(0.0009)
ρ	-0.8699(0.0025)	γ_{J_x}	19.1631(0.0034)
η_u	0.0352(0.0001)	γ_{J_λ}	-0.1579(0.0000)
η_d	0.0320(0.0000)	$\sigma_{2,2}$	0.0015(0.0000)
p	0.4543(0.0003)	$\sigma_{2,3}$	0.0010(0.0000)
κ_λ	7.6358(0.0090)	$\sigma_{3,2}$	0.0010(0.0000)
θ_λ	0.0218(0.0002)	$\sigma_{3,3}$	0.0008(0.0000)

This table reports the parameters estimates by MCEM with standard errors in parentheses.

[8]There are two transition equations. But the jump timing is also a latent process which we need to take into account.

For ease comparison, we reported the parameters under two measures in Table 5.3. The relationship of parameters under two measures are given below,

$$\begin{aligned}
\kappa_v^{\mathbb{Q}} &= \kappa_v^{\mathbb{P}} + \gamma_v \sigma_v \\
\theta_v^{\mathbb{Q}} &= \frac{\kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}}{\kappa_v^{\mathbb{Q}}} \\
\eta_u^{\mathbb{Q}} &= \frac{\eta_u^{\mathbb{P}}}{1 + \eta_u^{\mathbb{P}} \gamma_{J_x}} \\
\eta_d^{\mathbb{Q}} &= \frac{\eta_d^{\mathbb{P}}}{1 - \eta_d^{\mathbb{P}} \gamma_{J_x}} \\
p^{\mathbb{Q}} &= \frac{p(1 - \gamma_{J_x} \eta_d^{\mathbb{P}})}{p(1 - \gamma_{J_x} \eta_d^{\mathbb{P}}) + (1-p)(1 + \gamma_{J_x} \eta_u^{\mathbb{P}})} \\
\kappa_\lambda^{\mathbb{Q}} &= \kappa_\lambda^{\mathbb{P}} + \gamma_\lambda \sigma_\lambda \\
\theta_\lambda^{\mathbb{Q}} &= \frac{\kappa_\lambda^{\mathbb{P}} \theta_\lambda^{\mathbb{P}}}{\kappa_\lambda^{\mathbb{Q}}} \\
\eta^{\mathbb{Q}} &= \frac{\eta^{\mathbb{P}}}{1 + \eta^{\mathbb{P}} \gamma_{J_\lambda}}
\end{aligned}$$

and other parameters, ρ, σ_v, and σ_λ, remain the same under two measures.

Table 5.3: Parameters Comparison under Two Measure

	$\mathbb{P}$ Measure	$\mathbb{Q}$ Measure
κ_v	2.3910	1.5467
θ_v	0.0949	0.1467
η_u	0.0352	0.0210
η_d	0.0320	0.0827
p	0.4543	0.1613
κ_λ	7.6358	45.6493
$\varkappa_\lambda$	4.6123	8.5145
θ_λ	0.0218	0.0036
ϑ_λ	0.0361	0.0196
η	5.5406	44.2755

$\varkappa_\lambda = \kappa_\lambda - (1-p)\eta$ and $\vartheta_\lambda = \frac{\kappa_\lambda \theta_\lambda}{\varkappa_\lambda}$ are the mean reverting speed and long run mean of the jump intensity after compensating the random measure.

As in previous studies, we also get negative risk premium γ_v for stochastic variance, which results in a slower mean reverting speed κ_v and a higher long run mean θ_v for stochastic variance under the $\mathbb{Q}$ measure. The leverage effect between stochastic variance and log return is characterised by the correlation ρ. Compared with

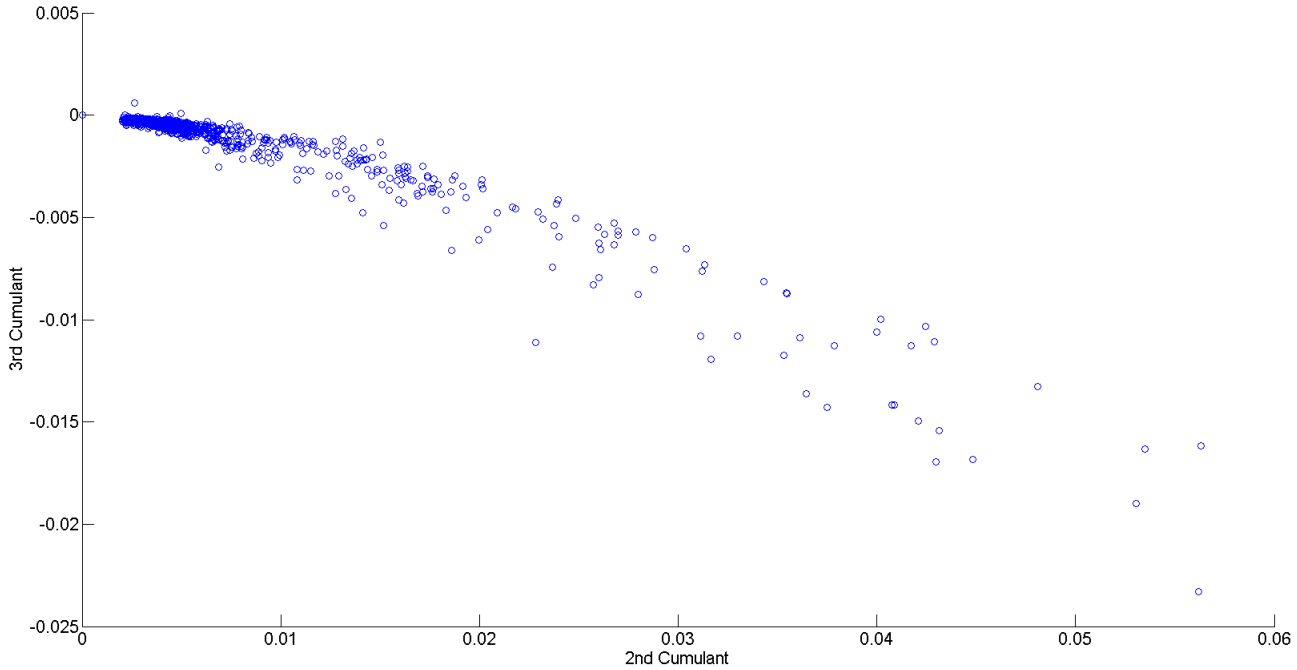
other studies which use only return series for estimation (e.g., Fulop et al. (2012) and Eraker (2004)), in which ρ is around -0.6, our result is more negative. This is because we also use option data in the joint estimation and the correlation implied by option price is typically stronger than those estimated from only return series. Similar results can be found in Bergomi (2004), Duffie et al. (2000), and Carr and Wu (2008).

The most drastic differences between two measures are the jump part. We observe a big negative shift in the distribution of jump size in log returns, causing a much higher skewness implied by option prices. The premium paid for the skewness is further amplified by the huge increase of jump size in jump intensities. An interesting observation is in the change of mean reverting speed of the jump intensity, $\varkappa_\lambda = \kappa_\lambda - (1-p)\eta$ and the long run mean $\vartheta_\lambda = \frac{\kappa_\lambda \theta_\lambda}{\varkappa_\lambda}$. Unlike the stochastic variance, the jump intensity decays faster under the $\mathbb{Q}$ measure than under $\mathbb{P}$ measure. However, the big jump size under $\mathbb{Q}$ measure suggests that investors pay high premium for the big loss and believe that the jump impact has a faster decay rate. In terms of implied volatility surface, the results correspond to a high skewness for short term maturities but flatten out quickly.

Figure 5.2 plots the log return series with the smoothed jump size. The identification of jumps depends on the observed return as well as the instantaneous variance. Therefore, the continuous big fluctuations during the sub-prime crisis are not all identified as jumps. Only the first few days of the crisis period are identified as jumps, when the market has just started to fall into panic. More importantly, the identification of jumps also depends on the information contained in the option prices. For example, there is a big loss in the stock market before October 2008, which was identified as jump. The reason might be that the market was still optimistic and did not change the view about the future movements. Such inconsistency in the two markets is part of the reason why we disentangle the dependency of jump size in log return and jump intensity. In another word, a big loss in stock markets does not necessarily trigger sudden big changes in option markets.

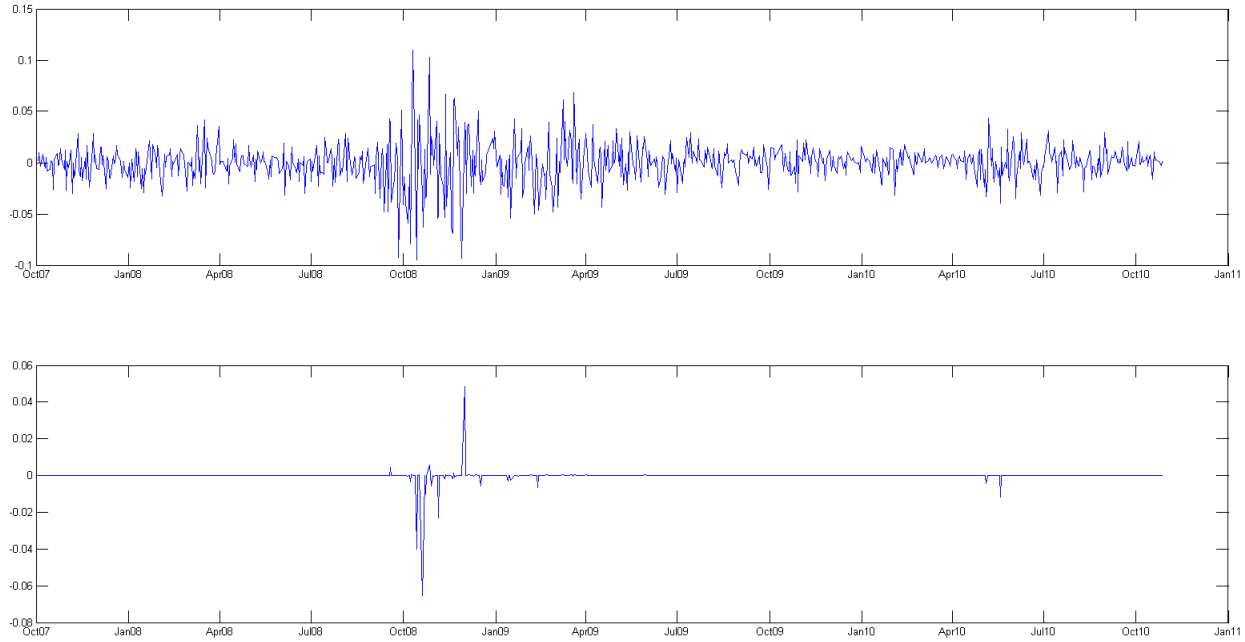
In contrast, Carr and Wu (2008) and Fulop et al. (2012) assume that jump size in return is proportional to the jump size in the jump intensity. Such specification works fine when the estimation uses only option prices or return series but not both. It can help identify the risk premium for the jump size in return if the estimation only use option data; and it help relieve the identification issue when the estimation only use return series, since otherwise jump size in jump intensity has to be identified, which

Figure 5.1: Scatter Plot of Variance against Skewness



This graph shows the scatter plot of 2nd cumulant and 3rd cumulant, which are variance and unnormalised skewness respectively, implied by options maturing in 2 month.

Figure 5.2: Log Return Series and Mean of Jump Size



is very difficult only given the return series. However, since we use both option data and return series in the joint estimation, we can relax this constrain. For the sake of parsimony, we do not impose other dependence structures.

One thing I want to bring to notice is the impact of model choice and data to the estimation of jumps. Comparing with the results in chapter 3, the distribution of jumps differs significantly from that reported in Table 5.3. As mentioned in Section 2.1, since the idea behind the estimation is to match the distribution implied by the model and that observed in the market, the estimation of jumps will depend on the model and data. If the model assumption and market data change, the estimation of jumps will certainly be different.

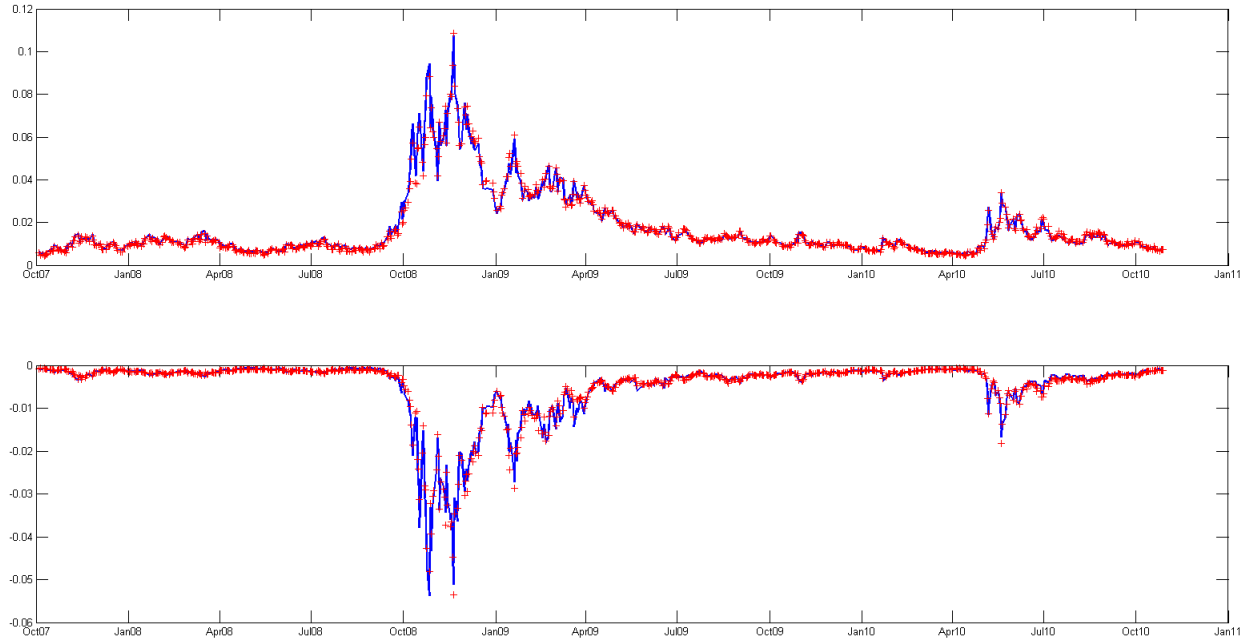
In Figures 5.3 and 5.4, we plot the time series of four cumulants derived from market option prices together with the filtered counterparts by the model. The filtered processes track the market cumulants series well in both calm and crisis periods. Compared with Figure 5.2, we can see that the periods with higher skewness, the third cumulant, started with negative jumps, suggesting a strong relationship between the realized jumps and high skewness during the crisis.

Figure 5.5 plots the time series of filtered stochastic variance and jump intensity. It is clear that stochastic variance explains most of the market dynamics in the calm period while the jump intensity is small and negligible. In the crisis period, jump intensity becomes very prominent. Compared with previous plot, we notice that after triggered by a realized jump at the start of the crisis period, the afterward big fluctuation in jump intensity afterward is well explained by the diffusion part in the jump intensity. Because the diffusion noise in the jump intensity is proportional to $\sqrt{\lambda_t}$; the higher the jump intensity, the bigger the fluctuations. Such movements are not necessarily associated with or driven by jumps in the return series.

To see the different roles played by variance and jump intensity, we plot the contribution to the cumulants from these two factors. Figure 5.6 and 5.7 show the proportion of the cumulants that are explained by jump intensity for maturities of 2 months and 3 months respectively.

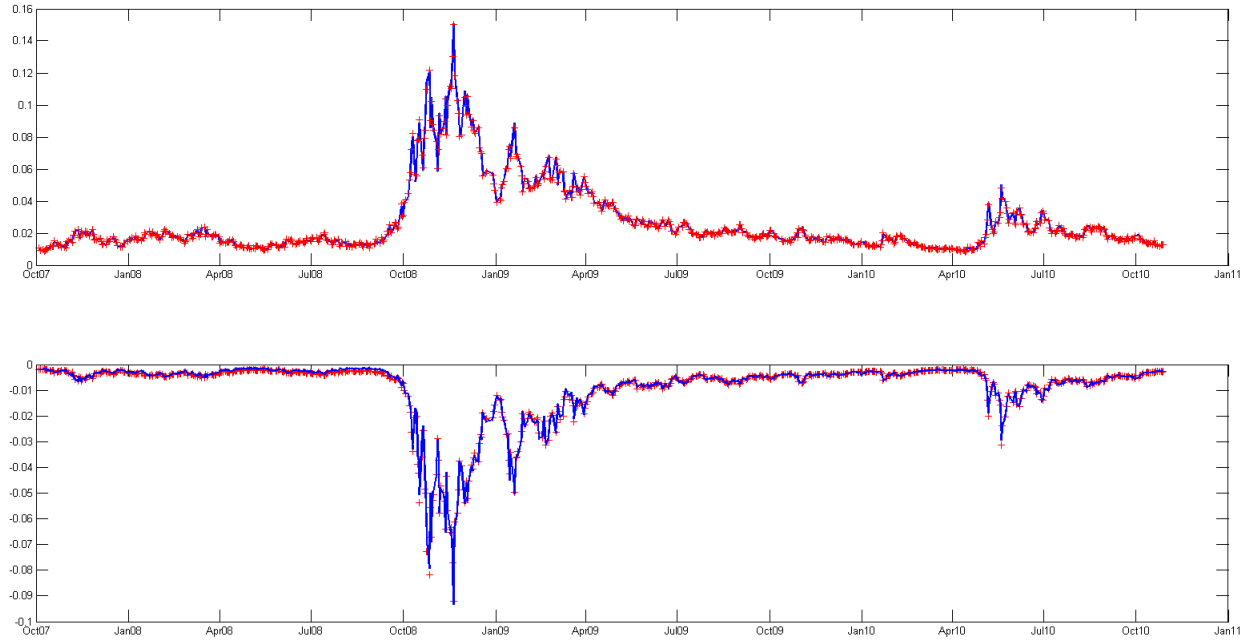
Before the sub-prime crisis, we can see both the second and third cumulants are mainly explained by stochastic variance. However, in the sub-prime crisis, jump intensity explains most of the skewness. This is because that the leverage effect between stochastic variance and log return is not enough to account for the big skewness implied by short maturity options. While for the second cumulant, stochastic variance is always the most important factor. If we do not use the skewness information in

Figure 5.3: Cumulants Series with Maturity of 2 Months



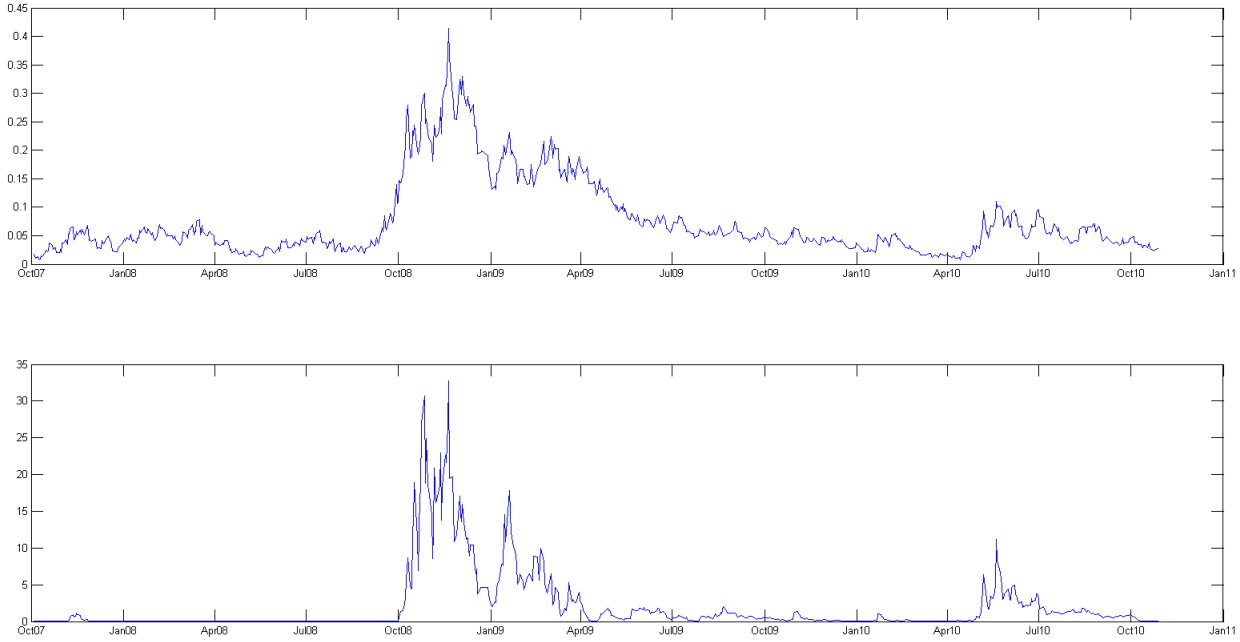
The top panel plots the second cumulant series with maturity of 2 months. The bottom panel plots the third cumulant series with maturity of 2 months. The blue line is the cumulants derived from option data and the red cross is the filtered process by the model.

Figure 5.4: Cumulants Series with Maturity of 3 Months



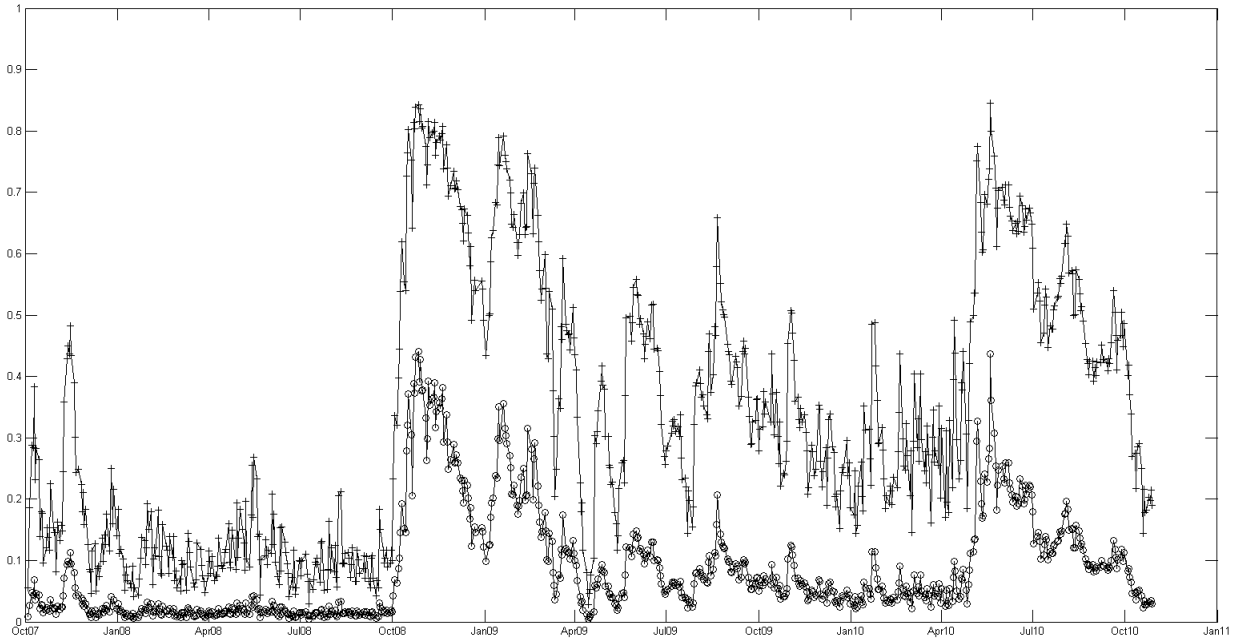
The top panel plots the second cumulant series with maturity of 3 months. The bottom panel plots the third cumulant series with maturity of 3 months. The blue line is the cumulants derived from option data and the red cross is the filtered process by the model.

Figure 5.5: Time Series of Stochastic Variance and Jump Intensity



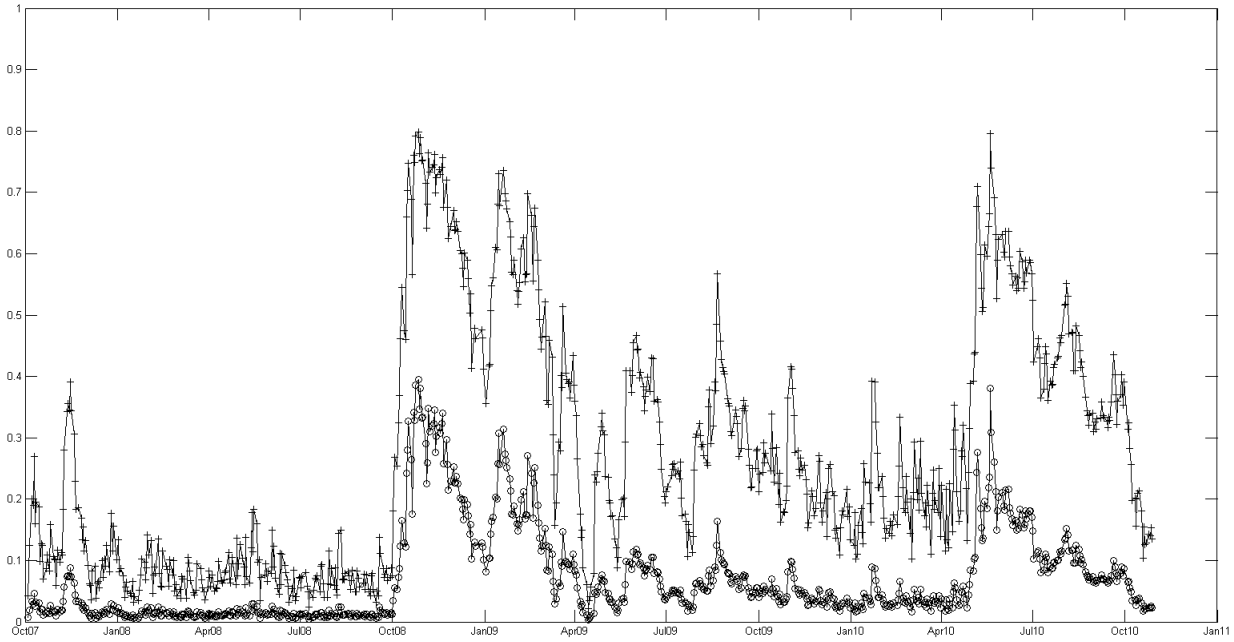
The top panel plots the time series of filtered stochastic variance and the bottom panel plots the time series of filtered stochastic jump intensity.

Figure 5.6: Cumulants Contributed by Jump Intensity for Two Months Maturity



The line marked with round is the percentage of contribution to 2nd cumulant by jump intensity; The line marked with cross is the percentage of contribution to 3rd cumulant by jump intensity.

Figure 5.7: Cumulants Contributed by Jump Intensity for Three Months Maturity



The line marked with round is the percentage of contribution to 2nd cumulant by jump intensity; The line marked with cross is the percentage of contribution to 3rd cumulant by jump intensity.

the estimation, it might not be possible to identify the jump intensity process, since stochastic variance is capable of fully explaining the second cumulant.

5.4.3 Risk Premium in Variance Swap

Since we have the estimates for the risk premium for both stochastic variance and jump intensity, we can revisit the research question, "how the risk premium of variance swap reacts to realized jump?"

Variance swap is a forward contract on variance. Demeterfi et al. (1999) show that the expectation of future variance, or the fair price of a variance swap, is a weighted integral of option prices across all strikes, assuming that there is no jump in equity price. Such findings strongly influence the method used by CBOE to calculate the spot VIX index (see CBOE (2009)), which is defined as the square root of the variance swap. Carr and Wu (2009) establish the relationship between log contract and variance swap with presence of jumps. If admitting jumps, the variance swap calculated by CBOE is,

$$\begin{aligned}\mathbb{E}[\mathrm{VS}^{t,T}|\mathscr{F}_t] &= (T-t)(\theta_v + 2\mathbb{E}[e^{J_x}-1-x]\vartheta) \\ &+ \frac{1-e^{-\kappa_v(T-t)}}{\kappa_v}(v_t-\theta_v) \\ &+ 2\mathbb{E}[e^{J_x}-1-J_x]\frac{1-e^{-\varkappa_\lambda(T-t)}}{\varkappa_\lambda}(\lambda_t-\vartheta_\lambda)\end{aligned} \tag{5.11}$$

From the original definition of variance swap, $\mathbb{E}[\mathrm{VS}^{t,T}|\mathscr{F}_t] = \int_t^T \langle d\ln S_t\rangle$, we have

$$\begin{aligned}\mathbb{E}[\mathrm{VS}^{t,T}|\mathscr{F}_t] &= (T-t)(\theta_v + \mathbb{E}[J_x^2]\vartheta) \\ &+ \frac{1-e^{-\kappa_v(T-t)}}{\kappa_v}(v_t-\theta_v) \\ &+ \mathbb{E}[J_x^2]\frac{1-e^{-\varkappa_\lambda(T-t)}}{\varkappa_\lambda}(\lambda_t-\vartheta_\lambda)\end{aligned} \tag{5.12}$$

The difference between these two definitions are of the third order of the jump size, i.e, $O(\mathbb{E}[J_x^3])$. Here, we will omit the difference, which is very small. In the following part of this section, we will use the second definition of variance swap and calculate the time varying risk premium for the variance swap. Following Todorov (2010), we

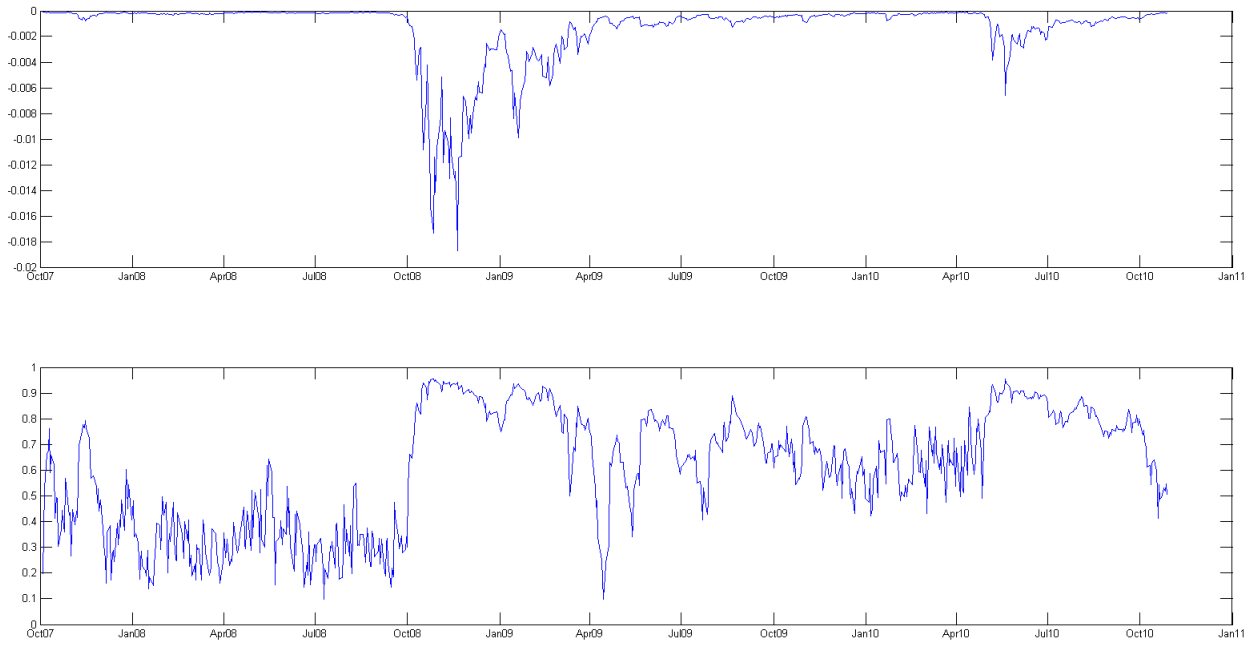
define the variance swap risk premium as the difference between the expectations under two measures,

$$VR_\tau(t) \quad = \quad \mathbb{E}^{\mathbb{P}}[\mathrm{VS}^{t,T}|\mathscr{F}_t] - \mathbb{E}^{\mathbb{Q}}[\mathrm{VS}^{t,T}|\mathscr{F}_t] \tag{5.13}$$

where $\tau = T - t$, and $\mathbb{E}^{\mathbb{P}}[\mathrm{VS}^{t,T}|\mathscr{F}_t]$ and $\mathbb{E}^{\mathbb{Q}}[\mathrm{VS}^{t,T}|\mathscr{F}_t]$ can be calculated by using parameters under two measures reported in Table 5.3 and time series of state variables, v_t and λ_t, in Figure 5.5. Figure 5.8 plots the risk premium associated with one month variance swap. Since the risk premium is also an affine function of the state variables, the risk premium will jump as long as the state variables jumps. Todorov (2010) finds an inconsistent decaying speed of variance swap price after realized jumps. In our two factor specification, we confirm this observation. The premium shoots up from the beginning of the sub-prime crisis, as well as the proportion of the contribution from jump intensity. After the initiate shock in the sub-prime crisis, the percentage of jump intensity contribution decays slows till the next shock due to the Greek debt crisis in 2010. Since the proportion of jump intensity contribution decays after jumps, we expect a slightly faster decaying speed of jump intensity than that of stochastic variance, $\kappa_v < \varkappa_\lambda$, which is consistent with the parameters estimates reported in Table 5.2.

In Figure 5.8, we see that the contribution to the variance swap risk premium by jumps accounts for about 30% of the total risk premium before sub-prime crisis, and more than 90% during the sub-prime crisis. Even after the crisis, it is persistently above 50%. However, as shown in 5.6 and 5.7, the contribution to the second cumulant (variance) by jumps is only below 10% in most time and up to 40% even during the crisis. This observation suggests that for each unit of variance caused by diffusion part and jumps, the risk premium by jumps is much higher, i.e., investors pay a higher risk premium the same amount of variance that comes from jumps. The reason is that, as explained by Bakshi and Madan (2006), the risk premium for variance comes from higher moments, i.e., skewness and kurtosis. As shown in 5.6 and 5.7, jumps contribute much more to the third cumulant (unnormalised skewness) than to the second cumulant (variance), so the risk premium for each unit of variance caused by jumps is much higher.

Figure 5.8: One Month Variance Swap Risk Premium and Contribution by Jump Intensity



Top panel plots the risk premium for variance swap in total; Button panel plots the contribution by jump intensity to the variance swap risk premium.

5.5 Conclusions

We proposed a stochastic variance with self exciting jump (SVSEJ) model for explaining the time varying risk premium of variance swap based on premium embedded in short term maturity option prices. Instead of using option prices directly, we track the cumulants implied by options to distinguish the different roles of stochastic variance and self exciting jump. It is shown that the stochastic variance is able to fit market dynamics very well in normal period and explain the time varying risk premium. However, in the crisis period, the market behaves in a different way. The high skewness of the log return under $\mathbb{Q}$ measure, especially those implied in the short term maturity options, can only be explained by jumps with high jump intensity.

To cater for these two periods with calm or crisis market conditions, we use a Poisson process to capture rare jump events that act as the turning point from calm period to crisis periods. Such assumption is based on what one can observe in the history of stock markets. Jumps are not only triggered by an abrupt big movement in the stock return, but also characterised by the very skewed volatility surface. A jump in our model may be described as the event, which makes investors change their behaviour pattern, i.e, from seeking profit to paying for protection.

The self exciting jump feature in our model explains the high skewness triggered by a realized jump in the crisis period. We introduce another diffusion term in the jump intensity process to describe the big fluctuation of jump intensity in this period, which represents investor's uncertainty against the future jumps in the crisis period. A mean reverting term (with some other conditions) guarantees the stationarity of the process, ensuring the market will revert back to normal condition eventually. The estimation of the model shows that the reverting speeds of the stochastic variance and self exciting jump are different under both measures, suggesting different risk premium dynamics between calm and crisis periods.

In this study, we only focus on the dynamics and risk premium of short maturities. One of the reasons is that we do not have enough information from long term maturity derivatives, which are less liquid. It is possible that we may need extra factors to explain options dynamics with mid or long term maturities. For example, the long run mean of the stochastic variance at long maturity implied by our model is a constant, while we can see big changes day by day of the ATM implied vol of long maturity options. A stochastic long run mean might be a potential third factor if one wants to include long term maturity. Other improvements that can be further done concerns

the estimation method with particle filter, especially for the sampling of jumps in the prediction step. One can also take advantage of parallel computing to improve the optimization efficiency in the EM algorithm.

Appendix

A. Solution of PIDE for Moment Generating Function

We derive the solution form for the moment generating function in equation (5.5), which has the following form,

$$G(\omega, x_t, v_t, \lambda_t, t, T) = e^{\omega x_t + A(\omega,t,T) + B(\omega,t,T)v_t + C(\omega,t,T)\lambda_t} \tag{5.14}$$

Given the initial condition at time T that $G(\omega, x_T, V_T, \lambda_T, T, T) = e^{x_T\omega}$, we have $A(\omega, T, T) = 0$, $B(\omega, T, T) = 0$, and $C(\omega, T, T) = 0$. Substitute this solution form into equation (4.17), we have

$$\begin{aligned}
& -\frac{1}{2}v_t\omega - \lambda_t\mathbb{E}[e^J - 1]\omega + \frac{1}{2}v_t\omega^2 + \kappa(\theta - v_t)B(\omega,t,T) + \frac{1}{2}\sigma^2 v_t B(\omega,t,T)^2 \\
+ \;& \rho\sigma v_t\omega B(\omega,t,T) + \kappa_\lambda(\theta_\lambda - \lambda_t)C(\omega,t,T) \\
+ \;& \lambda_t\int_{\mathbb{R}\times\mathbb{R}^+}(e^{J\omega + 1_{J<0}J_\lambda C(\omega,t,T)} - 1)\nu(dJ, dJ_\lambda) \\
= \;& -\frac{\partial A(\omega,t,T)}{\partial t} - \frac{\partial B(\omega,t,T)}{\partial t}v_t - \frac{\partial C(\omega,t,T)}{\partial t}\lambda_t
\end{aligned}$$

Regroup the equation with respect to the state variables v_t and λ_t,

$$\begin{aligned}
& (-\frac{1}{2}\omega + \omega^2 - \kappa B(\omega,t,T) + \frac{1}{2}\sigma^2 B(\omega,t,T)^2 + \rho\sigma\omega B(\omega,t,T) + \frac{\partial B(\omega,t,T)}{\partial t})v_t \\
+ \;& (-\mathbb{E}[e^J - 1]\omega - \kappa_\lambda C(\omega,t,T) + \int_{\mathbb{R}\times\mathbb{R}^+}(e^{J\omega + 1_{J<0}J_\lambda C(\omega,t,T)} - 1)\nu(dJ, dJ_\lambda) \\
+ \;& \frac{\partial C(\omega,t,T)}{\partial t})\lambda_t + \kappa_v\theta_v B(\omega,t,T) + \kappa_\lambda\theta_\lambda C(\omega,t,T) + \frac{\partial A(\omega,t,T)}{\partial t} \\
= \;& 0
\end{aligned}$$

Since v_t and λ_t are stochastic and not identical, to let the RHS always equals to zero, we have

$$
\begin{aligned}
\frac{\partial A(\omega,t,T)}{\partial t} &= -\kappa_v\theta_v B(\omega,t,T) - \kappa_\lambda\theta_\lambda C(\omega,t,T) && (5.15)\\
\frac{\partial B(\omega,t,T)}{\partial t} &= \frac{1}{2}(\omega-\omega^2) + (\kappa-\rho\sigma\omega)B(\omega,t,T) && (5.16)\\
&\quad -\frac{1}{2}\sigma^2 B(\omega,t,T)^2 \\
\frac{\partial C(\omega,t,T)}{\partial t} &= \kappa_\lambda C(\omega,t,T) \\
&\quad -\int (e^{J_x\omega+1_{J_x<0}J_\lambda C(\omega,t,T)}-1)\nu(dJ_x,dJ_\lambda) && (5.17)\\
&\quad +\mathbb{E}[e^{J_x}-1]\omega && (5.18)
\end{aligned}
$$

Therefore, the PIDE is decomposed as a set of ODEs. The second nonlinear ODE (5.16) is known as the Riccati differential equation, which has analytical solution, while the third ODE (5.17), which corresponds to the self exciting jump intensity, has no analytical solution and has to be solved numerically.[9]

B. Price of Log Contract and Variance Swap

The relationship between moment generating function and log contract is given by,

$$
\begin{aligned}
\mathbb{E}[\ln F_T|\mathscr{F}_t] &= \frac{\partial G(\omega,x_t,v_t,\lambda_t,t,T)}{\partial\omega}|_{\omega=0} \\
&= \ln F_t + \frac{\partial A(\omega,t,T)}{\partial\omega}|_{\omega=0} && (5.19)\\
&\quad +\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0}v_t + \frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0}\lambda_t
\end{aligned}
$$

Taking derivative w.r.t ω on equations (5.15), (5.16), and (5.17), we have,

[9]The numerical method used here is the Runge-Kutta solver.

$$
\begin{aligned}
\frac{\partial^2 A(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= -\kappa_v\theta_v\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} - \kappa_\lambda\theta_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
\frac{\partial^2 B(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= \frac{1}{2} - \rho\sigma B(0,t,T) + \kappa\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
&\quad -\sigma^2 B(0,t,T)\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
\frac{\partial^2 C(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= \kappa_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
&\quad -\int_{\mathbb{R}\times\mathbb{R}^+}(J_x + 1_{J_x<0}J_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0})e^{1_{J_x<0}J_\lambda C(0,t,T)}\nu(dJ_x,dJ_\lambda) \\
&\quad +\mathbb{E}[e^{J_x}-1]
\end{aligned}
$$

Given the initial conditions $A(\omega,T,T)=0$, $B(\omega,T,T)=0$, and $C(\omega,T,T)=0$, we have $A(0,t,T)$, $B(0,t,T)$, and $C(0,t,T)$ are all equal to zero, which simplify the above ODE set as,

$$
\begin{aligned}
\frac{\partial^2 A(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= -\kappa_v\theta_v\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} - \kappa_\lambda\theta_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
\frac{\partial^2 B(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= \frac{1}{2} + \kappa_v\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
\frac{\partial^2 C(\omega,t,T)}{\partial t\partial\omega}|_{\omega=0} &= \kappa_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0} \\
&\quad -\int_{\mathbb{R}\times\mathbb{R}^+}(J_x + 1_{J_x<0}J_\lambda\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0})\nu(dJ_x,dJ_\lambda) + \mathbb{E}[e^{J_x}-1]
\end{aligned}
$$

Taking integral w.r.t t and using the initial conditions, we have,

$$
\begin{aligned}
\frac{\partial A(\omega,t,T)}{\partial\omega}|_{\omega=0} &= \frac{1-(T-t)\kappa_v - e^{-\kappa_v(T-t)}}{2\kappa_v}\theta_v \\
&\quad +\vartheta\frac{1-(T-t)\varkappa_\lambda - e^{-\varkappa_\lambda(T-t)}}{\varkappa_\lambda^2}\kappa_\lambda\theta_\lambda \\
\frac{\partial B(\omega,t,T)}{\partial\omega}|_{\omega=0} &= \frac{e^{-\kappa_v(T-t)}-1}{2\kappa_v} \\
\frac{\partial C(\omega,t,T)}{\partial\omega}|_{\omega=0} &= \xi\frac{e^{-\varkappa_\lambda(T-t)}-1}{\varkappa_\lambda}
\end{aligned}
$$

where $\varkappa_\lambda = \kappa_\lambda - P(J<0)\eta$ and $\xi = \mathbb{E}[e^{J_x}-1-J_x]$. Substituting above results into

equation (5.19), the price of log contract is given by,

$$\begin{aligned}\mathbb{E}[\ln F_T|\mathscr{F}_t] &= \ln F_t \\ &- \frac{1}{2}(T-t)\theta_v - \frac{1-e^{-\kappa_v(T-t)}}{2\kappa_v}(v_t-\theta_v) \\ &- \xi(T-t)\vartheta_\lambda - \xi\frac{1-e^{-\varkappa_\lambda(T-t)}}{\varkappa_\lambda}(\lambda_t-\vartheta_\lambda)\end{aligned} \tag{5.20}$$

where $\vartheta_\lambda = \frac{\kappa_\lambda}{\varkappa_\lambda}\theta_\lambda$ is the long run mean of the jump intensity. Using the relationship between log contract and variance swap, (see Carr and Wu (2009)),

$$\mathbb{E}[\ln F_T|\mathscr{F}_t] = \ln F_t - \frac{1}{2}\mathbb{E}[VS^{t,T}|\mathscr{F}_t] + \int_t^T \int_{\mathbb{R}\times\mathbb{R}^+} (e^{J_x} - 1 - J_x - \frac{J_x^2}{2})\nu(dJ_x, dt)$$

Noticing that $\xi = \mathbb{E}[e^{J_x} - 1 - J_x] = \mathbb{E}[\frac{J_x^2}{2}] + \mathbb{E}[e^{J_x} - 1 - J_x - \frac{J_x^2}{2}]$ and using the assumption that jump timing and size are independent, we substitute above relationship into equation (5.20) to have,

$$\begin{aligned}\mathbb{E}[\mathrm{VS}^{t,T}|\mathscr{F}_t] &= (T-t)(\theta_v + \mathbb{E}[J_x^2]\vartheta_\lambda) \\ &+ \frac{1-e^{-\kappa_v(T-t)}}{\kappa_v}(v_t-\theta_v) \\ &+ \mathbb{E}[J_x^2]\frac{1-e^{-\varkappa_\lambda(T-t)}}{\varkappa_\lambda}(\lambda_t-\vartheta_\lambda)\end{aligned} \tag{5.21}$$

We can see from equation (5.21) that over the time interval from t to T, as the contribution to the variance, diffusion and jump parts have the same function form in terms of their mean reversion rate and long run mean, except that the variance from jumps is scaled by $\mathbb{E}[J_x^2]$, which is the variance for each jump. So using variance swap prices alone is not enough to separate the two state variables in our model.

For the second and third cumulants we tracked in the estimation, we follow the same procedure and derive the close form by using Wolfram Mathematica,

$$
\begin{aligned}
\frac{\partial^2 A(\omega,t,T)}{\partial\omega^2}|_{\omega=0} &= \frac{e^{-2(T-t)\kappa_v}}{8\kappa_v^3}\theta_v((1+4e^{(T-t)\kappa_v}-5e^{2(T-t)\kappa_v})\sigma^2 \qquad (5.22)\\
&+ 2e^{(T-t)\kappa_v}(8(-1+e^{(T-t)\kappa_v})\rho\sigma\\
&+ (2+e^{(T-t)\kappa_v})(T-t)\sigma^2)\kappa_v+8e^{2(T-t)\kappa_v}(T-t)\kappa_v^3\\
&- 8e^{(T-t)\kappa_v}(-1+(T-t)\rho\sigma+e^{(T-t)\kappa_v}(1+(T-t)\rho\sigma))\kappa_v^2)\\
&+ \frac{e^{-2(T-t)\varkappa_\lambda}}{\varkappa_\lambda^3}\vartheta(e^{(T-t)\varkappa_\lambda}\varkappa_\lambda^2(1+e^{(T-t)\varkappa_\lambda}(-1+(T-t)\varkappa_\lambda))J_2\\
&- (-1+p)\eta\xi(\eta(1+4e^{(T-t)\varkappa_\lambda}(1+(T-t)\varkappa_\lambda)\\
&+ e^{2(T-t)\varkappa_\lambda}(-5+2(T-t)\varkappa_\lambda))\xi\\
&+ 2e^{(T-t)\varkappa_\lambda}\varkappa_\lambda(2+(T-t)\varkappa_\lambda+e^{(T-t)\varkappa_\lambda}(-2+(T-t)\varkappa_\lambda))\eta_d))\\
\frac{\partial^2 B(\omega,t,T)}{\partial\omega^2}|_{\omega=0} &= \frac{e^{-2(T-t)\kappa_v}}{4\kappa_v^3}(-\sigma^2+e^{2(T-t)\kappa_v}\sigma^2+4e^{(T-t)\kappa_v}\rho\sigma\kappa_v \qquad (5.23)\\
&- 2e^{(T-t)\kappa_v}(T-t)\sigma^2\kappa_v-4e^{(T-t)\kappa_v}\kappa_v^2+4e^{2(T-t)\kappa_v}\kappa_v^2\\
&+ 4e^{(T-t)\kappa_v}(T-t)\rho\sigma\kappa_v^2-4e^{2(T-t)\kappa_v}\rho\sigma\kappa_v)\\
\frac{\partial^2 C(\omega,t,T)}{\partial\omega^2}|_{\omega=0} &= \frac{e^{-2(T-t)\varkappa_\lambda}}{\varkappa_\lambda^3}(e^{(T-t)\varkappa_\lambda}(-1+e^{(T-t)\varkappa_\lambda})\varkappa_\lambda^2 J_2 \qquad (5.24)\\
&- 2(-1+p)\eta\xi(\eta(-1+e^{2(T-t)\varkappa_\lambda}-2e^{(T-t)\varkappa_\lambda}(T-t)\varkappa_\lambda)\xi\\
&+ e^{(T-t)\varkappa_\lambda}\varkappa_\lambda(-1+e^{(T-t)\varkappa_\lambda}-(T-t)\varkappa_\lambda)\eta_d))
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3 A(\omega,t,T)}{\partial\omega^3}|_{\omega=0} &= -\frac{e^{-3(T-t)\kappa_v}}{16\kappa_v^5}\sigma\theta_v(-(-1-6e^{(T-t)\kappa_v} \\
&- 15e^{2(T-t)\kappa_v}+22e^{3(T-t)\kappa_v})\sigma^3 \\
&+ 6e^{(T-t)\kappa_v}\sigma((-5-8e^{(T-t)\kappa_v}+13e^{2(T-t)\kappa_v})\rho\sigma \\
&+ \sigma((1-8e^{(T-t)\kappa_v}+7e^{2(T-t)\kappa_v})\rho \\
&+ (1+3e^{(T-t)\kappa_v}+e^{2(T-t)\kappa_v})(T-t)\sigma))\kappa_v \\
&- 6e^{(T-t)\kappa_v}(2(-1+(T-t)\rho\sigma)\sigma \\
&+ 2e^{2(T-t)\kappa_v}(12\rho\rho\sigma+5\sigma+2(T-t)\rho\sigma^2 \\
&+ (T-t)\rho\sigma^2)-e^{(T-t)\kappa_v}(24\rho\rho\sigma \\
&+ 8\sigma-12t\rho\sigma\sigma-4t\rho\sigma^2+(T-t)^2\sigma^3))\kappa_v^2 \\
&+ 12e^{2(T-t)\kappa_v}(-(T-t)(-4-2e^{(T-t)\kappa_v}+(T-t)\rho\sigma)\sigma \\
&+ \rho(-8+8t\rho\sigma+4e^{(T-t)\kappa_v}(2+(T-t)\rho\sigma)-(T-t)^2\sigma^2))\kappa_v^3 \\
&- 24e^{2(T-t)\kappa_v}(T-t)\rho(2+2e^{(T-t)\kappa_v}-(T-t)\rho\sigma)\kappa_v^4) \\
&- \frac{e^{-3(T-t)\varkappa_\lambda}}{\varkappa_\lambda^5}\vartheta(-e^{2(T-t)\varkappa_\lambda}\varkappa_\lambda^4(1+e^{(T-t)\varkappa_\lambda}(-1+(T-t)\varkappa_\lambda))J_3 \\
&- 3e^{(T-t)\varkappa_\lambda}(-1+p)\eta\kappa^2 J_2(\eta(1+4e^{(T-t)\varkappa_\lambda}(1+(T-t)\varkappa_\lambda) \\
&+ e^{2(T-t)\varkappa_\lambda}(-5+2(T-t)\varkappa_\lambda))\xi+e^{(T-t)\varkappa_\lambda}\varkappa_\lambda(2+(T-t)\varkappa_\lambda \\
&+ e^{(T-t)\varkappa_\lambda}(-2+(T-t)\varkappa_\lambda))\eta_d) \\
&+ (-1+p)\eta\xi(\eta^2(-\varkappa_\lambda(-1+9e^{(T-t)\varkappa_\lambda}+9e^{2(T-t)\varkappa_\lambda}(1+2(T-t)\varkappa_\lambda) \\
&+ e^{3(T-t)\varkappa_\lambda}(-17+6t\kappa))+2(-1+p)\eta(1+6e^{(T-t)\varkappa_\lambda}(1+(T-t)\varkappa_\lambda) \\
&+ e^{3(T-t)\varkappa_\lambda}(-22+6t\kappa)+3e^{2(T-t)\varkappa_\lambda}(5+6t\kappa+2(T-t)^2\varkappa_\lambda^2)))\xi^2 \\
&- 3e^{(T-t)\varkappa_\lambda}\eta\kappa(\varkappa_\lambda(1+4e^{(T-t)\varkappa_\lambda}(1+(T-t)\varkappa_\lambda) \\
&+ e^{2(T-t)\varkappa_\lambda}(-5+2(T-t)\varkappa_\lambda)) \\
&- 2(-1+p)\eta(2+(T-t)\varkappa_\lambda+2e^{(T-t)\varkappa_\lambda}(2+(T-t)\varkappa_\lambda)^2 \\
&+ e^{2(T-t)\varkappa_\lambda}(-10+3(T-t)\varkappa_\lambda)))\xi\eta_d \\
&- 3e^{2(T-t)\varkappa_\lambda}\varkappa_\lambda^2(-(-1+p)\eta(6+4t\kappa+(T-t)^2\varkappa_\lambda^2 \\
&+ 2e^{(T-t)\varkappa_\lambda}(-3+(T-t)\varkappa_\lambda)) \\
&+ 2\kappa(2+(T-t)\varkappa_\lambda+e^{(T-t)\varkappa_\lambda}(-2+(T-t)\varkappa_\lambda)))\eta_d^2))
\end{aligned}
\tag{5.25}
$$

$$
\begin{aligned}
\frac{\partial^3 B(\omega,t,T)}{\partial\omega^3}|_{\omega=0} &= \frac{3e^{-3(T-t)\kappa_v}}{16\kappa_v^5}\sigma(-(-1-2e^{(T-t)\kappa_v}+e^{2(T-t)\kappa_v} \\
&+ 2e^{3(T-t)\kappa_v})\sigma^3 \\
&+ 2e^{(T-t)\kappa_v}\sigma(4(-2+e^{(T-t)\kappa_v}+e^{2(T-t)\kappa_v})\rho\sigma \\
&+ \sigma(2(-1+e^{(T-t)\kappa_v})^2\rho+(2+e^{(T-t)\kappa_v})(T-t)\sigma))\kappa_v \\
&- 2e^{(T-t)\kappa_v}(4(-1+(T-t)\rho\sigma)\sigma+4e^{2(T-t)\kappa_v}(2\rho\rho\sigma+\sigma) \\
&- e^{(T-t)\kappa_v}(8\rho\rho\sigma-8t\rho\sigma\sigma+(T-t)^2\sigma^3))\kappa_v^2 \\
&+ 4e^{2(T-t)\kappa_v}((T-t)(4-(T-t)\rho\sigma)\sigma \\
&+ \rho(-4+4e^{(T-t)\kappa_v}+4t\rho\sigma-(T-t)^2\sigma^2))\kappa_v^3 \\
&+ 8e^{2(T-t)\kappa_v}(T-t)\rho(-2+(T-t)\rho\sigma)\kappa_v^4)
\end{aligned} \tag{5.26}
$$

$$
\begin{aligned}
\frac{\partial^3 C(\omega,t,T)}{\partial\omega^3}|_{\omega=0} &= \frac{e^{-3(T-t)\varkappa_\lambda}}{\varkappa_\lambda^5}(e^{2(T-t)\varkappa_\lambda}(-1+e^{(T-t)\varkappa_\lambda})\varkappa_\lambda^4 J_3 \\
&+ 3e^{(T-t)\varkappa_\lambda}(-1+p)\eta\kappa^2 J_2(2\eta(-1+e^{2(T-t)\varkappa_\lambda} \\
&- 2e^{(T-t)\varkappa_\lambda}(T-t)\varkappa_\lambda)\xi \\
&+ e^{(T-t)\varkappa_\lambda}\varkappa_\lambda(-1+e^{(T-t)\varkappa_\lambda}-(T-t)\varkappa_\lambda)\eta_d) \\
&- 3(-1+p)\eta\xi(\eta^2(\varkappa_\lambda(-1+6e^{(T-t)\varkappa_\lambda}-2e^{3(T-t)\varkappa_\lambda} \\
&+ e^{2(T-t)\varkappa_\lambda}(-3+6(T-t)\kappa)) \\
&+ 2(-1+p)\eta(-1+2e^{3(T-t)\varkappa_\lambda}-2e^{(T-t)\varkappa_\lambda}(1+2(T-t)\varkappa_\lambda) \\
&+ e^{2(T-t)\varkappa_\lambda}(1-2(T-t)\varkappa_\lambda-2(T-t)^2\varkappa_\lambda^2)))\xi^2 \\
&- 2e^{(T-t)\varkappa_\lambda}\eta\kappa(-\varkappa_\lambda(1-e^{2(T-t)\varkappa_\lambda}+2e^{(T-t)\varkappa_\lambda}(T-t)\varkappa_\lambda) \\
&- (-1+p)\eta(-3+3e^{2(T-t)\varkappa_\lambda}-2(T-t)\varkappa_\lambda \\
&- 2e^{(T-t)\varkappa_\lambda}(T-t)\varkappa_\lambda(2+(T-t)\varkappa_\lambda)))\xi\eta_d \\
&- e^{2(T-t)\varkappa_\lambda}\varkappa_\lambda^2(-2\kappa(1-e^{(T-t)\varkappa_\lambda}+(T-t)\varkappa_\lambda) \\
&- (-1+p)\eta(-2+2e^{(T-t)\varkappa_\lambda}-2(T-t)\varkappa_\lambda-(T-t)^2\varkappa_\lambda^2))\eta_d^2))
\end{aligned} \tag{5.27}
$$

where $J_2 = \mathbb{E}[J_x^2]$ and $J_3 = \mathbb{E}[J_x^3]$.

C. Cumulants derived from Option Prices

As shown in Carr et al. (1998), for any twice differentiable function $f(F_T)$ of the final future price F_T,

$$
\begin{aligned}
f(F_T) &= f(F_t) + f'(F_t)(F_T - F_t) \\
&+ \int_0^{F_t} f''(K)(K - F_T)^+ dK + \int_{F_t}^{\infty} f''(K)(F_T - K)^+ dK
\end{aligned}
$$

Let $f(\cdot) = \left(\frac{\ln \cdot}{\ln F_t}\right)^i$, where i denotes moment order, and taking expectation on both sides, we can calculate the ith moment of the log return,

$$
\mathbb{E}[f(F_T)] \quad + \quad \int_0^{F_t} f''(K)\text{Put}(K)dK + \int_{F_t}^{\infty} f''(K)\text{Call}(K)dK \tag{5.28}
$$

where $\text{Call}(K)$ and $\text{Put}(K)$ are option prices at strike price K without discounting. So, for the first three moments, we have,

$$
\begin{aligned}
m_1\left(\frac{\ln F_T}{\ln F_t}\right) &= -\int_0^{F_t} \frac{1}{K^2}\text{Put}(K)dK - \int_{F_t}^{\infty} \frac{1}{K^2}\text{Call}(K)dK \\
m_2\left(\frac{\ln F_T}{\ln F_t}\right) &= \int_0^{F_t} \frac{2}{K^2}\left(1 - \ln\frac{K}{F_t}\right)\text{Put}(K)dK \\
&+ \int_{F_t}^{\infty} \left(1 - \ln\frac{K}{F_t}\right)\text{Call}(K)dK \\
m_3\left(\frac{\ln F_T}{\ln F_t}\right) &= \int_0^{F_t} \frac{3}{K^2}\left(2\ln\frac{K}{F_t} - \ln^2\frac{K}{F_t}\right)\text{Put}(K)dK \\
&+ \int_{F_t}^{\infty} \frac{3}{K^2}\left(2\ln\frac{K}{F_t} - \ln^2\frac{K}{F_t}\right)\text{Call}(K)dK
\end{aligned}
$$

To convert moments to cumulants, we have,

$$
\begin{aligned}
\kappa_1 &= m_1 \\
\kappa_2 &= m_2 - m_1^2 \\
\kappa_3 &= m_3 - 3m_2m_1 + 2m_1^3
\end{aligned}
$$

Chapter 6

Consistent Pricing and Hedging with Two Surfaces

6.1 Introduction

The Volatility Index (VIX) was introduced by the Chicago Board Options Exchange (CBOE) in 1993, and has been viewed as a measure for the short term volatility in the stock market. In 2003, CBOE switched to a new method to calculate VIX based on a portfolio of options on the S&P 500 index (SPX). The new method makes it possible to replicate the VIX index payoff, generating a huge interest to trade and hedge the invisible volatility.

Shortly after the new VIX was adopted, future contracts written on VIX became available in 2004, and options on VIX started trading in 2006. Derivatives on VIX have become increasingly popular and the average daily trading volume of both options and futures has increased by about 20 times from 2006 to 2011. VIX derivatives are not only a hedging instrument as they were first created, but also became a new asset class for investment and speculation. Such a change in market practice also reflects the development of pricing methods.

The earlier papers treat volatility as a separate underlying and price its derivatives in isolation from the stock price process. Grunbichler and Longstaff (1996) assume volatility follows a CIR process and derive the price of futures and options based on the close form of the transition distribution. Psychoyios et al. (2010) include jumps in the CIR process to model VIX dynamics directly. The presence of jump improves the fit to VIX option prices significantly, since the original CIR process does not

generate enough positive skewness in the implied volatility of VIX options. The cost of adding jumps is that the transition distribution of the underlying spot VIX is no longer available, but, the characteristic function of VIX distribution is still available in close form, which gives semi-close forms for the VIX future and option prices.

Another category of models does not treat VIX as a separately asset, but first derive it from the dynamics of SPX index, and then derive prices of VIX derivatives. Zhang and Zhu (2006) use Heston model to derive the dynamics of VIX future. Sepp (2008) adds jumps to the CIR process for stochastic variance and shows how to price futures and options on VIX. Some recent works, e.g., Lu and Zhu (2009) use multi-factor model to study the term structure of VIX futures.

This second approach of VIX modelling has becomes more popular in recent years. Since such method prices derivatives on SPX and VIX consistently, one can measure the sensitivity of equity derivatives w.r.t the change of VIX. This is not possible if VIX is treated as a separate asset. We call this type of joint modelling and pricing the consistent pricing approach.

In this chapter, we adopt the consistent pricing approach, but unlike previous studies, which focus on static calibration of the VIX surface,[1] we also study the term structure of VIX futures dynamics. Specifically, we are interested in the correlation term structure of VIX futures, which is an important measure for hedging. In addition, we also show how to calculate the exposure of VIX option w.r.t VIX future.

The rest of this chapter is organized as follows. A general pricing framework based on characteristic function is introduced briefly in Section 6.2. In Section 6.3, we show how to derive the characteristic function for both VIX and SPX consistently based on affine models. Three affine models, one one-factor model and two multi-factor models are introduced in this section. Correlation term structure implied by different models is calculated in Section 6.4. In this section, we start with the correlation term structure of VIX future, which is conditional on the state variables (or factors). We show that the correlation term structure implied by one factor model is unrealistic. For the multi-factor model, we show how to calculate the unconditional correlation term structure, which can be used to calibrate to market data. The hedge ratio for a VIX option w.r.t VIX future is also derived in this section. Finally, a brief discussion and conclusion are provided in Section 6.5.

[1] In the rest part of this chapter, we use VIX surface to denote the implied volatility surface of VIX options

6.2 Consistent Pricing Method

In finance, the pricing of contingent claims is normally done by taking expectation under the risk neutral measure. In many cases, the distribution function for underlying state is not available, while the characteristic function still exists. Since the characteristic function is an alternative way to represent the distribution, one can price contingent claims by using the characteristic function instead. The key is to find a Fourier transform for the payoff function. Sepp (2003) reviews the class of affine model, and shows the general approach for deriving the characteristic function and using it to price contingent claim. This general pricing framework cannot be applied to some other non-affine models. For example, the derivation for the characteristic function for the 3/2 model discussed in Carr and Sun (2007) is more complicated.

In the following, we will first briefly explain how to use the characteristic function to price SPX derivatives and then extend the pricing approach to price VIX futures and options.

6.2.1 Derivatives on Equity

For a European option, the current price is the discounted expectation of future payoff under the risk neutral measure,

$$\begin{aligned} f(X_t, t) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[f(X_T, T)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\int_{\mathbb{R}^n} f(X_T, T)p(X_T|\mathcal{F}_t)dX_T \end{aligned} \tag{6.1}$$

where $X_t = \ln S_t$. Let us denote $G(\omega, X_t, T)$ as the characteristic function of the transition distribution $p(X_T|X_t)$, which is defined by the generalised Fourier Transform,

$$\begin{aligned} G(\omega, X_t, T) &= \int_{-\infty}^{\infty} e^{X_T\omega}p(X_T|\mathcal{F}_t)dX_T \\ p(X_T|\mathcal{F}_t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{X_T\omega}G(\omega, X_t, T)d\Im(\omega) \end{aligned}$$

where ω is a complex number and $\Im(\cdot)$ denotes the imaginary part.[2]

[2]The notation in this chapter is consistent with chapter 4.

So if we define $\hat{f}(\omega)$ the Fourier transform of the payoff function $f(X_T, T)$, or alternatively $f(X_T, T)$ is the inverse Fourier transform of $\hat{f}(\omega)$. Equation (6.1) can be rewritten as,

$$\begin{aligned} f(X_t, t) &= \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-\omega X_T} d\omega p(X_T|\mathcal{F}_t) dX_T \\ &= \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) \int_{\mathbb{R}} e^{-\omega X_T} p(X_T|\mathcal{F}_t) dX_T d\omega \\ &= \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) G(-\omega, X_t, T) d\omega \end{aligned} \tag{6.2}$$

Since the expression of $\hat{f}(\omega)$ and $G(-\omega, X_t, T)$ are normally known, we can evaluate the integral in equation (6.2) numerically. Many efficient methods are developed to perform this integration efficiently, for example, FFT (Fast Fourier Transform) by Carr and Madan (1999) and CFT (Cosine Fourier Transform) by Fang and Oosterlee (2008).

6.2.2 Pricing of VIX Future and Option

The pricing for variance derivatives is not as straightforward as the pricing of equity derivatives, since the payoff is not associated directly with the variables in the model dynamics. For derivatives such as VIX futures and options, the underlying is the expectation of continuous accumulated variance. The model we usually use to price options on equity price only track instantaneous variance, not the accumulated variance. To price derivatives on both expected accumulated variance and equity price consistently, we need to derive the expected accumulated variance as a function of the state variables.

Assuming the dynamics of equity price with time varying variance and jump intensity as follows,

$$d \ln S_t = (r - \frac{v_t}{2} - \lambda m_1) dt + \sqrt{v_t} dW_t + \int_{\mathbb{R} \times \mathbb{R}_+} J_x \mu(dJ_x, dt)$$

where $m_1 = \mathbb{E}[e^{J_x} - 1]$ and $\mu(dJ_x, dt)$ is the random measure for the compound Poisson jump J_x with jump intensity λ_t. So the accumulated variance over the period from t

to $t + \Delta t$ is given by,

$$V = \int_t^{t+\Delta t} \left(v_s ds + \int_{-\infty}^{\infty} J_x^2 \mu(dJ_x, ds) \right) \tag{6.3}$$

where Δt is usually set as 30 days for contracts like variance swaps and VIX futures. By assuming no dependance between jump timing and size, we have $\nu(dJ_x, dt) = \nu(dJ_x)\lambda_t dt$, where $\nu(dJ_x, dt)$ is the corresponding compensator and $\nu(dJ_x)$ is the probability density function of jump size. Taking expectation and annualising on equation (6.3), we have,

$$\begin{aligned} \bar{V}(v_t, \lambda_t) &= \frac{1}{\Delta t}\mathbb{E}[V|\mathcal{F}_t] \\ &= \mathbb{E}\left[\int_t^{t+\Delta t} \left(v_s + \lambda_s \int_{-\infty}^{\infty} J_x^2 \nu(dJ_x) \right) ds \right] \\ &= \mathbb{E}\left[\int_t^{t+\Delta t} \left(v_s + \lambda_s \mathbb{E}[J_x^2] \right) ds \right] \end{aligned} \tag{6.4}$$

which $\bar{V}(v_t, \lambda_t)$ is the price for an annualized Δt variance swap starting at time t. For many stochastic variance and stochastic jump intensity models, e.g., Heston and Bates models, there is close form solution. Based on the above definition, one can further define VIX futures and options,

$$\begin{aligned} F_{VIX}(t,T) &= \mathbb{E}^{\mathbb{Q}}[\sqrt{\bar{V}(v_T, \lambda_T)}|\mathcal{F}_t] \\ C_{VIX}(t,T) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\left(\sqrt{\bar{V}(v_T, \lambda_T)} - K \right)^+ |\mathcal{F}_t] \end{aligned}$$

which are priced at time t, but settled at time T.[3] Therefore, we can derive the VIX future and option as the expectation of the respective payoff functions under the joint distribution of $p(v_T, \lambda_T|\mathcal{F}_t)$. Following the approach described in Section 6.2.1, we need to derive the Fourier Transform of $\sqrt{\bar{V}(v_T, \lambda_T)}$ and $\left(\sqrt{\bar{V}(v_T, \lambda_T)} - K \right)^+$, and also the characteristic function of $\bar{V}(v_T, \lambda_T)$.

Before we proceed, it is important to point out a subtlety of the VIX definition. VIX, as reported by CBOE, is based on a replicating strategy involving vanilla op-

[3]Noting that for VIX future, the reference period is from T to $T + \Delta t$, observed at t.

tions. Under the assumption of no jump, the derivation is perfect. However, when there are jumps, the replicated variance is different from equation (6.3) . As shown in Carr and Wu (2009), the VIX derived from options, which is reported by CBOE, is,

$$\bar{V} = \frac{1}{\Delta t}\mathbb{E}\left[\int_t^{t+\Delta t}\left(v_s + \lambda_s \int_{-\infty}^{\infty} 2\left(e^{J_x} - 1 - J_x\right)\nu(dJ_x)\right)ds\right] \tag{6.5}$$

The difference between above equation and equation (6.3) is of the third order of jump size, since[4]

$$\begin{aligned}\int_{-\infty}^{\infty} 2\left(e^{J_x} - 1 - J_x\right)\nu(dJ_x) &= 2\mathbb{E}[e^{J_x} - 1 - J_x] \\ &= \mathbb{E}[J_x^2] + O(J_x^3)\end{aligned}$$

Although $O(J_x^3)$ is small and trivial compared with $\mathbb{E}[J_x^2]$, to be consistent with the CBOE definition of the underlying, we will use equation (6.5) as the definition of VIX for the rest of the chapter. And we denote $m_2 = 2\mathbb{E}[e^{J_x} - 1 - J_x]$.

6.3 Characteristic Function of Equity Price and Variance

This chapter develops a framework for consistent pricing of equity and volatility derivatives based on characteristic functions. We show how the closed form solutions for derivative prices are derived for three models; a single factor model with jumps in both stock price and variance, a two-factor model proposed in the previous chapter, where the jump intensity follows a self exciting jump process with stochastic variance, and another two-factor model with two Heston type stochastic variance, one of which share the jump component with the stock price. The number of factors here refer to the factors that drive the variance. The characteristic functions of the factors and equity price are presented as the solution to a Partial Integral Differential Equation (PIDE), and then we show how to solve such PIDE.

[4]Using Taylor expansion, $e^x = 1 + x + \frac{x^2}{2} + O(x^3)$.

6.3.1 One-Factor Model

As pointed out by Sepp (2008) and Psychoyios et al. (2010), a positive jump is needed in the variance process to generate the positive skewness observed in the VIX volatility surface. So we use the following specification,

$$\begin{aligned} d\ln S_t &= (r - \frac{v_t}{2} - \lambda m_1)dt + \sqrt{v_t}dW_t^1 + \int J_x \mu(dJ_x, dJ_v, dt) \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2 + \int J_v \mu(dJ_x, dJ_v, dt) \\ J_x &\sim \text{Normal}(\epsilon + \rho_J J_v, \delta^2) \\ J_v &\sim \text{Exponential}(\eta) \end{aligned} \tag{6.6}$$

where $m_1 = \mathbb{E}[e^{J_x} - 1] = \frac{1}{1-\rho_J\eta}e^{\epsilon+\delta^2/2} - 1$, $\mu(dJ_x, dJ_v, dt)$ is the random measure for jumps in both price and volatility which have the same jump timing, λ is the jump intensity, and ρ_J defines the dependence between jump in log price and variance. The model above is studied in by Bates (2000), Pan (2002), and Eraker (2004), and nests many other well known stochastic volatility models. For example, it is the Heston model if there are no jumps. If jumps are only allowed in equity price but not in variance, the model becomes Bates (1996). However, we should note that despite the generality of the model, it is still a one factor model.

First, following equation (4.32), the prices of VIX future and option can be written as,

$$\begin{aligned} F_{VIX}(t,T) &= \mathbb{E}[\sqrt{\bar{V}(v_T)}|\mathcal{F}_t] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sqrt{\pi}}{2(-\nu)^{3/2}}\phi_{\bar{V}}(-\nu, t, T)d\Im(\nu) \end{aligned} \tag{6.7}$$

$$\begin{aligned} C_{VIX}(t,T) &= e^{-r(T-t)}\mathbb{E}[\left(\sqrt{\bar{V}(v_T)} - K\right)^+ |\mathcal{F}_t] \\ &= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty}^{\infty} \frac{1 - \text{erf}(K\sqrt{-\nu})}{2(-\nu)^{3/2}}\phi_{\bar{V}}(-\nu, t, T)d\Im(\nu) \end{aligned} \tag{6.8}$$

where $\phi_{\bar{V}}(-\nu, t, T)$ is the characteristic function of the expectation of accumulated variance $\bar{V}(v_T)$, erf is the error function, and $\frac{\sqrt{\pi}}{2(-\nu)^{3/2}}$ and $\frac{1-\text{erf}(K\sqrt{-\nu})}{2(-\nu)^{3/2}}$ are the characteristic functions of the payoff of VIX future and option respectively, which are

directly derived from the definition of characteristic function. To derive $\phi_{\bar{V}}(-\nu, t, T)$, we write $\bar{V}$ as a function of v_T. Given equation (6.5), we have,

$$\begin{aligned}\bar{V}(v_T) &= \left(\frac{e^{-\kappa\Delta t}-1}{\kappa\Delta t}(\frac{\kappa\theta+\lambda\eta}{\kappa}-v_T)+\frac{\kappa\theta+\lambda\eta}{\kappa}\right)+\lambda m_2 \\ &= C_0+C_1 v_T\end{aligned}$$

where $m_2 = 2\mathbb{E}^{\mathbb{Q}}[e^{J_x}-1-J_x] = 2\left(\frac{1}{1-\rho_J\eta}e^{\epsilon+\delta^2/2}-1-\epsilon-\rho_J\eta\right)$ and $C_0 = \lambda m_2 + \frac{\kappa\theta+\lambda\eta}{\kappa\Delta t}(\frac{e^{-\kappa\Delta t}-1}{\kappa}+\Delta t)$, $C_1 = -\frac{(e^{-\kappa\Delta t}-1)}{\kappa\Delta t}$. We notice that the expectation of accumulated variance is an affine function of v_T. If we know the characteristic function of v_T, it is straightforward to derive that of $\bar{V}$. On the other hand, because we are interested in consistent pricing of derivatives on both equity and variance, we need to derive the joint characteristic function for S_T and v_T. The characteristic function of the joint distribution of states is defined as $G(\omega_0, \omega_1, t, T) = \mathbb{E}[e^{\ln S_T\omega_0+v_T\omega_1}|\mathcal{F}_t]$, and due to Feynman-Kac formula, it is the solution to the PIDE,

$$\begin{aligned}G_t+(r-\frac{v_t}{2}-\lambda m_1)G_x+\frac{1}{2}v_tG_{xx} &+ \\ \kappa(\theta-v_t)G_v+\frac{1}{2}\sigma^2 v_tG_{vv}+\rho\sigma v_tG_{xv} &+ \\ \lambda\int_{\mathbb{R}\times\mathbb{R}^+}\delta G\nu(dJ_x,dJ_v) &= 0 \end{aligned} \tag{6.9}$$

where $x = \ln S_t$, $\delta G = G(x+J_x, v+J_v) - G(x,v)$ and $\nu(dJ_x, dJ_v)\lambda dt$ is the compensator for the random measure $\mu(dJ_x, dJ_v, dt)$, and the initial condition at $t = T$ is given by,

$$G(\omega_0, \omega_1, T, T) = e^{\omega_0 \ln S_T+\omega_1 v_T} \tag{6.10}$$

As shown in Sepp (2003), the solution for the PIDE (6.9) takes the form of $\exp(\omega_0 \ln S_t + r(T-t)\omega_0 + A(\omega_0, \omega_1, t, T) + B(\omega_0, \omega_1, t, T)v_t)$. So the integral in the

last term of equation (6.9) becomes,

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}^+} \delta G\nu(dJ_x, dJ_v) &= \int_{\mathbb{R}\times\mathbb{R}^+} (G(x+J_x, v+J_v) - G(x,v))\nu(dJ_x, dJ_v) \\ &= G(x,v)\int (e^{\omega_0 J_x + BJ_V} - 1)\nu(dJ_x, dJ_v) \\ &= G(x,v)\left(\frac{e^{\epsilon\omega_0 + \delta^2\omega_0^2/2}}{1-\eta(B+\rho_J\omega_0)} - 1\right)\end{aligned}$$

Substituting the proposed solution form into the PDE, we have the following ODE set,

$$\begin{aligned}\dot{A}_t &= -\kappa\theta B_t - \lambda(\frac{e^{\epsilon\omega_0+\delta^2\omega_0^2/2}}{1-\eta(B_t+\rho_J\omega_0)} - 1) \\ &\quad -\omega_0\lambda(e^{\epsilon+\delta^2/2} - 1) \qquad (6.11) \\ \dot{B}_t &= \frac{1}{2}(\omega_0 - \omega_0^2) + (\kappa - \rho\sigma\omega_0)B_t \\ &\quad -\frac{1}{2}\sigma^2 B_t^2 \qquad (6.12)\end{aligned}$$

where $\dot{}$ denotes the derivative w.r.t to time t. And the initial conditions are,

$$\begin{aligned}A_T &= 0 \qquad (6.13) \\ B_T &= \omega_1 \qquad (6.14)\end{aligned}$$

which is derived from equation (6.10).

To price derivatives on equity price and variance, the joint characteristic function is not necessary, since the payoff only depends on the final state of either equity price or instantaneous variance. Therefore, we can solve the characteristic functions for $\ln S_T$ and v_T separately, which is much easier. If we force $\omega_1 = 0$, we have the solution for $A(\omega_0, 0, t, T)$ and $B(\omega_0, 0, t, T)$, which is the solution for the characteristic function of equity price. The solution as shown below can be found in Pan (2002) and Sepp (2003),

$$
\begin{aligned}
A(\omega_0,0,t,T) &= -\frac{\kappa\theta}{\sigma^2}\left(\psi_+(T-t)+2\ln\frac{\psi_-+\psi_+e^{-\zeta(T-t)}}{2\zeta}\right) \qquad (6.15)\\
&- \lambda(T-t)\left(\omega_0(\frac{e^{\epsilon+\delta^2/2}}{1-\rho_J\eta}-1)+1\right)\\
&+ \lambda e^{\epsilon\omega_0+\delta^2\omega_0^2/2}\big(\frac{\psi_-}{\psi_-L+\eta U}(T-t)-\\
&- \frac{2\eta U}{(\zeta L)^2-(ML+\eta U)^2}\ln(1-\frac{\psi_+-\eta U}{2\zeta L}(1-e^{-\zeta(T-t)}))\big)\\
B(\omega_0,0,t,T) &= -(\omega_0-\omega_0^2)\frac{1-e^{-\zeta(T-t)}}{\psi_-+\psi_+e^{-\zeta(T-t)}} \qquad (6.16)\\
\zeta &= \sqrt{(\kappa-\rho\sigma\omega_0)^2+\sigma^2(\omega_0-\omega_0^2)}\\
\psi_\pm &= \zeta\mp(\kappa-\rho\sigma\omega_0)\\
\xi &= \sqrt{\kappa_\lambda^2-2\sigma_\lambda^2\Lambda}\\
\chi_\pm &= \xi\mp\kappa_\lambda
\end{aligned}
$$

To solve the characteristic function for v_T, we can set $\omega_0 = 0$, and the solution for ODE (6.11) and (6.12) is given by,

$$
\begin{aligned}
A(0,\omega_1,t,T) &= \frac{2}{\sigma^2}\kappa\theta\ln(\frac{2\kappa}{\sigma^2\omega_1(e^{-\kappa(T-t)}-1)+2\kappa}) \qquad (6.17)\\
&+ \frac{2\eta\lambda}{2\eta\kappa-\sigma^2}\ln(\frac{\omega_1(\sigma^2-2\eta\kappa)e^{-\kappa(T-t)}+2\kappa-\omega\sigma^2}{2\kappa(1-\omega\eta)})\\
B(0,\omega_1,t,T) &= \frac{2\kappa\omega_1e^{-\kappa(T-t)}}{\sigma^2\omega_1(e^{-\kappa(T-t)}-1)+2\kappa} \qquad (6.18)
\end{aligned}
$$

Similar results can be found in Sepp (2008). Since $\bar{V}(v_T) = C_0 + C_1v_T$, we have the

characteristic function for the expectation of accumulated variance,

$$\begin{aligned}
\phi_{\bar{V}}(\nu,t,T) &= \int_{-\infty}^{\infty} e^{\bar{V}\nu} p(\bar{V}|\mathcal{F}_t) dX_T \\
&= \int_{-\infty}^{\infty} e^{(C_0+C_1 v_T)\nu} p(v_T|\mathcal{F}_t) dX_T \\
&= e^{\nu C_0} \int_{-\infty}^{\infty} e^{v_T C_1 \nu} p(v_T|\mathcal{F}_t) dX_T \\
&= e^{\nu C_0} G(0, C_1\nu, t, T) \qquad (6.19)
\end{aligned}$$

So we can apply $G(-\omega_0, 0, t, T)$ to equation (6.2) to price Europe option on the SPX index, and substitute equation (6.19) to equations (6.7) and (6.8) to price VIX future and VIX option. In this manner, the pricing of options on SPX and VIX are consistent.

6.3.2 Multi-factor Models

In this section, we will study two multi-factor Models. We derive the characteristic function for both stock price and variance.

6.3.2.1 Stochastic Variance Self-Exciting Jump

The first model we propose here is a Self Exciting jump model,

$$\begin{aligned}
d\ln S_t &= (r - \frac{v_t}{2} - m_1\lambda_t)dt + \sqrt{v_t}dW_t + \int J_x \mu(dJ_x, dJ_\lambda, dt) \\
dv_t &= \kappa_1(\theta_1 - v_t)dt + \sigma_1\sqrt{v_t}dZ_t \\
d\lambda_t &= \kappa_2(\theta_2 - \lambda_t)dt + \int_{\mathbb{R}\times\mathbb{R}^+} 1_{J_x<0} J_\lambda \mu(dJ_x, dJ_\lambda, dt) \\
J_x &\sim \text{Double Exponential}(\eta_u, \eta_d, p) \\
J_\lambda &\sim \text{Exponential}(\eta)
\end{aligned}$$

where Double Exponential$(\eta_u, \eta_d, p) = 1_{x>0}\frac{p}{\eta_u}e^{-x/\eta_u} + 1_{x<0}\frac{1-p}{\eta_d}e^{-x/\eta_d}$ and $0 < p < 1$ and $\eta, \eta_u, \eta_d > 0$. The two factors of the model are stochastic variance and stochastic jump intensity. The stochastic variance process is driven by a CIR process, as in the Heston model. The stochastic jump intensity is driven by negative realized jump in equity price. Once there is a negative jump in equity price, there is a corresponding

jump in jump intensity. We assume jump size in equity price and that in jump intensity are independent. So the corresponding PIDE for the characteristic function for joint distribution of v_T and λ_T is given as,

$$\begin{aligned} & G_t + (r - \frac{1}{2}v_t - \lambda_t m_1)G_x + \frac{1}{2}v_t G_{xx} + \kappa(\theta - v_t)G_v \\ + \; & \frac{1}{2}\sigma^2 v_t G_{vv} + \rho\sigma v_t G_{xv} + \kappa_\lambda(\theta_\lambda - \lambda_t)G_\lambda \\ + \; & \lambda_t \int [G(\omega_0, \omega_1, \omega_2, x + J_x, v_t, \lambda_t + 1_{J<0}J_\lambda, t, T) \\ - \; & G(\omega_0, \omega_1, \omega_2, x, v_t, \lambda_t, t, T)]\nu(dJ_x, dJ_\lambda) \\ = \; & 0 \end{aligned} \tag{6.20}$$

$$G_T = e^{\omega_0 x_T + \omega_1 v_T + \omega_2 \lambda_T} \tag{6.21}$$

The solution form is $G = e^{\omega_0 x_t + r(T-t)\omega_0 + A + Bv_t + C\lambda_t}$. By substituting it into the PIDE (6.20), the corresponding ODE set is,

$$\begin{aligned} \dot{A}_t &= \kappa_1\theta_1 B_t + \kappa_2\theta_2 C_t \\ \dot{B}_t &= \frac{1}{2}(\omega_0 - \omega_0^2) + (\kappa_1 - \rho\sigma\omega_0)B_t - \frac{1}{2}\sigma^2 B_t^2 \\ \dot{C}_t &= -\kappa_2 C_t + \Lambda(\omega_0, C_t) \end{aligned} \tag{6.22}$$

where

$$\begin{aligned} \Lambda(\omega_0, C_t) &= \int e^{\omega_0 J_x + 1_{J_x<0}J_\lambda C_t}\pi(J_x, J_\lambda)dJ_x dJ_\lambda - 1 \\ &= \left(\frac{p}{1 - \omega_0\eta_u} + \frac{1-p}{1 + \omega_0\eta_d}\frac{1}{1 - C_t\eta}\right) - 1 \end{aligned}$$

and initial conditions at time T are,

$$\begin{aligned} A_T &= 0 \\ B_T &= \omega_1 \\ C_T &= \omega_2 \end{aligned}$$

which are derived from equation (6.21).

Since the model is the sum of a Heston stochastic variance and Self Exciting jump and there is no dependence between them, we can find the solutions for B and C separately. It is easy to see that the solution of B is just the degenerate case of equations (6.16) and (6.18). However, there is no close form solution for C, and it has to be solved using numerical method, e.g. Runge Kutta method.

Finally, we derive the characteristic function for the expectation of accumulated variance. Following the steps in the previous section, we first write the expectation of accumulated variance starting at time T as a function of state variables, v_T and λ_T,

$$\begin{aligned}
\bar{V}(v_T, \lambda_T) &= C_0(\Delta T) + C_1(\Delta T)v_T + C_2(\Delta T)\lambda_T \\
C_0 &= \theta_1(1 - C_1) + \theta_2(1 - C_2) \\
C_1 &= \frac{1 - e^{-\kappa_1 \Delta T}}{\kappa_1 \Delta T} \\
C_2 &= m\frac{1 - e^{-\kappa_2 \Delta T}}{\kappa_2 \Delta T}
\end{aligned}$$

Given the independence of v_T and λ_T, we have,

$$\phi_{\bar{V}}(\nu, t, T) = e^{\nu C_0}\phi_v(C_1\nu, t, T)\phi_\lambda(C_2\nu, t, T) \tag{6.23}$$

for the characteristic function of the expectation of accumulated variance settled at time T, where $\phi_v(\nu, t, T) = G(0, \nu, 0, x_t, v_t, \lambda_t, t, T)$ and $\phi_\lambda(\nu, t, T) = G(0, 0, \nu, x_t, v_t, \lambda_t, t, T)$ are the characteristic functions of v_T and λ_T respectively.

6.3.2.2 Double Heston with Jumps

The second multi-factor model is the double Heston model with jumps in one of the two stochastic variance processes.

$$
\begin{aligned}
d\ln S_t &= (r - \frac{v_{1,t}+v_{2,t}}{2} - m_1\lambda)dt + \sqrt{v_{1,t}}dW_{1,t} + \sqrt{v_{2,t}}dW_{2,t} \\
&+ \int J_x \mu(dJ_x, dJ_v, dt) \\
dv_{1,t} &= \kappa_1(\theta_1 - v_{1,t})dt + \sigma_1\sqrt{v_{1,t}}dZ_{1,t} + \int J_v \mu(dJ_x, dJ_v, dt) \\
dv_{2,t} &= \kappa_2(\theta_2 - v_{2,t})dt + \sigma_2\sqrt{v_{2,t}}dZ_{2,t} \\
[dW_{1,t}, dZ_{1,t}] &= \rho_1 dt \\
[dW_{2,t}, dZ_{2,t}] &= \rho_2 dt \\
J_x &\sim N(\epsilon + \rho_J J_v, \delta^2) \\
J_v &\sim Exp(\eta)
\end{aligned}
$$

This model is equivalent to the one factor model in Section (6.3.1) plus another CIR stochastic variance process. The corresponding PIDE is,

$$
\begin{aligned}
& G_t + (r - \frac{1}{2}(v_{1,t}+v_{2,t}) - \lambda m_1)G_x + \frac{1}{2}v_{1,t}G_{xx} + \frac{1}{2}v_{2,t}G_{xx} \\
+ \;& \kappa_1(\theta_1 - v_{1,t})G_{v_1} + \frac{1}{2}\sigma_1^2 v_{1,t}G_{v_1v_1} + \rho_1\sigma_1 v_{1,t}G_{xv_1} \qquad (6.24) \\
+ \;& \kappa_2(\theta_2 - v_{2,t})G_{v_2} + \frac{1}{2}\sigma_2^2 v_{2,t}G_{v_2v_2} + \rho_2\sigma_2 v_{2,t}G_{xv_2} \qquad (6.25) \\
+ \;& \lambda \int [G(\omega_0, \omega_1, \omega_2, x + J_x, v_t + J_v, t, T) \\
- \;& G(\omega_0, \omega_1, \omega_2, x, v_{1,t}, t, T)]\nu(dJ_x, dJ_v) \\
= \;& 0 \\
G_T &= e^{\omega_0 x_T + \omega_1 v_{1,T} + \omega_2 v_{2,T}} \qquad (6.26)
\end{aligned}
$$

By substituting the solution form $G = e^{\omega_0 x_t + r(T-t)\omega_0 + A + B_1 v_{1,t} + B_2 v_{2,t}}$, we have the following ODE set,

$$
\begin{aligned}
\dot{A}_t &= -\kappa_1\theta_1 B_{1,t} - \kappa_2\theta_2 B_{2,t} \qquad (6.27)\\
&\quad -\lambda\left(\frac{e^{\epsilon\omega_0+\delta^2\omega_0^2/2}}{1-\eta(B_{1,t}+\rho_J\omega_0)} - 1\right) - \omega_0\lambda(e^{\epsilon+\delta^2/2}-1)\\
\dot{B}_{1,t} &= \frac{1}{2}(\omega_0-\omega_0^2) + (\kappa_1 - \rho_1\sigma_1\omega_0)B_{1,t} \qquad (6.28)\\
&\quad -\frac{1}{2}\sigma_1^2 B_{1,t}^2\\
\dot{B}_{2,t} &= \frac{1}{2}(\omega_0-\omega_0^2) + (\kappa_2 - \rho_1\sigma_2\omega_0)B_{2,t} \qquad (6.29)\\
&\quad -\frac{1}{2}\sigma_2^2 B_{2,t}^2
\end{aligned}
$$

and the initial conditions at time T are,

$$
\begin{aligned}
A_T &= 0\\
B_{1,T} &= \omega_1\\
B_{2,T} &= \omega_2
\end{aligned}
$$

which are derived from equation (6.26).

The solution form for B_1 are the same as equation (6.16) and (6.18), and B_2 is the degenerating case of B_1 as λ goes to zero. Due to the independence of $v_{1,T}$ and $v_{2,T}$, the expectation of accumulated variance is,

$$
\begin{aligned}
\bar{V}(v_{1,T}, v_{2,T}) &= C_0 + C_1 v_{1,T} + C_2 v_{2,T}\\
C_0 &= \lambda\left(m_2 + \frac{\eta(e^{-\kappa_1\Delta T} + \kappa_1\Delta T - 1)}{\kappa_1^2\Delta t}\right)\\
&+ \theta_1(1-C_1) + \theta_2(1-C_2)\\
C_1 &= \frac{1-e^{-\kappa_1\Delta T}}{\kappa_1\Delta T}\\
C_2 &= \frac{1-e^{-\kappa_2\Delta T}}{\kappa_2\Delta T}
\end{aligned}
$$

Similar to the case of SVSEJ model, because of the independence of two state

variables, the characteristic function of the expectation of accumulated variance is,

$$\phi_{\bar{V}}(\nu, t, T) = e^{\nu C_0}\phi_{v_1}(C_1\nu, t, T)\phi_{v_2}(C_2\nu, t, T) \tag{6.30}$$

where $\phi_{v_1}(\nu, t, T) = G(0, \nu, 0, x_t, v_t, \lambda_t, t, T)$ and $\phi_{v_2}(\nu, t, T) = G(0, 0, \nu, x_t, v_t, \lambda_t, t, T)$. Applying equations (6.7) and (6.8), we can calculate the price for VIX future and VIX option.

6.4 Correlation and Hedging

In this section, we analyze the instantaneous correlation term structure for VIX futures implied by the models from the previous section. We first show that the one factor model implies an unrealistic constant correlation term structure for VIX futures, contrary to the empirical observation.

6.4.1 Flaw of Single Factor Models

Single factor stochastic variance models are very popular for pricing equity and FX derivatives. However, Single-Factor models have some implications that are undesirable. For example, if stochastic variance is the only driving factor for the volatility, the correlation term structure of variance derivatives is always one. We use the SVJJ model as example to illustrate this pitfall below. For the sake of simplicity, from now on, we use Vix_t^T instead of $F_{VIX}(t, T)$ to denote VIX future price observe at t and settled at T. Given the SVJJ specification in (6.6), we can derive the SDE for VIX future price, which is observed at t and settled at T,

$$dVix_t^T = \frac{\partial Vix_t^T}{\partial v_t}\sqrt{v_t}\sigma dZ_t + \int_{\mathbb{R}\times\mathbb{R}^+} \delta Vix_t^T \left(\mu(dJ_v, dt) - \upsilon(dJ_v, dt)\right)$$

where δVix_t^T is due to the jump in variance process, and $\delta Vix_t^T = Vix_t^T(v_t + J_v) - Vix_t^T(v_t)$.[5] Given the above dynamics for VIX future, the correlation between VIX futures of different maturities is

[5]Note that the dynamics of VIX future price is a martingale.

$$\begin{aligned}\varrho_t &= \frac{\left\langle dVix_t^{T_1}, dVix_t^{T_2}\right\rangle}{\sqrt{\left\langle dVix_t^{T_1}, dVix_t^{T_1}\right\rangle\left\langle dVix_t^{T_2}, dVix_t^{T_2}\right\rangle}} \\ &= \frac{\frac{\partial Vix_t^{T_1}}{\partial v_t}\frac{\partial Vix_t^{T_2}}{\partial v_t}\sigma^2 v_t + \lambda\int_{\mathbb{R}}\delta Vix_t^{T_1}\delta Vix_t^{T_2}\upsilon(dJ_v)}{\sqrt{\prod_{i=1}^{2}\left(\left(\frac{\partial Vix_t^{T_i}}{\partial v_t}\right)^2\sigma^2 v_t + \lambda\int_{\mathbb{R}}\left(\delta Vix_t^{T_i}\right)^2\upsilon(dJ_v)\right)}}\end{aligned} \tag{6.31}$$

Substituting the characteristic function (6.19) into equation (6.7) and taking derivative, we have

$$\frac{\partial Vix_t^T}{\partial v_t} = \frac{1}{2\sqrt{\pi}}\int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} B(C_1\nu,t,T)\phi_{\bar{V}}(\nu,t,T)\right]d\nu$$

where $\phi_{\bar{V}}(\nu,t,T)$ is given by (6.19). For the change due to jump, we use the linear approximation as follows,

$$\begin{aligned}\delta Vix_t^T &= Vix_t^T(v_t+J_v) - Vix_t^T(v_t) \\ &= \frac{1}{2\sqrt{\pi}}\int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}}\left(e^{BJ_v}-1\right)\phi_{\bar{V}}(\nu,t,T)\right]d\nu \\ &\approx \frac{J_v}{2\sqrt{\pi}}\int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} B(C_1\nu,t,T)\phi_{\bar{V}}(\nu,t,T)\right]d\nu\end{aligned}$$

Therefore, under the linear approximation for jump, the correlation is,

$$\begin{aligned}\varrho_t &\approx \frac{D_1D_2\sigma^2 v_t + 2D_1D_2\eta^2\lambda}{\sqrt{\left(D_1^2\sigma^2 v_t + 2D_1^2\eta^2\lambda\right)\left(D_2^2\sigma^2 v_t + 2D_2^2\eta^2\lambda\right)}} \\ &= 1\end{aligned}$$

where $D_i = \frac{1}{2\sqrt{\pi}}\int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} B(C_1\nu,t,T_i)\phi_{\bar{V}}(\nu,t,T_i)\right]d\nu$. Therefore, for two VIX futures with maturities T_1 and T_2, the correlation is always one, contradicting to the market observation. For the one factor model, if the price of one VIX future with certain maturity is given, the state variable can be derived and therefore price of the second VIX futures with a different maturity is also known. However, this is not the case for the multi-factor models.

6.4.2 VIX Future Dynamics and Correlation implied by Multiple Factor Model

6.4.2.1 Stochastic Variance Self-Exciting Jump

First we looked at the dynamics of VIX future under the Self-Exciting jump model, where the dynamics on the VIX future is derived as follows,

$$\begin{aligned}
dVix_t^T &= \bullet dt + \frac{\partial Vix_t^T}{\partial v_t} dv_t + \int_{\mathbb{R}\times\mathbb{R}^+} \delta Vix_t^T \mu^* \\
&= \bullet dt + \frac{dv_t}{2\sqrt{\pi}} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} e^{i\nu C_0} B(C_1)\phi_v(C_1)\phi_\lambda(C_2)\right] d\nu \\
&\quad + \frac{1}{2\sqrt{\pi}} \int_{\mathbb{R}\times\mathbb{R}^+} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} e^{i\nu C_0} \phi_v(C_1)\left(\phi_{\lambda+J_\lambda}(C_2) - \phi_\lambda(C_2)\right)\right] d\nu \times \mu^* \\
&= \frac{\sqrt{v_t}\sigma}{2\sqrt{\pi}} dW_t \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} B(C_1)\phi_{\bar{V}}(\nu,t,T)\right] d\nu \\
&\quad + \frac{1}{2\sqrt{\pi}} \int_{\mathbb{R}\times\mathbb{R}^+} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} \left(e^{J_\lambda C(C_2)} - 1\right)\phi_{\bar{V}}(\nu,t,T)\right] d\nu \times \mu^*
\end{aligned}$$

where $\mu^* = \mu(dJ_x, dJ_\lambda, dt) - \upsilon(dJ_x, dJ_\lambda, dt)$ is the compensated random measure for jump, and we write $B(C_i) = B(C_i\nu, t, T)$, $C(C_i) = C(C_i\nu, t, T)$, and $\phi(C_i) = \phi(C_i\nu, t, T)$ for short notation. $\bullet dt$ is some drift which will cancel the drift from other terms, since it is a future contract and total drift is zero. For the difference of jump, we have

$$\begin{aligned}
\delta Vix_t^T &= Vix_t^T(\lambda_t + J_v) - Vix_t^T(\lambda_t) \\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} e^{i\nu C_0}\phi_v(C_1)\phi_\lambda(C_2)\left(e^{CJ_\lambda} - 1\right)\right] d\nu \\
&\approx \frac{J_\lambda}{2\sqrt{\pi}} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} C(C_2)\phi_{\bar{V}}(\nu,t,T)\right] d\nu
\end{aligned}$$

in which the last equality is due to linear approximation. Therefore, the correlation can be calculated as,

$$
\begin{aligned}
\varrho_t &= \frac{\left\langle dVix_t^{T_1}, dVix_t^{T_2} \right\rangle}{\sqrt{\left\langle dVix_t^{T_1}, dVix_t^{T_1} \right\rangle \left\langle dVix_t^{T_2}, dVix_t^{T_2} \right\rangle}} \\
&= \frac{D_1 D_2 \sigma^2 v_t + 2E_1 E_2 P(J_x < 0)\eta^2 \lambda_t}{\sqrt{\left(D_1^2 \sigma^2 v_t + 2E_1^2 \eta^2 \lambda_t\right)\left(D_2^2 \sigma^2 v_t + 2E_2^2 \eta^2 \lambda_t\right)}} \\
&= \frac{\sigma^2 v_t + \frac{2E_1 E_2}{D_1 D_2}\eta^2 P(J_x < 0)\lambda_t}{\sqrt{\prod_{i=1}^{2}\left(\sigma^2 v_t + 2(\frac{E_i}{D_i})^2 \eta^2 P(J_x < 0)\lambda_t\right)}}
\end{aligned}
$$

where

$$
\begin{aligned}
D_i &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} B(C_1\nu, t, T_i)\phi_{\bar{V}}(\nu, t, T_i)\right] d\nu \\
E_i &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re\left[\left(\frac{1}{i\nu}\right)^{\frac{3}{2}} C(C_2\nu, t, T_i)\phi_{\bar{V}}(\nu, t, T_i)\right] d\nu
\end{aligned}
$$

and B and C are solutions to the ODE set (6.22). B has the same solution forms as equation (6.18), and there is no close form for C, which can be solved by ODE solver numerically. $\phi_{\bar{V}}$ is given by equation (6.23).

6.4.2.2 Double Heston with Jumps

For the Double Heston with Jumps model, the dynamics for VIX future is,

$$
\begin{aligned}
dVix_t^T &= \bullet dt + \frac{\partial Vix_t^T}{\partial v_{1,t}} dv_{1,t} + \frac{\partial Vix_t^T}{\partial v_{2,t}} dv_{2,t} \\
&= \frac{dv_{1,t}}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} e^{i\nu C_0} B(C_1) \phi_{v_1}(C_1) \phi_{v_2}(C_2) \right] d\nu \\
&+ \frac{dv_{2,t}}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} e^{i\nu C_0} B(C_2) \phi_{v_1}(C_1) \phi_{v_2}(C_2) \right] d\nu \\
&+ \int_{\mathbb{R}\times\mathbb{R}^+} \delta Vix_t^T \left(\mu(dJ_v, dt) - \upsilon(dJ_v, dt) \right) + \bullet dt \\
&= \frac{\sqrt{v_{1,t}}\sigma_1}{2\sqrt{\pi}} dZ_{1,t} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B(C_1) \phi_{\bar{V}}(\nu, t, T) \right] d\nu \\
&+ \frac{\sqrt{v_{2,t}}\sigma_2}{2\sqrt{\pi}} dZ_{2,t} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B(C_2) \phi_{\bar{V}}(\nu, t, T) \right] d\nu \\
&+ \frac{1}{2\sqrt{\pi}} \int_{\mathbb{R}\times\mathbb{R}^+} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} \left(e^{J_v C(C_2)} - 1 \right) \phi_{\bar{V}}(\nu, t, T) \right] d\nu \times \mu^*
\end{aligned}
$$

where $\mu^* = \mu(dJ_x, dJ_v, dt) - \upsilon(dJ_x, dJ_v, dt)$ is the compensated random measure for jump. As before, the total drift for a future contract is zero. Again, we use linear approximation for the jump part on stochastic variance,

$$
\begin{aligned}
\delta Vix_t^T &= Vix_t^T(v_t + J_v) - Vix_t^T(v_t) \\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} \left(e^{BJ_v} - 1 \right) \phi_{\bar{V}}(\nu, t, T) \right] d\nu \\
&\approx \frac{J_v}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B(\omega, t, T) \phi_{\bar{V}}(\nu, t, T) \right] d\nu
\end{aligned}
$$

As before, the correlation can be calculated as,

$$
\begin{aligned}
\varrho_t &= \frac{\left\langle dVix_t^{T_1}, dVix_t^{T_2} \right\rangle}{\sqrt{\left\langle dVix_t^{T_1}, dVix_t^{T_1} \right\rangle \left\langle dVix_t^{T_2}, dVix_t^{T_2} \right\rangle}} \\
&= \frac{D_1 D_2 \left(\sigma^2 v_{1,t} + 2\eta^2 \lambda \right) + E_1 E_2 \sigma^2 v_{2,t}}{\sqrt{\prod_{i=1}^2 \left(D_i^2 \left(\sigma^2 v_{1,t} + 2\eta^2 \lambda \right) + E_i^2 \sigma^2 v_{2,t} \right)}}
\end{aligned}
$$

where for $i = 1, 2$

$$
\begin{aligned}
D_i &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B_1(C_1\nu, t, T_i) \phi_{\bar{V}}(\nu, t, T_i) \right] d\nu \\
E_i &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B_2(C_2\nu, t, T_i) \phi_{\bar{V}}(\nu, t, T_i) \right] d\nu
\end{aligned}
$$

and B_1 and B_2 are solutions to the ODE set (6.27), (6.28), and (6.29). The solution forms are exactly the same as equation (6.18) with corresponding parameters, and $\phi_{\bar{V}}$ is given by equation (6.30).

6.4.3 Calibration

Here, we calibrate the two multi-factor models proposed in the previous section using European style options on SPX and VIX on 25/04/2008. There are 5 maturities, May, June, July, August, and November, that have liquid options traded. The option data is downloaded from OptionMetrics and we have excluded all the options without trading volume. We only use put option on equity and call option on VIX to capture the negative skew for SPX surface and the positive skew for VIX surface. For risk free interest rate, we assume a flat 2% for all maturities.

The calibration is done by minimizing the mean square error between market implied volatility and model implied volatility. There are a few reasons for not matching option prices instead. First, VIX future price is an output of the model, which we need to take into account in the calibration. Hence, we need to assign weights on option price and future price, which has a big influence in the calibration. Moreover, since the option prices for SPX and VIX are not comparable, their pricing errors will have an unbalanced influence on the calibration. By using implied volatility, we avoid any such issue since the implied volatilities for equity and VIX are of the same scale and it is reasonable to put equal weights on the two surfaces in the calibration.

The calibration results for the Self-Exciting jump model is reported in Table 6.1 and the fit of volatility surface is plotted in Figure 6.1. The calibration for Double Heston with Jumps model using the same data is reported in Table 6.2 and the fit of volatility surface is plotted in Figure 6.2. Given the parameter values, we can calculate the instantaneous correlation term structure conditional on the state variables for the two models. The result is plotted in Figure 6.3. The results show

Table 6.1: Calibrated Parameters of Self Exciting Jump Model

	Parameter Estimates
κ_1	1.4791
θ_1	0.0364
σ_1	0.5389
ρ_1	-0.8454
κ_2	34.3965
θ_2	6.2696
η	22.4183
η_u	0.0362
η_d	0.0422
p	0.3686

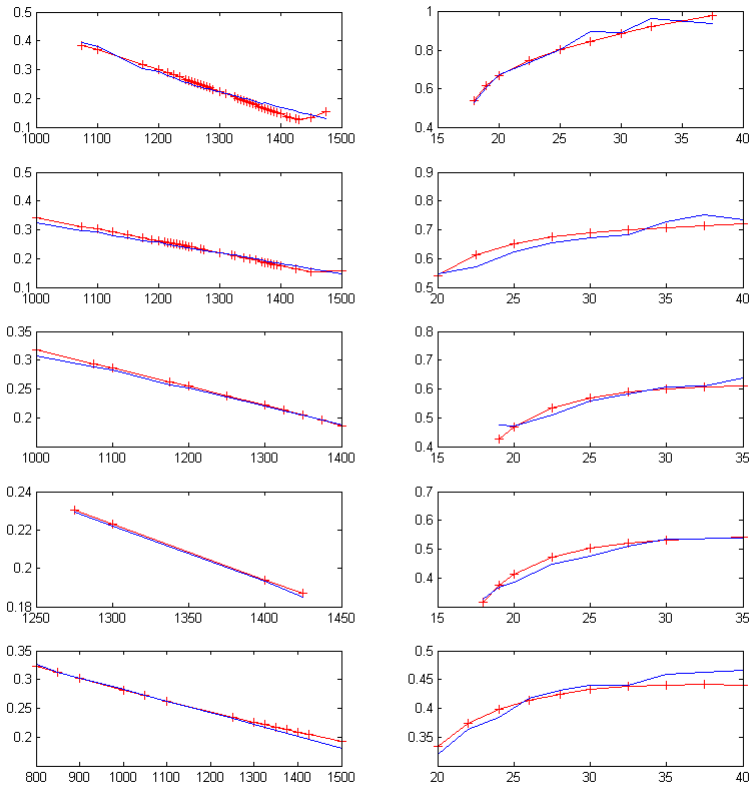
This table reports the calibrated parameters of Self Exciting Jump model. The calibration is done by min square error of implied volatility between market and model.

Table 6.2: Calibrated Parameters of Double Heston with Jump Model

	Parameter Estimates
κ_1	3.6721
θ_1	0.0195
σ_1	0.6981
ρ_1	-0.8562
η	0.0410
ρ_J	1.0395
κ_2	34.2975
θ_2	0.0234
σ_2	1.3896
ρ_2	-0.9999
λ	1.1576
ϵ	-0.0667
δ	0.0522

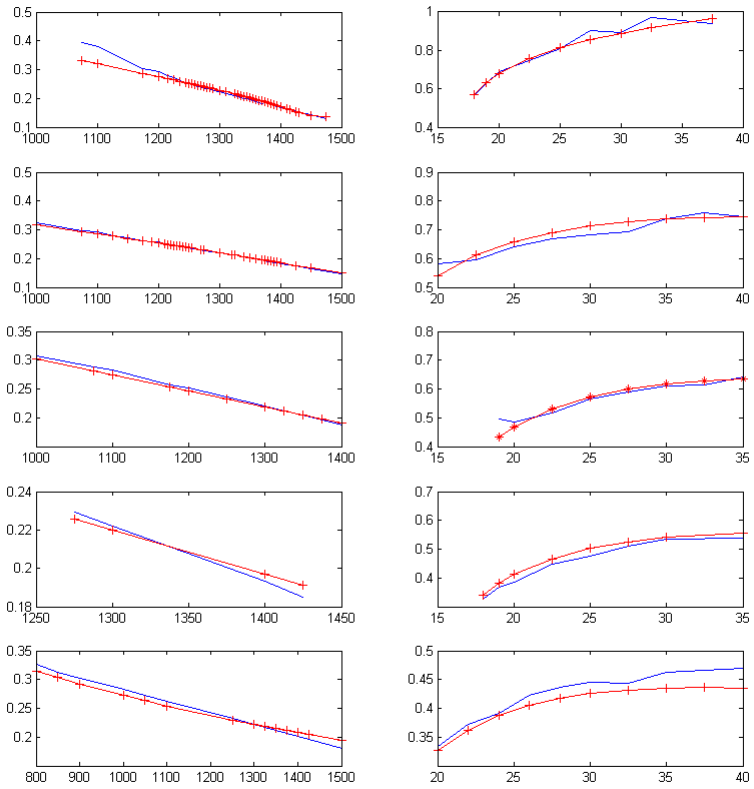
This table reports the calibrated parameters of Double Heston with Jump model. The calibration is done by min square error of implied volatility between market and model.

Figure 6.1: Fit of Implied Volatility by Self Exciting Jump Model



This graph plots the fit of implied volatility on 25/04/2008 with 5 maturities. The red line with crosses is the implied volatility by model and the blue line the implied volatility by market price. The left column is the fit of the implied vol of Equity put options; the right column is the fit of the implied vol of VIX call options. From top to bottom are the maturities in May, June, July, August, and November in 2008 respectively.

Figure 6.2: Fit of Implied Volatility by Double Heston with Jump Model



This graph plots the fit of implied volatility on 25/04/2008 with 5 maturities. The red line with crosses is the implied volatility by model and the blue line the implied volatility by market price. The left column is the fit of the implied vol of Equity put options; the right column is the fit of the implied vol of VIX call options. From top to bottom are the maturities in May, June, July, August, and November in 2008 respectively.

that both models calibrated well jointly to the two surfaces. However, they produce a different correlation term structure for the VIX futures. The SVSEJ model has a faster decay while Double Heston with Jumps is more linear and decays slower.

It should be noted that the calibration here only uses the option data on one particular date, which might not be enough for the model with many parameters. The main problem is that the calibration result could be unstable. For example, a change in the initial value of model parameters could end up with a relative big difference in the final results, or if we choose some other day, the calibrated results might also be significantly different. However, since the focus of this research is to examine the feature of different models, not to explain the empirical observation in the market, we will take the calibrated results for the further works in this chapter without questioning its accuracy.

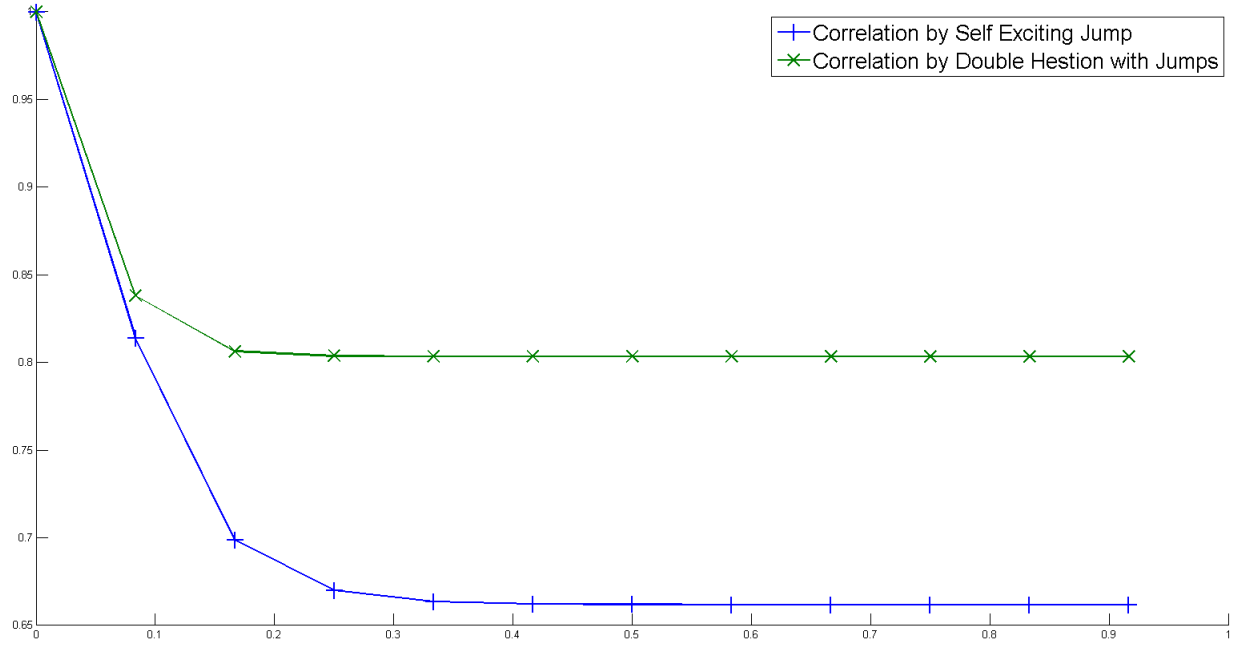
Another issue worth to mention is the the estimation of jumps. As mentioned in chapter 2 and 5, the idea of estimation is to match the distribution implied by the model and that observed in the market. And the distribution here is under the risk neutral measure. The estimated jump distribution has nothing to do with the observed jump distribution in the return series. One can argue that the jump distributions under two measures are linked by Radon–Nikodym derivative, i.e., the pricing kernel. But since I only use the option price to calibrate the model, the estimated jump distribution is the one that provides the best fit to the implied volatility, rather than a measure of how much the stock market could fluctuate.

6.4.4 Unconditional Correlation Term Structure of VIX Future

We have shown how to derive the conditional correlation of VIX futures in the previous sections. Since the conditional correlation term structure of VIX future is conditional on state variables and change over time and unobservable, the unconditional correlation might be of more interest, because one can calculate it from historical VIX future data. VIX future can be expressed as,

$$\begin{aligned} \varrho &= \mathbb{E}\left[\varrho_t\right] \\ &= \int \varrho_t p(f_1, f_2) df_1 df_2 \end{aligned} \tag{6.32}$$

Figure 6.3: Conditional Instantaneous Correlation Term Structure between VIX Spot and VIX Future



This graph plots the conditional instantaneous correlation term structure of VIX future on 25/4/2008.

where f_1 and f_2 are the state variables. Since the SVSEJ and Double Heston with Jumps models are Markov models, the state variables have stationary distributions under some trivial assumptions. We derived the stationary distribution for the state variables for these two models in Appendix A. The following is a summary of the stationary distributions used in our study.

1. For the CIR process in Heston model, the stationary distribution for variance is a Gamma distribution.

2. For the Heston with jump in variance process, the stationary distribution is not any known distribution. However, when the jump size in variance is small, specifically, $\eta < \frac{\sigma^2}{2\kappa}$, the distribution can be approximated by another Gamma distribution. When the jump size is big, the distribution of variance is a sum of two independent Gamma distributions.

3. The stationary distribution for the jump intensity of the self exciting jump process is also a Gamma distribution, but with a shift in its base line level θ.

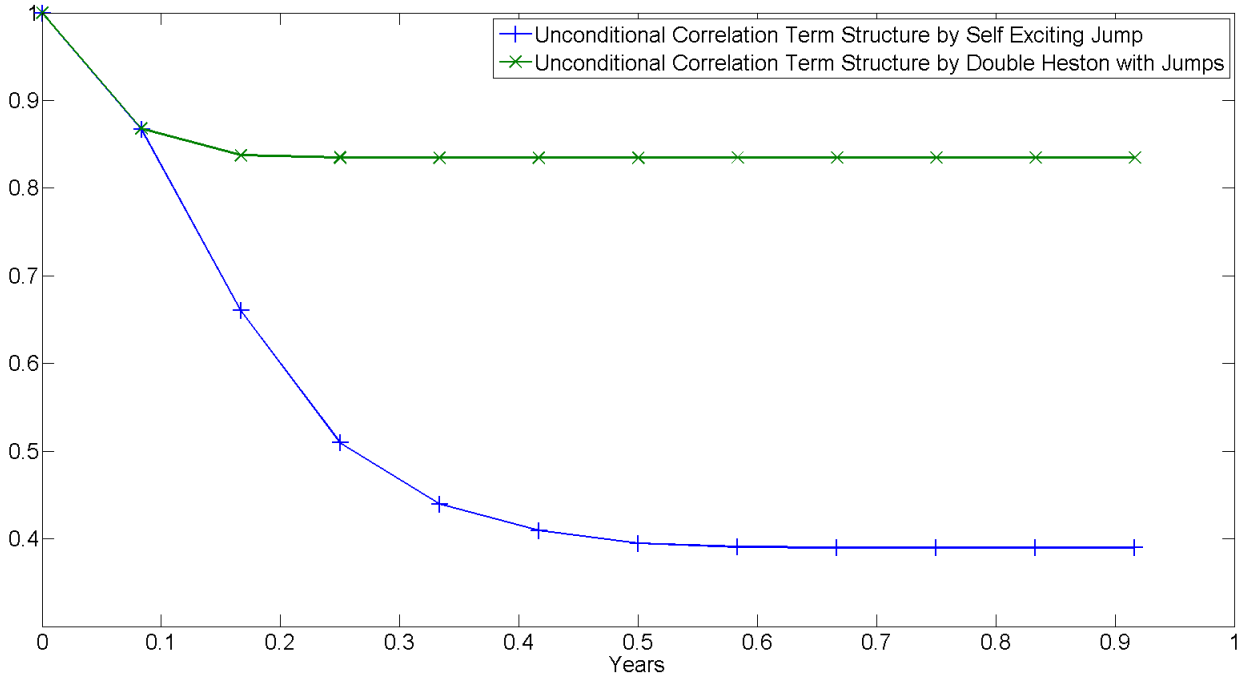
Since the volatility factors are independent in both models, the joint distribution of the factors are just the product of the marginal distributions. Using equation (6.32), we can calculate the unconditional correlation, given the model parameters and stationary distribution functions.

Figure 6.4 plots the unconditional term structure for VIX future. Not surprisingly, two models produce very different term structures. Unlike the conditional correlation, the term structure of unconditional correlation for VIX future can be calculated from market data. It will be interesting to compare with the correlation implied by model with that observed in the market to see which model fit the market better. But it should be noted that the unconditional distribution of factors, $p(f_1, f_2)$, is supposed to be under the $\mathbb{P}$ measure. In our study, we simply use the model parameters under the $\mathbb{Q}$ measure as a demonstration. To estimate the model under both measures, one can use the approach introduced in Chapter 5.

6.4.5 Hedging Delta of VIX Options

Here we compare the hedge ratio for VIX option w.r.t VIX future using the model studied in the previous section. In the one factor model without jumps, one can achieve perfect hedge, and the dynamical optimal hedge ratio is,

Figure 6.4: Unconditional Instantaneous Correlation Term Structure between VIX Spot and VIX Future



This graph plots the unconditional instantaneous correlation term structure of VIX future implied by two models.

$$h_t = \frac{\frac{\partial C}{\partial v_t}}{\frac{\partial F}{\partial v_t}}$$

where C is the derivative, i.e., VIX option, to be hedged and F is the hedging instrument, i.e., VIX future. For the multi-factor models, perfect hedging is impossible. The hedge ratio or delta which gives minimal hedging variance is,

$$h(K) = \frac{\left\langle dC_t^T(K), dF_t^T \right\rangle}{\langle dF_t^T, dF_t^T \rangle}$$

In SVSEJ, the dynamics of VIX future or option is given by,

$$d\bullet \approx \frac{\partial \bullet}{\partial v_t}\sqrt{v_t}\sigma dZ_t + \frac{\partial \bullet}{\partial \lambda_t}\int_{\mathbb{R}\times\mathbb{R}^+} 1_{J_x<0} J_\lambda \left(\mu(dJ_x, dJ_\lambda, dt) - \upsilon(dJ_x, dJ_\lambda, dt)\right)$$

whereas in Double Heston with Jumps model model, we have

$$\begin{aligned} d\bullet &\approx \frac{\partial \bullet}{\partial v_{1,t}}\left(\sqrt{v_{1,t}}\sigma dZ_{1,t} + \int_{\mathbb{R}\times\mathbb{R}^+} J_v \left(\mu(dJ_x, dJ_v, dt) - \upsilon(dJ_x, dJ_v, dt)\right)\right) \\ &+ \frac{\partial \bullet}{\partial v_{2,t}}\sqrt{v_{2,t}}\sigma dZ_{2,t} \end{aligned}$$

$\frac{\partial \bullet}{\partial v_t}$ can be derived given the close form of characteristic function as shown in Section 6.4. By following the same procedure in Section 6.4.2, we have the optimal hedge ratio for SVSEJ model as,

$$h(K) = \frac{D_F D_C \sigma^2 v_t + 2E_F E_C \eta^2 P(J_x < 0)\lambda_t}{\sqrt{(D_F^2 \sigma^2 v_t + E_F^2 \eta^2 P(J_x < 0)\lambda_t)}}$$

where

$$
\begin{aligned}
D_F &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B(C_1\nu, t, T_1) \phi_{\bar{V}}(\nu, t, T_1) \right] d\nu \\
D_C &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\frac{1 - \text{erf}(K\sqrt{-\nu})}{(i\nu)^{3/2}} B(C_1\nu, t, T_2) \phi_{\bar{V}}(\nu, t, T_2) \right] d\nu \\
E_F &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} C(C_2\nu, t, T_1) \phi_{\bar{V}}(\nu, t, T_1) \right] d\nu \\
E_C &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\frac{1 - \text{erf}(K\sqrt{-\nu})}{(-i\nu)^{3/2}} C(C_2\nu, t, T_2) \phi_{\bar{V}}(\nu, t, T_2) \right] d\nu
\end{aligned}
$$

The optimal hedge ratio for Double Heston with Jumps model is given by,

$$
h(K) = \frac{D_1 D_2 \left(\sigma^2 v_{1,t} + 2\eta^2 \lambda\right) + E_1 E_2 \sigma^2 v_{2,t}}{\sqrt{\left(D_1^2 \left(\sigma^2 v_{1,t} + 2\eta^2 \lambda\right) + E_1^2 \sigma^2 v_{2,t}\right)}}
$$

where

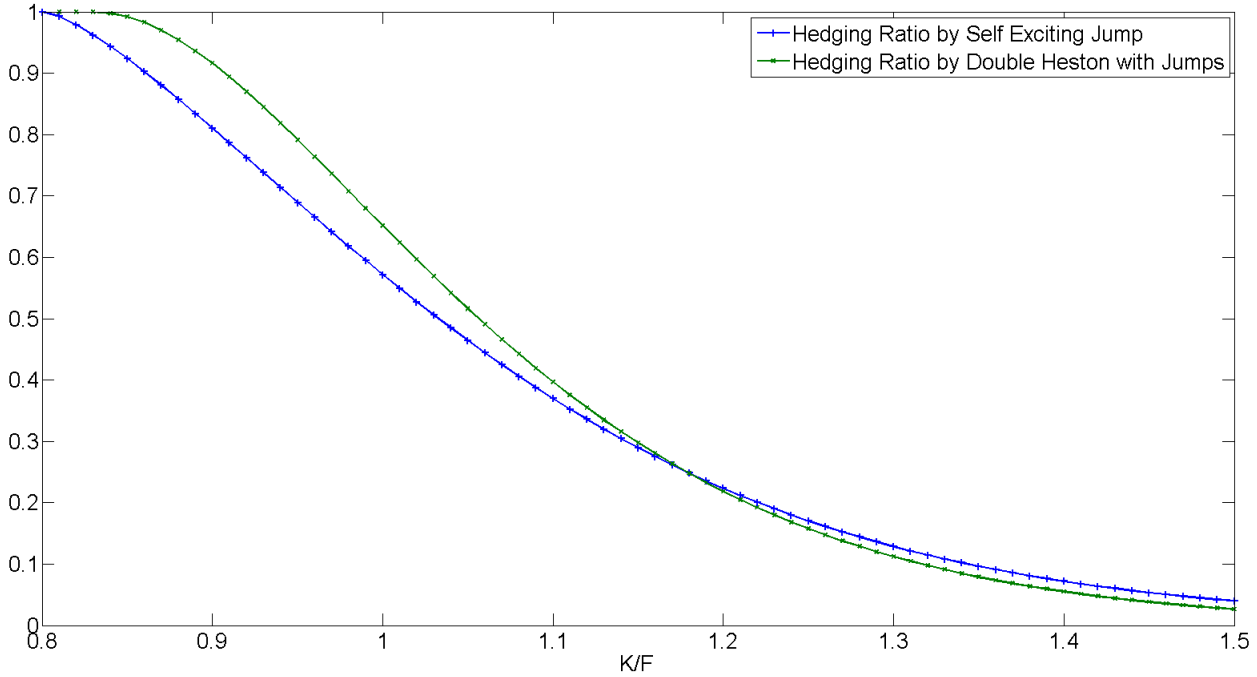
$$
\begin{aligned}
D_F &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B_1(C_1\nu, t, T_1) \phi_{\bar{V}}(\nu, t, T_1) \right] d\nu \\
D_C &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\frac{1 - \text{erf}(K\sqrt{-\nu})}{(i\nu)^{3/2}} B_1(C_1\nu, t, T_2) \phi_{\bar{V}}(\nu, t, T_2) \right] d\nu \\
E_F &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\left(\frac{1}{i\nu} \right)^{\frac{3}{2}} B_2(C_2\nu, t, T_1) \phi_{\bar{V}}(\nu, t, T_1) \right] d\nu \\
E_C &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \Re \left[\frac{1 - \text{erf}(K\sqrt{-\nu})}{(-i\nu)^{3/2}} B_2(C_2\nu, t, T_2) \phi_{\bar{V}}(\nu, t, T_2) \right] d\nu
\end{aligned}
$$

Figure 6.5 plots the hedge ratio implied by the two models based on the parameters estimated in the previous section. For ATM VIX option, the delta for the SVSEJ model is smaller than that for the Double Heston with Jumps model, suggesting a less exposure of VIX option to its underlying under SVSEJ model.

6.5 Conclusion

In this study, we provide a framework for pricing derivatives on SPX and VIX consistently, given the characteristic function of equity price and the factors that drive

Figure 6.5: Optimal Hedge Ratio



The graph plots the hedge ratio of VIX options using VIX future on 25/4/2008. Both maturities are one month.

variance. This method can be applied to any affine model such as Heston model with jumps, since the characteristic functions of affine models are easy to derive.

In this chapter, we are particular interested in the correlation term structure of VIX futures of different time to maturity. We studied two categories of affine models, a one-factor model and two multi-factor models. Based on a linear approximation of the jump size, we show that one-factor models always imply a perfect correlation between VIX futures of different maturities. However, we know that it is not the case in the market. So even one factor model can fit volatility surfaces well, it still implies a wrong dynamics for variance.

Next, we show that two different multi-factor models are able to fit the implied volatility surfaces of SPX option and VIX option well. Based on the calibrated model parameters, however, they produce very different correlation term structure. We also derive the unconditional correlation so that the correlation implied by models can be compared with that calculated from the market historical data.

To calculate the unconditional correlation term structure, we derive the stationary distribution of the factors, based on the Kolmogorov forward equation. By taking expectation of conditional correlation, we can calculate the unconditional correlation term structure. However, to compare with the correlation observed in the market, we need the stationary distributions of the factors under the $\mathbb{P}$ measure. In this study, we only use the parameters calibrated under $\mathbb{Q}$ measure to calculate the unconditional correlation of the factors as a demonstration.

In the last part, we demonstrated how to calculate the optimal hedge ratio for hedging VIX option with VIX future. Since two models implies different dynamics for VIX future, it is not surprised to see that hedge ratio are also different.

Appendix

A. Stationary distribution of CIR process with jumps

The CIR process with jumps is as follows,

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t + \int J_v \mu(dJ_v, dt)$$

The forward equation is,

$$\frac{\partial f(v,t)}{\partial t} = -\frac{\partial}{\partial v}\kappa(\theta - v)f(v,t) + \frac{1}{2}\frac{\partial^2}{\partial v^2}\sigma^2 v f(v,t) - \lambda f(v,t) + \lambda \int f(v - J_v, t)\nu(dJ_v)$$

The stationary distribution is the solution of

$$0 = -\frac{\partial}{\partial v}\kappa(\theta - v)f(v) + \frac{1}{2}\frac{\partial^2}{\partial v^2}\sigma^2 v f(v) - \lambda f(v) + \lambda \int f(v - J_v)\nu(dJ_v)$$

where $\frac{\partial f(v,t)}{\partial t} = 0$, since it is stationary. The difficulty arises due to the integral. We apply Fourier transform to solve this problem. The Fourier transform is defined as,

$$F(\omega) = \mathcal{F}[f(x)] = \int e^{\omega x} f(x)dx$$

where ω is complex number.

So the OIDE becomes,

$$\begin{aligned} 0 &= \omega\left[\kappa\theta F(\omega) - \kappa F'(\omega)\right] + \frac{1}{2}\omega^2\sigma^2 F'(\omega) - \lambda F(\omega) + \lambda F(\omega)\int e^{J_v\omega}\nu(dJ_v) \\ &= (\omega\kappa\theta + \lambda\frac{\eta\omega}{1-\eta\omega})F(\omega) + (\frac{1}{2}\omega^2\sigma^2 - \omega\kappa)F'(\omega) \end{aligned}$$

Therefore, we have,

$$(\omega\kappa - \frac{1}{2}\omega^2\sigma^2)F'(\omega) = (\omega\kappa\theta + \lambda\frac{\eta\omega}{1-\eta\omega})F(\omega)$$

The condition of determined solution is,

$$F(0) = 1$$

The solution for $F(\omega)$, which is the characteristic function of the stationary distribution is,

$$F(\omega) = (1 - \frac{\sigma^2}{2\kappa}\omega)^{-2\kappa\theta/\sigma^2}\left(\frac{1-\eta\omega}{1-\frac{\sigma^2}{2\kappa}\omega}\right)^{-2\eta\lambda/(2\eta\kappa-\sigma^2)}$$

When $\eta > \frac{\sigma^2}{2\kappa}$ and $\lambda < \kappa\theta\frac{2\eta\kappa-\sigma^2}{\sigma^2\eta}$, we can rearrange the expression as,

$$F(\omega) = (1 - \frac{\sigma^2}{2\kappa}\omega)^{2\eta\lambda/(2\eta\kappa-\sigma^2)-2\kappa\theta/\sigma^2} (1 - \eta\omega)^{-2\eta\lambda/(2\eta\kappa-\sigma^2)}$$

which is product of two Characteristic functions of Gamma distribution. So $v_{1,t}$ is the sum of two Gamma distributed random variables.

When $\eta < \frac{\sigma^2}{2\kappa}$, there is no such simple solution, so we expand the Characteristic function and take the first order approximation of η, we have,

$$\Phi(\omega) = (1 - \frac{\sigma^2}{2\kappa}\omega)^{-2(\lambda\eta+\kappa\theta)/\sigma^2}$$

which is just the Characteristic function of a Gamma distribution.

For the second state variable, it is just the degenerate case of $v_{1,t}$ as $\lambda \to 0$, so the Characteristic function is,

$$\Phi(\omega) = (1 - \frac{\sigma^2}{2\kappa}\omega)^{-2\kappa\theta/\sigma^2}$$

which is the characteristic function of Gamma distribution $\Gamma(\frac{2\kappa\theta}{\sigma^2}, \frac{\sigma^2}{2\kappa})$ for the CIR process. Here, it is clear to see that when $2\kappa\theta/\sigma^2 < 1$, the probability of $v_{2,t}$ of being zero is positive.[6]

B. Stationary distribution of self exciting jump intensity

The intensity process for self exciting jump is as follows,

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \int_{\mathbb{R}\times\mathbb{R}^+} 1_{J_x<0} J_\lambda \mu(dJ_x, dJ_\lambda, dt)$$

The forward equation is,

$$\frac{\partial f(\lambda,t)}{\partial t} = -\frac{\partial}{\partial\lambda}\kappa(\theta-\lambda)f(\lambda,t) - \lambda f(\lambda,t) + P(J_x<0)\int(\lambda - J_\lambda)f(\lambda - J_\lambda, t)\nu(dJ_\lambda)$$

The stationary distribution is the solution of

[6]Strictly speaking, it is the density that is positive, not the probability. There is still a debate whether one should stick with such so called Feller condition when using CIR process.

$$0 = -\frac{\partial}{\partial\lambda}\kappa(\theta-\lambda)f(\lambda) - \lambda f(\lambda) + P(J_x<0)\int(\lambda-J_\lambda)f(\lambda-J_\lambda)\nu(dJ_\lambda)$$

Applying Fourier transform, we have,

$$0 = \omega\left[\kappa\theta F(\omega) - \kappa F'(\omega)\right] - P(J_x<0)\left[F'(\omega) - F'(\omega)\int e^{J_\lambda\omega}\nu(dJ_\lambda)\right]$$

The ODE to be solved is,

$$(\frac{P(J_x<0)\eta\omega}{1-\eta\omega} - \kappa\omega)F'(\omega) = -\omega\kappa\theta F(\omega)$$

The solution is,

$$F(\omega) = e^{\theta\omega}\left(1 - \frac{\eta\kappa}{\kappa-\eta P(J_x<0)}\omega\right)^{-\theta P(J_x<0)/\kappa}$$

which corresponds to a Gamma distribution, $\Gamma(\frac{\theta P(J_x<0)}{\kappa}, \frac{\eta\kappa}{\kappa-\eta P(J_x<0)})$ shifted with θ. Unlike normal Gamma distribution, which is supported in $(0,\infty)$, the shifted Gamma distribution here is supported in (θ,∞).

Chapter 7

Conclusion

This thesis investigates three different topics in finance; portfolio selection in the presence of jumps, time varying variance risk premium, and correlation term structure of VIX futures with consistent pricing for two volatility surfaces. The key contributions are as follows:

First, we provide a non-parametric framework for portfolio selection when assets can have systemic co-jumps as well as independent idiosyncratic jumps. First, a MCMC method is used to estimate jumps in the weekly returns of eleven MSCI stock market index return which are categorized in two groups; developed and emerging markets groups. The estimation of eleven univariate jump components are used to study their dependence structure and portfolio optimisation. The evidence of co-jump is found to be stronger in the developed markets group; while idiosyncratic jumps are more common in the emerging markets group. The frequency of jumps and jump size are severely mis-estimated under the strict assumption of jump dependence in existing literature. The bias is stronger for emerging markets, as they are prone to large idiosyncratic jumps. Developed markets, are more harmonised, the assumption of perfectly correlated jumps is less damaging. The omission of idiosyncratic jumps has less impact on portfolio that consists of only developed markets.

The second contribution of my work is the finding of the connection between variance risk premium and realized jumps. I proposed a stochastic variance with self exciting jump model to explain the time varying risk premium. Instead of using option prices directly, I track the cumulants implied by options and I distinguish the different risk premium attached to the stochastic variance and the self exciting jump. It is shown that the stochastic volatility fits market dynamics very well in the calm

period and explains most of the time varying risk premium. However, in the crisis period, the market behaves in a different way. The high skewness of the implied volatility, especially for short maturities, can only be explained by the self-exciting jump component. In the joint estimation based on market return as well as option prices, jumps are not only determined by a abrupt big movement in the stock return, but also driven by abnormal change in option prices, which reflect investor's change in expectation. A jump in our model is defined as such events that make investors change their expectations. The self exciting jump feature in our model can explain very well the abrupt change of skewness in the crisis period, which fit investors' behaviour during crisis.

In my third piece of works, I show that one factor model may fit the implied volatility surface perfectly, but implies perfect correlation among VIX futures with different maturities. I further proposed two multi-factor models, both of which are able to fit the SPX and VIX implied volatility surfaces well. Based on the calibrated model parameters, we are able to calculate the implied term structure of correlation between VIX spot and VIX future, which is very important for hedging. Since the implied correlation term structure is conditional on the current state of factors and changes every day, I further derived the stationary distributions of the factors, which allow me to calculate the unconditional correlation term structure in order to compare it with the market realised correlation. In the last part, I also demonstrated how to calculate the optimal hedge ratio for hedging VIX option with VIX future.

There are still many unanswered questions left for future research. For the portfolio optimisation, one might consider stochastic volatility model to take into account the volatility clustering effect. Furthermore, if time varying jump intensity is allowed, the dependence between jump intensity can be another interesting way to model the dependence of jump processes in the multivariate setting. For the research on variance risk premium, I only study the variance dynamics and risk premium of options with 2 and 3 months maturities. It will be useful to examine how the long term variance changes over time. A stochastic long run mean might be a potential third factor. Last but not least, one could calculate the unconditional correlation term structure using the stationary distribution of state variables under the physical measure, since the observed dynamics of the state variables is under physical measure. Again, a third factor might be necessary for a better fit of the VIX future correlation term structure.

Bibliography

Aase, K. K. (1984). Optimum portfolio diversification in a general continuous-time model. *Stochastic Processes and their Applications*, 18(1):81–98.

Aggarwal, R., Inclan, C., and Leal, R. (1999). Volatility in emerging stock markets. *The Journal of Financial and Quantitative Analysis*, 34(1):33–55.

Ait-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2010). Modeling financial contagion using mutually exciting jump processes. Technical report, National Bureau of Economic Research, Inc.

Ang, A. and Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies*, 15(4):1137–1187.

Ang, A. and Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494.

Asgharian, H. and Bengtsson, C. (2006). Jump spillover in international equity markets. *Journal of Financial Econometrics*, 4(2):167–203.

Asgharian, H. and Nossman, M. (2011). Risk contagion among international stock markets. *Journal of International Money and Finance*, 30(1):22–38.

Aït-Sahalia, Y., Cacho-Diaz, J., and Hurd, T. R. (2009). Portfolio choice with jumps: A closed-form solution. *Annals of Applied Probability*, 19:556–584.

Bakshi, G. and Madan, D. (2006). A theory of volatility spreads. *Management Science*, 52(12):1945–1956.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Bates, D. S. (1996). Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics*, 94(1-2):181–238.

Bergomi, L. (2004). Smile dynamics i. *Risk*.

Bergomi, L. (2005). Smile dynamics ii. *Risk*.

Bergomi, L. (2007). Smile dynamics iii. *Risk*.

Bradley P. Carlin, Nicholas G. Polson, D. S. S. (1992). A monte carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association*, 87:493–500.

Carr, P. and Madan, D. B. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4):61–73.

Carr, P., Stanley, M., and Madan, D. (1998). Towards a theory of volatility trading. pages 417–427.

Carr, P. and Sun, J. (2007). A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, 10(2):87–150.

Carr, P. and Wu, L. (2008). Leverage effect, volatility feedback, and self-exciting market disruptions: Disentangling the multi-dimensional variations in s&p 500 index options. *SSRN eLibrary*.

Carr, P. and Wu, L. (2009). Variance risk premiums. *Review of Financial Studies*, 22(3):1311–1341.

CBOE (2009). The cboe volatility index - vix. Technical report.

Cont, R. and Da Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60.

Das, S. R. and Uppal, R. (2003). Systemic risk and international portfolio choice. Technical report, Santa Clara University and London Business School.

Das, S. R. and Uppal, R. (2004). Systemic risk and international portfolio choice. *Journal of Finance*, 59(6):2809–2834.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999). More than you ever wanted to know about volatility swaps. Quantitative strategies research notes, Goldman Sachs.

Doucet, A., de Freitas, N., and Gordon, N., editors (2001). *Sequential Monte Carlo Methods in Practice (Statistics for Engineering and Information Science)*. Springer, 1 edition.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical Finance*, 5(1):13–32.

Duan, J.-C. and Yeh, C.-Y. (2011). Price and volatility dynamics implied by the vix term structure. Technical report, National University of Singapore.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376.

Eraker, B. (1998). Mcmc analysis of diffusion models with application to finance. *Journal of Business and Economic Statistics*, 19:177–191.

Eraker, B. (2004). Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, 59(3):1367–1404.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance*, 58(3):1269–1300.

Fang, F. and Oosterlee, C. W. (2008). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31:826–848.

Fouque, J., Papanicolaou, G., Sircar, R., and Sølna, K. (2011). *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Mathematics, Finance and Risk Series. Cambridge University Press.

Fulop, A., Li, J., and Yu, J. (2012). Bayesian learning of impacts of self-exciting jumps in returns and volatility. Working Papers 03-2012, Singapore Management University, School of Economics.

Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. In *Global Derivatives & Risk Management*.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *Iee Proceedings F Radar and Signal Processing*, 140.

Grunbichler, A. and Longstaff, F. A. (1996). Valuing futures and options on volatility. *Journal of Banking & Finance*, 20(6):985–1001.

Guidolin, M. and Timmermann, A. (2008). International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies*, 21(2):889–935.

Guo, D. (1998). The risk premium of volatility implicit in currency options. *Journal of Business and Economic Statistics*, 16:498–507.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343.

Hürzeler, M. and Künsch, H. R. (1998). Monte carlo approximations for general state-space models. *Journal Of Computational And Graphical Statistics*, 7(2):175–193.

Jeanblanc-Picque, M. and Pontier, M. (1990). Optimal portfolio for a small investor in a market model with discontinuous prices. *Applied Mathematics and Optimization*, 22(3):287–310.

Jondeau, E. and Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management*, 12(1):29–55.

Josep Perelló, R. S. and Masoliver, J. (2008). Option pricing under stochastic volatility: the exponential ornstein-uhlenbeck model. *Journal of Statistical Mechanics: Theory and Experiment*.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

Lu, Z. and Zhu, Y. (2009). Volatility components: The term structure dynamics of vix futures. *Journal of Futures Markets*, 30:230–256.

Maheu, J. M. and McCurdy, T. H. (2004). News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance*, 59(2):755–793.

Malik, S. and Pitt, M. K. (2011). Particle filters for continuous likelihood evaluation and maximisation. *Journal of Econometrics*, 165(2):190–209.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63:3–50.

Psychoyios, D., Dotsis, G., and Markellos, R. (2010). A jump diffusion model for vix volatility options and futures. *Review of Quantitative Finance and Accounting*, 35(3):245–269.

Pukthuanthong, K. and Roll, R. (2012). Internationally correlated jumps. Technical report, European Central Bank.

Sepp, A. (2003). Fourier transform for option pricing under affine jump-diffusions: An overview. *SSRN eLibrary*.

Sepp, A. (2008). Vix option pricing in a jump-diffusion model. *Risk*.

Todorov, V. (2010). Variance risk-premium dynamics: The role of jumps. *Review of Financial Studies*, 23(1):345–383.

Zhang, J. E. and Zhu, Y. (2006). Vix futures. *Journal of Futures Markets*, 26:521–531.