Option Pricing with Stochastic Volatility Models

Jungyeon Yoon

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research (Statistics).

Chapel Hill
2008

Approved by

Advisor: Chuanshu Ji

Advisor: Eric Renault

Reader: Amarjit Budhiraja

Reader: Vladas Pipiras

Reader: Ron Gallant

ABSTRACT

JUNGYEON YOON: Option Pricing with Stochastic Volatility Models
(Under the direction of Chuanshu Ji and Eric Renault)

Despite the success and the user-friendly features of Black-Scholes (BS) pricing, many empirical results in the option pricing literature have shown the departures from the BS model. The motivation of this dissertation starts from these departures.

In the first part of dissertation, we take the popular approach of stochastic volatility and jump models that are known to give good explanations to the empirical phenomenon. In order to keep analytic tractability, we derive the Generalized Black-Scholes (GBS) formula by a proper conditioning in a general mixture framework. By taking advantage of this new version of option pricing formula, we propose an approximation scheme that is well suited for the conditional Monte Carlo method. The simulation study and Markov Chain Monte Carlo (MCMC) algorithm give an evidence of a huge computational time reduction without much loss of accuracy.

In the second part, we provide a new prospective on the forecasting ability and information content of the BS implied volatility in the presence of nonzero leverage effect. The leverage effect, which is the correlation between the return and volatility process, is introduced to model the observed Black-Scholes implied volatility (BSIV) smile and its skewness. We provide a simple theoretical framework that explains and justifies the use of BSIV from at-the-money option for the volatility forecast. Based on this and simulation study, which show the sensitivity of the concavity of option price with respect to the underlying stock price (the “gamma effect”), we propose a new approach to improve option pricing accuracy by a proper account for the gamma effect.

ACKNOWLEDGEMENTS

I would especially like to thank my advisors, Chuanshu Ji and Eric Renault for their generous advice and support, for stimulating discussions and for sharp comments. It has been a privilege working with them.

I wish to express my sincere gratitude to the members of my committee: Amarjit Budhiraja, Vladas Pipiras, and Ron Gallant for their suggestions that led me to improve this work.

I have been fortunate to work with SAS Stat Studio group as a research assistant for past three years. I have truly enjoyed interacting with each one of them. I greatly appreciate support and encouragement from Bob Rodriguez and Rick Wicklin.

I am the most grateful to my fiance, Keejeong Bae, for his endless support and love during all the stage of my time at Chapel Hill and for his patience when we had to spend five years apart.

I am deeply indebted to my family, who has always supported me, especially to my parents and Keejeong's parents for their love and prayers.

I thank all faculty members and fellow graduate students at UNC Chapel Hill Statistics department for my great experience. I also thank all my friends for helping me out in many ways. Special thanks to Bob and Makiko Humphreys, Sandra Campo-Manton, Sonja Murrell, Brandy Boe, Hyunsoo Jin and Seoyoung Park for friendship, understanding and laughter.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

An approximation scheme for option pricing with stochastic volatility and jump models

1.1 Introduction

In the empirical option pricing literature, departures from the Black and Scholes (BS) model are often explained by the importance of both stochastic volatility and jumps in stock returns. More often than not, this observation leads empirical researchers to give up the tractability of the BS formula to compute option prices and price sensitivities. Instead, we argue in this paper that most of the user-friendly features of BS pricing and hedging can be kept thanks to proper conditioning on the future path of relevant state variables. Our approach can actually be seen as a way to revisit and to correct a natural strategy of mixture of BS formulas that is quite popular among practitioners. Typically, to account for excess kurtosis and skewness in stock log-returns, a fast empirical approach amounts to consider that the option price is given by a weighted average:

$$\alpha BS(\sigma_1) + (1-\alpha) BS(\sigma_2), \tag{1.1}$$

where $BS(\sigma)$ denotes the BS option price computed with the value for the volatility parameter σ. The rationale for (1.1) is to consider that a mixture of two normal distributions with standard errors σ_1 and σ_2 and weights α and $(1-\alpha)$ respectively may account for both skewness and excess kurtosis. The problem with this naive approach is that it does not take into account any risk premium associated to the mixture component.

A convenient way to specify various kinds of risk premiums is to introduce a Stochastic Discount Factor (SDF) $m_{t,T}$ (see Hansen and Richard (1987)) such that the price p_t paid at time t for a terminal payoff g_T at time $T > t$ is given by:

$$p_t = E[m_{t,T} g_T | F_t], \tag{1.2}$$

where F_t stands for the information available to investors at time t. Then, it is well known (see e.g. Garcia *et al.* (2007)) that the BS option pricing formula for an European call written on a stock with price S_T at time T (without dividend) is valid insofar as the joint conditional distribution given F_t of $[\log(m_{t,T}), \log(S_T)]$ is normal. More generally, if we want to accommodate a mixture of normal distributions with a mixing variable $U_{t,T}$, we can write:

$$p_t = E[E(m_{t,T} g_T | F_t, U_{t,T}) | F_t], \tag{1.3}$$

where, for each possible value $u_{t,T}$ of $U_{t,T}$, the BS formula is valid to compute:

$$E[m_{t,T} g_T | F_t, U_{t,T} = u_{t,T}].$$

In other words, it is true that, as in (1.1), the final conditional expectation operator (given F_t) in (1.3) displays the option price as a weighted average of different BS prices with the weights corresponding to the probabilities of the possible values $u_{t,T}$ of the mixing variable $U_{t,T}$. However, the naive approach (1.1) is applied in a wrong way when forgetting that the additional conditioning information $U_{t,T} = u_{t,T}$ should lead to modify some key inputs in the BS option pricing formula.

Suppose that investors are told that the mixing variable $U_{t,T}$ will take the value $u_{t,T}$. Then, the current stock price would no longer be:

$$S_t = E[m_{t,T} S_T | F_t],$$

but

$$S_t^*(u_{t,T}) = E[m_{t,T} S_T | F_t, U_{t,T} = u_{t,T}]. \tag{1.4}$$

For the same reason, the pure discount bond which delivers \$1 at time T will no longer be priced at time t:

$$B(t,T) = E[m_{t,T}|F_t],$$

but

$$B^*(t,T)(u_{t,T}) = E[m_{t,T}|F_t, U_{t,T} = u_{t,T}]. \tag{1.5}$$

In other words, various BS option prices which are averaged in a mixture approach like (1.1) must be computed, no longer with actual values $B(t,T)$ and S_t of the current bond and stock prices, but with values $B^*(t,T)(u_{t,T})$ and $S_t^*(u_{t,T})$ not directly observed but computed from (1.4) and (1.5). In particular, the key inputs, underlying stock price and interest rate, should be different in various applications of the BS formulas like $BS(\sigma_1)$ and $BS(\sigma_2)$ in (1.1).

This remark is crucial for the conditional Monte Carlo approach put forward in this paper. We extend in this respect the work of Willard (1997) in the context of option pricing with stochastic volatility. Revisiting a formula initially derived by Romano and Touzi (1997), Willard (1997) notes that the variance reduction technique, known as conditional Monte Carlo, can be applied even when the conditioning factor (the stochastic volatility process) is instantaneously correlated with the stock return as it is the case when leverage effect is present. He stresses that "by conditioning on the *entire path* of the noise element in the volatility (instead of just the average volatility), we can still write the option's price as an expectation over Black-Scholes prices by appropriately adjusting the arguments to the Black-Scholes formula" However, Willard (1997) does not note that the "appropriate adjustment" of the stock price as an argument of the BS formulas is precisely akin to the replacement of S_t by $S_t^*(U_{t,T})$ given by (1.4). Moreover, he does not appropriately adjust the interest rate according to (1.5) and works with a fixed risk neutral distribution. We will see in Section 2 that this approach is in general flawed when the short term interest rate is stochastic.

The contribution of this paper is threefold. First, we correct and generalize the mixture/conditional extension of BS option pricing by characterizing the appropriate adjustment of both bond and stock prices. While our adjustment of stock price is just a generalization of Willard (1997) to accommodate jumps, the adjustment of bond price is akin to a proper conditioning of the risk neutral distribution. This resulting generalization of BS pricing, dubbed Generalized Black-Scholes (GBS), has already been put forward by Garcia, Luger and Renault (2003) in the context of stochastic volatility. Our second contribution is to extend GBS pricing to a general model of stochastic volatility and jumps. This extension is important because it displays the relevant concept of average volatility over the lifetime of the option. In this respect, it paves the way for new applications of the recent techniques (see e.g. Barndorff-Nielsen and Shephard (2004)) for separate identification of the two components of quadratic variation, namely (continuous) integrated variance and sum of squared jumps. Finally, we take advantage of this new version of GBS option pricing to design an approximation scheme well suited for conditional Monte Carlo simulation of option prices under a popular stochastic volatility model in Heston (1993). Our approximation takes a spirit from Cheng, Gallant, Ji and Lee (2007) and Broadie and Kaya (2006). The approximation scheme is illustrated through its usefulness to implement Markov Chain Monte Carlo (MCMC) estimation procedures. Since our model is at least as general as the stochastic volatility models in the common literature (e.g. Bates (2000), Pan (2002)), the computational efficiency of our procedure must be compared with the available competitors. Even though the extensive comparison is still in progress, we give evidence of huge computational time reduction without much to pay in terms of accuracy.

The chapter is organized as follows. Section 1.2 sets up a general mixture framework and develops an adjusted bond and stock prices. Section 1.3 provides the GBS option pricing formula. Section 1.4 introduces an approximation scheme for GBS formula. Section 1.5 illustrates the MCMC estimation. Technical details, including proofs of all lemmas and propositions, are collected in the appendices.

1.2 Adjustment of bond and stock prices in a general mixture framework

1.2.1 Proper conditioning of the risk-neutral distribution

In a continuous time framework, the SDF $m_{t,T}$ is generally seen as the relative increment of a pricing kernel process:

$$m_{t,T} \quad = \quad \pi_T/\pi_t.$$

As already announced, the key idea of the mixture model is to define a conditioning variable $U_{t,T}$ such that the pricing kernel process follows a geometric Brownian motion under the conditional probability distribution given $U_{t,T}$. The mixing variable $U_{t,T}$ will typically show up as a function of a state variable path $(X_u)_{t\leq u\leq T}$. More precisely, let us assume that the pricing kernel process is defined by:

$$d\log(\pi_t) \quad = \quad h(X_t)dt + a(X_t)dW_{1t} + b(X_t)dW_{2t} + c_t dN_t, \tag{1.6}$$

where:

(W_{1t}, W_{2t}) is a two-dimensional standard Brownian motion,

(N_t) is a Poisson process with intensity $\lambda(X_t)$ depending on the state variable X_t, and the jumps sizes c_t are i.i.d. normal variables independent of the state variable process (X_t).

The model (1.6) is devised such that, given the state variables path $(X_u)_{t\leq u\leq T}$, the log of SDF $\log(m_{t,T})$ is normally distributed insofar as the number $(N_T - N_t)$ of jumps which occurred in the time interval $(t, T]$ is given. In order to introduce some instantaneous correlation between the state variable and the pricing kernel process, we also include the first Brownian motion W_{1t} as a component of X_t:

$$X_t \quad = \quad (Z_t, W_{1t}), \tag{1.7}$$

where Z_t is a vector of additional covariates possibly needed to describe the stochastic time variations of the coefficients in (1.6). All the stochastic processes considered so far are assumed to be $(F_t)-$adapted where (F_t) is an increasing family of $\sigma-$fields, assumed to fulfill the usual hypotheses (see e.g. Protter (2004)) and to summarize the information available at time t to all investors. As usual, the drift of the pricing kernel process is tightly related to the short term interest rate. For sake of expositional simplicity, we have precisely assumed that this drift is a deterministic function of X_t (through the functions $h(\cdot)$ and $\lambda(\cdot)$) to be sure that the short term interest rate is a function $r(X_t)$ of the current state variables. More precisely, we are able to define the short term interest rate function $r(\cdot)$ by the standard condition:

$$1 = E\left[m_{t,T}\exp(\int_t^T r(X_u)du)|F_t\right]. \tag{1.8}$$

This is akin to say that:

$$B_t^\pi = \pi_t \exp\left(\int_0^t r(X_u)du\right)$$

is a $(F_t)-$martingale. From this martingale condition, it is easy to deduce that:

Lemma 1.2.1. *The short term interest rate is given by:*

$$r(X_t) = -h(X_t) - \frac{1}{2}\left[a^2(X_t) + b^2(X_t)\right] - \lambda(X_t)E[\exp(c_t) - 1].$$

Note that Lemma 1.2.1 actually characterizes a local martingale property but for all practical purposes we will assume that it ensures that B_t^π is a $(F_t)-$martingale. Within the information structure defined by the filtration (F_t), the variable with unit conditional expectation $m_{t,T}\exp(\int_t^T r(X_u)du)$ defines the Radon-Nikodym derivative of a risk-neutral distribution. This risk neutral distribution is valid to price at time t future payoffs at time T when the information available to all investors is F_t. Suppose now that investors are told at time t that the future path of state variables will be $(X_u)_{t\leq u\leq T}$. Then, this additional information must of course be used to update the pricing measure and, by the law of iterated expectations, a risk neutral distribution will now be given by its Radon-Nikodym derivative:

$$\eta_{t,T} = E[m_{t,T}|F_t \vee \sigma((X_u)_{t \le u \le T}), (N_u)_{t \le u \le T}] \exp\left(\int_t^T r(X_u)du\right).$$

The projected SDF $E[m_{t,T}|F_t \vee \sigma((X_u)_{t \le u \le T}), (N_u)_{t \le u \le T}]$ can actually be interpreted as a SDF valid to price the payoffs whose only randomness goes through the value (X_u) of state variables and (N_u), the jump times, where $t \le u \le T$. The advantage of conditional log-normality is to allow us to derive a closed-form formula for the updated risk neutral distribution:

Proposition 1.2.1. *The Radon-Nikodym derivative of a risk neutral distribution conditional on $F_t \vee \sigma((X_u)_{t \le u \le T})$, and $(N_u)_{t \le u \le T}$ is given by:*

$$\begin{aligned} \eta_{t,T} &= \exp\left[\int_t^T a(X_u)dW_{1u} - \frac{1}{2}\int_t^T a^2(X_u)du\right] \\ &\quad \exp\left[(N_T - N_t)\log\left[E\left(e^{c_t}\right)\right] - \left[E\left(e^{c_t}\right) - 1\right]\int_t^T \lambda(X_u)du\right] \end{aligned}$$

Note that, following the notation of the introduction, an alternative interpretation of this result is to say that the adjusted bond price $B^*_{t,T}$ is:

$$B^*_{t,T} = B^*_{t,T}(U_{t,T}) = \exp\left[-\int_t^T r(X_u)du\right]\eta_{t,T},$$

where the mixing variable $U_{t,T}$ summarizes the values of the relevant conditioning variables, namely:

$$\textstyle\int_t^T a(X_u)dW_{1u},\ \int_t^T a^2(X_u)du,\ \int_t^T \lambda(X_u)du,\ \int_t^T r(X_u)du \text{ and } (N_T - N_t).$$

Irrespective of the interpretation, the key issue is to correctly update the quantities of interest to account for the additional information $(X_u, N_u)_{t \le u \le T}$. For instance, even in a purely continuous stochastic volatility setting, the risk neutral distribution must be rescaled with the exponential martingale factor:

$$\exp\left[\int_t^T a(X_u)dW_{1u} - \frac{1}{2}\int_t^T a^2(X_u)du\right].$$

Omitting to do so would amount to assume that the part of the Brownian innovation that is instantaneously perfectly correlated to the conditioning state variables has a zero weight in the pricing kernel, which is quite a restrictive assumption about the sources of risk that are actually compensated. The restrictive content of such an assumption will become obvious if we consider, as in Willard (1997), that this instantaneous correlation corresponds to leverage effect in the stochastic volatility process of the stock return of interest. This assumption is made precise in the next subsection.

1.2.2 Adjustment of stock price

The key idea of GBS option pricing approach through iterated expectations as shown in Garcia *et al.* (2003) is to get a conditioning over the relevant state variables in order to be back to a geometric Brownian motion. In other words, the state variable X_t and N_t must summarize the randomness in the stock return volatility and in jump dates. Thus, we define a stock price process (S_t):

$$d(\log S_t) = \mu(X_t)dt + \alpha(X_t)dW_{1t} + \beta(X_t)dW_{2t} + \gamma_t dN_t, \tag{1.9}$$

where the jump sizes γ_t are i.i.d. normal variables independent of the state process (X_t). Consider

$$V_t = \alpha^2(X_t) + \beta^2(X_t),$$

which is interpreted as the square volatility process. The key is that conditioning by $F_t \vee \sigma((X_u)_{t\leq u\leq T})$ and $(N_u)_{t\leq u\leq T}$ will make the volatility path deterministic. Note also that a non-zero coefficient $\alpha(X_t)$ captures the instantaneous correlation between the continuous part of the price innovation and volatility. In other words,

$$\rho(X_t) = \frac{\alpha(X_t)}{\sqrt{\alpha^2(X_t) + \beta^2(X_t)}} = \frac{\alpha(X_t)}{\sqrt{V_t}}$$

can be interpreted as a leverage effect coefficient. The stock price equation is:

$$1 = E\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t}|F_t\right]$$

That is, the discounted price process $\varphi_t^\pi = \pi_t S_t$ is a F_t-martingale. From the martingale condition, it is deduced that

Lemma 1.2.2.

$$\mu(X_t) = r(X_t) - a(X_t)\alpha(X_t) - b(X_t)\beta(X_t) - \frac{1}{2}(\alpha^2(X_t) + \beta^2(X_t)) + \lambda(X_t)\left[E\left(e^{c_t}\right) - E\left(e^{c_t+\gamma_t}\right)\right],$$

For all practical purposes, we will assume throughout that the condition in Lemma 1.2.2 ensures that φ_t^π is a martingale. The drift coefficient $\mu(X_t)$ is actually not free but determined in Lemma 1.2.2. Following the notation given in the introduction, we can now state the second result for the adjusted stock price, S_t^*.

Proposition 1.2.2.

$$S_t^* = S_t^*(U_{t,T}) = E\left[\frac{\pi_T}{\pi_t} S_T | F_t \vee \sigma((X_u)_{t\le u\le T}), (N_u)_{t\le u\le T}\right] = S_t \xi_{t,T},$$

with:

$$\begin{aligned} \xi_{t,T} \quad = \quad & \exp\left[\int_t^T (a(X_u) + \alpha(X_u))dW_{1u} - \frac{1}{2}\int_t^T (a(X_u) + \alpha(X_u))^2 du\right] \\ & \exp\left[(N_T - N_t)\log(E[e^{c_t+\gamma_t}]) - E[e^{c_t+\gamma_t} - 1]\int_t^T \lambda(X_u)du\right]. \end{aligned}$$

Similarly, the mixing variable $U_{t,T}$ summarizes the values of the relevant conditioning variables, namely:

$\int_t^T (a(X_u) + \alpha(X_u))dW_{1u}$, $\int_t^T (a(X_u) + \alpha(X_u))^2 du$, $\int_t^T \lambda(X_u)du$ and $(N_T - N_t)$.

1.2.3 An equilibrium pricing interpretation

Asset pricing models typically use either no-arbitrage arguments or equilibrium to derive explicit pricing formula. Equilibrium approach is often built upon certain general equilibrium models, which balance consumption and investment, and incorporate preferences towards various risk factors by defining a utility function. Given a utility function, the price of an asset would come as a part of the solution to the optimization problem. The approach provides a better understanding for the preference of an investor or agent, especially the risk premiums can be interpreted more explicitly. However, a general equilibrium model tends to make empirical work (fitting a model based on real financial data) more difficult due to strict restrictions imposed by the framework. In contrast, no-arbitrage approach determines the value of a derivative (e.g. option, future) through its relation to other assets (e.g. stock, bond or other derivatives) whose prices are taken as given. Black-Scholes theory is a typical example. It is a practical way to price assets, and easy to link to real data, but does not offer the insight of an agent's preference as absolute pricing does. In this section, we like to show that the specification of risk premium in the previous section reconciles with the equilibrium approach.

Consider an economy in which an investor tries to solve a utility maximization problem $\max_c E\left[\int_0^\infty J(c_t,t)dt\right]$, where $J(c_t,t)$ is the utility function of the consumption process (c_t). An endowment, which follows an exogeneous process, is available to the investor. In equilibrium, the investor finds it optimal to just consume the exogenous endowment; i.e. $c_t = \delta_t$. The economy is endowed with a stochastic flow of the consumption good. The equilibrium price, $p(t)$, must satisfy the Euler equation

$$p(t) = \frac{E(J_c(\delta_T,T)p(T)|F_t)}{J_c(\delta_t,t)}, \tag{1.10}$$

where J_c is the partial derivative of $J(c,t)$ with respect to c. We assume the endowment (δ_t) follows the similar process to the stock price S_t:

$$d(\log \delta_t) = \mu_\delta(X_t)dt + \alpha_\delta(X_t)dW_{1t} + \beta_\delta(X_t)dW_{2t} + \gamma_t^\delta dN_t. \tag{1.11}$$

For simplicity, we adopt a utility function in Naik and Lee (1990). The utility function has a form:

$$J(c,t) = \exp(-\theta t)\frac{c^\omega}{\omega},$$

where $0 < \omega < 1$. If $\omega = 0$, then $J(c,t) = \exp(-\theta t)\log(c)$. It exhibits a constant relative risk aversion. Since $J_c(\delta_t,t) = \exp(-\theta t)\delta_t^{\omega-1}$, a pricing kernel is, from (1.10),

$$\pi_t = \exp(-\theta t)\delta_t^{\omega-1}.$$

By Ito's formula, the pricing kernel process is

$$d(\log \pi_t) = (-\theta + (\omega - 1)\mu_\delta)\, dt + (\omega - 1)\left(\alpha_\delta dW_{1t} + \beta_\delta dW_{2t}\right) + (\omega - 1)\gamma_t^\delta dN_t \qquad (1.12)$$

Lemma 1.2.3. *The equivalence of Lemma 1.2.1 for (1.12) is*

$$r(X_t) = \theta - (\omega - 1)\mu_\delta(X_t) - \frac{1}{2}(\omega - 1)^2\left(\alpha_\delta^2(X_t) + \beta_\delta^2(X_t)\right) - \lambda(X_t)E\left[\exp((\omega-1)\gamma_t^\delta) - 1\right].$$

Lemma 1.2.4. *The equivalence of Lemma 1.2.2 is*

$$\begin{aligned}\mu(X_t) &= r(X_t) - (\omega - 1)\left(\alpha_\delta(X_t)\alpha(X_t) + \beta_\delta(X_t)\beta(X_t)\right) - \frac{1}{2}\left(\alpha^2(X_t) + \beta^2(X_t)\right) \\ &\quad + \lambda(X_t)\left[E(\exp((\omega-1)\gamma_t^\delta) - \exp((\omega-1)\gamma_t^\delta + \gamma_t)\right]\end{aligned}$$

By Lemma 1.2.3, the equilibrium pricing kernel is

$$\begin{aligned}d(\log \pi_t) &= \left[-r(X_t) - \frac{1}{2}(\omega-1)^2 - \lambda(X_t)E\left(\exp((\omega-1)\gamma_t^\delta - 1\right)\right] dt \\ &\quad + (\omega-1)\left(\alpha_\delta(X_t)dW_{1t} + \beta_\delta(X_t)dW_{2t}\right) + (\omega-1)\gamma_t^\delta dN_t.\end{aligned}$$

The equilibrium pricing kernel will be used in pricing bond, stock and option prices. Again, by a proper conditioning by $F_t \vee \sigma((X_u)_{t\leq u\leq T})$ and $(N_u)_{t\leq u\leq T}$, we are back to Brownian Motion. Thus, the joint probability distributions of $\left[\log\left(\frac{\delta_T}{\delta_t}\right), \log\left(\frac{S_T}{S_t}\right)\right]$ and furthermore $\left[\log\left(\frac{\pi_T}{\pi_t}\right), \log\left(\frac{S_T}{S_t}\right)\right]$ are bivariate normal. We can derive a GBS option pricing formula similarly to the previous setting.

Proposition 1.2.3. *The adjusted bond price is:*

$$\begin{aligned}B_{t,T}^* &= E\left[\frac{\pi_T}{\pi_t}|F_t \vee \sigma((X_u)_{t\leq u\leq T})\right] \\ &= \exp\left[-\int_t^T r_u du\right]\eta_{t,T},\end{aligned}$$

with

$$\eta_{t,T} = \exp\left[(\omega-1)\int_t^T \alpha_\delta(X_u)dW_{1u} - \frac{1}{2}(\omega-1)^2\int_t^T \alpha_\delta^2(X_u)du\right]$$

$$\exp\left[(N_T - N_t)\log\left[E\left(\exp((\omega-1)\gamma_t^\delta)\right)\right] - \left[E\left(\exp((\omega-1)\gamma_t^\delta) - 1\right)\right]\int_t^T \lambda(X_u)du\right].$$

From Lemma 1.2.4, we get the adjusted stock price.

Proposition 1.2.4. *The adjusted stock price:*

$$S_t^* = E\left[\frac{\pi_T}{\pi_t} S_T | F_t \vee \sigma((X_u)_{t\le u\le T})\right] = S_t \xi_{t,T},$$

with:

$$\begin{aligned}\xi_{t,T} &= \exp\left[\int_t^T ((\omega-1)\alpha_\delta(X_u) + \alpha(X_u))dW_{1u} - \frac{1}{2}\int_t^T ((\omega-1)^2\alpha_\delta(X_u) + \alpha(X_u))^2 du\right] \\ &\quad \exp\left[(N_T - N_t)\log[E(\exp(\gamma_t + (\omega-1)\gamma_t^\delta))] + E[\exp(\gamma_t + (\omega-1)\gamma_t^\delta) - 1]\int_t^T \lambda(X_u)du\right]\end{aligned}$$

The proofs of Proposition 1.2.3 and Proposition 1.2.4 are not included since they are basically the same as those of Proposition 1.2.1 and Proposition 1.2.2. Note that the preference parameter ω appear explicitly in the adjusted bond prices and stock price and therefore in the option pricing formula. Except for some factorization with ω, everything remains the same as no-arbitrage approach in Section 1.2.1 and Section 1.2.2 given specification of the endowment process. We do not discuss thorough study on the relationship between the level of risk aversion and risk premium along with the specification of endowment process since they are beyond the scope of our study.

1.3 The GBS option pricing formula

Except for a few cases, such as the stochastic volatility in Heston (1993), most dynamics of the state variables in derivative pricing do not yield closed form solutions. Monte Carlo simulation can be used to get a derivative price. Standard Monte Carlo simulation would generate many sample paths of state variables, evaluate the payoff of the derivative on each path, then take an average over them. The average gives an estimate of a derivative price. In this section, we focus on an European call option and derive a BS like option pricing formula by the proper conditioning. The conditional Monte Carlo approach has several advantages over competitors. First, it can be easily applicable in both pricing and calculating sensitivities (the "Greeks"). Second, there is a big improvement in computational efficiency. It is faster and gives a smaller variance than the traditional Monte Carlo method, as shown in Willard (1997).

Following the notation in the introduction, the payoff g_T of an European call option written on a stock price S_t is $g_T = \max(0, S_T - K)$, where K is a strike price. The proof of the option pricing formula in Proposition 1.3.1 is in Appendix E.

Proposition 1.3.1. *The price of an European call option has the form:*

$$C_t = E\left[\frac{\pi_T}{\pi_t}\max(0, S_T - K)|F_t\right],$$

where F_t is the information available at time t to the investor. We obtain the GBS option pricing formula for the stochastic volatility and jumps model:

$$\begin{aligned} C_t &= E[\widetilde{BS}(S_t\xi_{t,T}, (\overline{\sigma_{t,T}})^2)|F_t] \\ &= E[S_t\xi_{t,T}\Phi(d_1) - KB^*_{t,T}\Phi(d_2)|F_t], \end{aligned}$$

where

$\widetilde{BS}(\cdot,\cdot)$ *is a BS-like option pricing formula,*

$\xi_{t,T}$ *and* $B^*_{t,T}$ *defined as in Proposition 1.2.1 and Proposition 1.2.2(or Proposition 1.2.3 and Proposition 1.2.4),*

$$d_1 = \tfrac{1}{\overline{\sigma_{t,T}}}\left[\log(\tfrac{S_t\xi_{t,T}}{KB^*_{t,T}}) + \tfrac{1}{2}(\overline{\sigma_{t,T}})^2\right],$$

$$d_2 = d_1 - \overline{\sigma_{t,T}},$$

and

$$\overline{\sigma_{t,T}} = \sqrt{\int_t^T (1-\rho^2(X_u))V_u du + (N_T - N_t)Var[\gamma_i]}$$

Like the previous work by Willard (1997), Romano and Touzi (1997), this option pricing formula extends to more general cases and still preserves tractability that BS pricing formula has. The characteristics of SDF explicitly shows up in the option pricing formula. That is, we have nice interpretations on how risk premiums and/or preference parameters play a role in pricing and sensitivities. Also, it is worthwhile noting that $\overline{\sigma_{t,T}}$ displays separate identification of the two components of quadratic variation, namely (continuous) integrated variance and sum of squared jumps, which is similar to quadratic variation in Barndorff-Nielsen and Shephard (2004).

1.4 An approximation scheme for GBS formula

We propose an approximation scheme, which results in a significant dimension reduction by simulating certain key summary statistics in the model instead of generating the entire volatility path. Hence, it is particularly well suited to our GBS pricing formula and the related conditional Monte Carlo simulation. Cheng, Gallant, Ji and Lee (2007) demonstrates the promise of this approach by applying it to the MCMC estimation of a log-linear stochastic volatility model. Although the idea of the approximation scheme may be quite general, we implement it in a setting of the square-root volatility process, which enjoys a great popularity in the financial asset pricing. Our approximation scheme can be used in the stochastic volatility (SV) model and extended to the SV model with jumps (SVJ), as in Bates (2000) and Pan (2002). The jump times follows a Poisson process with volatility dependent intensity (λV_t).

Only for the certain specification of the GBS formula can be simply rewritten as

$$\begin{aligned} C_t &= E[\widetilde{BS}(S_t\xi_{t,T}, (\overline{\sigma_{t,T}})^2 | (W_{1\tau})_{t\le\tau\le T}, N_T - N_t)] \\ &= E\left[\widetilde{BS}\left(S_t\xi_{t,T}, (\overline{\sigma_{t,T}})^2 | \int_t^T V_u du, \int_t^T \sqrt{V_u} dW_{1u}, N_T - N_t\right)\right]. \end{aligned}$$

Note that $N_T - N_t$ is a Poisson random variable with an intensity parameter $(\lambda \int_t^T V_u du)$ given $\int_t^T V_u du$, in SVJ model. Thus, the conditional Monte Carlo simulation to calculate an option price with the GBS formula is done by repeating the following steps many times:

step-(i) By simulation, get two volatility integrals, $\int_t^T V_u du$ and $\int_t^T \sqrt{V_u} dW_u$.

step-(ii) Generate a sample of $N_T - N_t$ from Poisson distribution with a parameter, $\lambda \int_t^T V_u du$.

step-(iii) Place the simulated values $\int_t^T \sqrt{V_u} dW_{1u}$, $\int_t^T V_u du$ and $N_T - N_t$ in the the GBS formula.

Note that the computational efficiency and the accuracy of the option price calculation heavily depend on generation of two volatility integrals, $\int_t^T V_u du$ and $\int_t^T \sqrt{V_u} dW_u$. These integrals are high-dimensional and create a significant computational challenges. For instance, MCMC estimation of the model parameters, which is illustrated in the next section, repeatedly needs to calculate option prices. Moreover, extra option calibration and time are required if the Metropolis-Hastings algorithm in MCMC estimation for the nonstandard conditional posteriors cannot be avoided when both option and return data are used. Thus, it is crucial for the performance of MCMC that these integrals can be obtained fast.

The square-root process (V_t) has a form of:

$$dV_t = k_v(\theta - V_t)dt + \sigma_v \sqrt{V_t} dW_{1t}. \tag{1.13}$$

θ is the long-run mean, k_v represents the mean reversion, and σ_v is a parameter which determines the variance of the process. Under the assumption $2k_v\theta > \sigma_v^2$, V_t process is always positive. The square-root process take an important role in many financial models, including the interest rate model in Cox, Ingersoll and Ross (1985), stochastic volatility model in Heston (1993). It has been widely used in derivative pricing models since the reformulation of the original Fourier integrals made computations of European option prices numerically stable and efficient. Yet, many practical applications of models with this dynamics, which involve the pricing and hedging, require the introduction of Monte Carlo method. Since our GBS option pricing formula is under the scope of the conditional Monte Carlo method, we start with an illustration of the method. We outline the existing discretization schemes from the literature designed for the square-root volatility process, for later comparative experiments. Then, we introduce our approximation scheme, followed by numerical comparisons.

1.4.1 Path simulation schemes

The simple Euler's discretization does not guarantee the positivity of the process. There have been research into the simulation scheme for the squared volatility process.

First, the volatility path can be generated by using the exact transition law of process. The distribution of V_t given $V_s, s < t$ is up to a scale factor, a noncentral chi-square distribution. Details of this method are illustrated in Glasserman (2003). This method is rarely used in practice because of the computational inefficiency.

Roger (1995) proposed the method using multi-dimensional Ornstein-Uhlenback process. The drawback of this method is that we need to impose extra constraints on the parameters in the square-root volatility process. Kahl and Jackel (2005) propose an application of an implicit Milstein scheme:

$$V_{t+\Delta} = \frac{V_t + k_v\theta\Delta + \sigma_v\sqrt{V_t}\varpi_v\sqrt{\Delta} + \frac{1}{4}\sigma_v^2\Delta(\varpi_v^2 - 1)}{1 + k\Delta},$$

where ϖ_v is a standard normal random variable. It requires a restriction on volatility parameters, $4k_v\theta > \sigma_v$, which is rarely satisfied in practice. There is another scheme suggested by Brigo and Alfonsi (2005):

$$V_{t+\Delta} = \left[\frac{\sigma_v\varpi_v\sqrt{\Delta} + \sqrt{\sigma_v^2\varpi_v^2\Delta + 4(V_t + (k_v\theta - \sigma^2/2)\Delta)(1 + k_v\Delta)}}{2(1 + k_v\Delta)}\right]^2.$$

It does not require more than $2k_v\theta > \sigma_v^2$, which should be assumed to preserve a positivity of volatility process. In addition, this method has nice convergence property(refer to Alfonsi (2005)), which guarantees its small bias if we use a very small time interval. For these reasons, we choose the simulation scheme by Brigo and Alfonsi (2005) to evaluate the performance of our proximation scheme in Section 1.4.4. All these methods show slightly different behaviors and convergence properties. In any case, if we use one of these methods, we cannot avoid simulating whole paths and its computation cost is high, especially in MCMC estmation when we need many iterations.

1.4.2 Proposed approximation scheme

On the contrary to the path simulation schemes, the scheme proposed by Broadie and Kaya (2006) allows us to avoid generating a whole path and reduce bias significantly. Since the conditional distribution of $\int_t^T V_u du$ is not known in a close form, they use the numerical inverse algorithm for the distribution, which turns out too complex and slow. We take some essential parts of their scheme and propose a much simpler scheme. We propose an approximation scheme on the conditional distribution of $\int_t^T V_u du$. Simply, add three more steps to step$-(i)$ in the previously illustrated conditional Monte Carlo method. Other steps remain the same.

step-(i-1) Generate a sample from the distribution of V_T given V_t: using non-central chi-squared distribution.

step-(i-2) Generate a sample from the approximated distribution of $\int_t^T V_u du$ given V_t.

step-(i-3) Recover $\int_t^T \sqrt{V_u} dW_{1u}$ from given V_T, V_t, and $\int_t^T V_u du$.

The proposed scheme is an approximation with two reasons. First, the gamma distribution for $\int_t^T V_u du$ is not exact. Second, V_T and $\int_t^T V_u$ are simulated separately, ignoring the correlation between the two. More details of the simulation steps are in the following subsections.

Generating V_T given V_t

Based on the results of Feller (1971) and Cox, Ingersoll and Ross (1985), the exact transition law of the square-root process is known. The distribution of V_T given V_t for some $t < T$ is a non-central chi-squared distribution up to a scale factor.

$$V_T = \frac{\sigma_v^2(1 - \exp(-k_v(T-t)))}{4k_v} \chi_{p_1}'^2(p_2),$$

where $\chi_{p_1}'^2(p_2)$ denotes the non-central chi-square random variable with p_1 degrees of freedom, and non-centrality parameter p_2,

$$p_1 = \frac{4\theta k_v}{\sigma_v^2},$$

and

$$p_2 = \frac{4k_v \exp(-k_v(T-t))V_t}{\sigma_v^2(1-\exp(-k_v(T-t)))}$$

Generating $\int_t^T V_u du$ given V_t from the proposed approximation

We observe that $\int_t^T V_u \, du$ has a skewed distribution. Since $\int_t^T V_u \, du$ is always positive, we propose the gamma distribution as an approximation. Figure 1.1 shows how the skewed distribution of $\int_t^T V_u \, du$ is well fitted with the gamma distribution. From Garcia, Lewis and Renault (2001), we know the explicit form of the conditional mean and variance of $\int_t^T V_u \, du$. These two moments give the parameter values of the gamma distribution by the method of moments.

Proposition 1.4.1. *The conditional mean is*

$$E\left[\int_t^T V_u du | F_t\right] = \;=\; V_t A_{t,T} + B_{t,T},$$

where

$$\begin{aligned} A_{t,T} &= \frac{1}{k_v}\left(1 - e^{-k_v(T-t)}\right) \\ B_{t,T} &= \theta(T-t) - \frac{\theta}{k_v}\left(1 - e^{-k_v(T-t)}\right). \end{aligned}$$

The conditional variance is obtained as

$$Var\left[\int_t^T V_u du | F_t\right] = V_t C_{t,T} + D_{t,T},$$

where

$$\begin{aligned} C_{t,T} &= \frac{\sigma_v^2}{k_v^2}\left[\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right] \\ D_{t,T} &= \frac{\sigma_v^2\theta}{k_v^2}\left[(T-t)\left(1 + 2e^{-k_v(T-t)}\right) + \frac{1}{2k_v}\left(e^{-k_v(T-t)} + 5\right)\left(e^{-k_v(T-t)} - 1\right)\right] \end{aligned}$$

The proof of Proposition 1.4.1 is in Appendix F.1. Using the conditional mean and variance, we can calculate the parameters in the gamma distribution. The scale parameter is expressed as $\frac{Var\left[\int_t^T V_u du|F_t\right]}{E\left[\int_t^T V_u du|F_t\right]}$ and the shape parameter as $\frac{E^2\left[\int_t^T V_u du|F_t\right]}{Var\left[\int_t^T V_u du|F_t\right]}$. We are ready to sample $\int_t^T V_u du$.

Recover $\int_t^T \sqrt{V_u} dW_{1u}$

$$\int_t^T \sqrt{V_u} dW_{1u} = \left(\frac{1}{\sigma_v}\right)\left(V_T - V_t - k_v\theta(T-t) + k_v \int_t^T V_u du\right),$$

is simply an integrated form of (1.13). Given V_T, V_t, and $\int_t^T V_u du$, we have $\int_t^T V_u du$.

1.4.3 Improved approximation scheme using bivariate gamma distribution

In the previous section, the approximation scheme generates $\int_t^T V_u du$ and V_T separately by ignoring the correlation between these two. First, we want to check whether the correlation is negligible by calculation and the simulation based on the estimates in the literature.

Proposition 1.4.2.

$$Corr\left[\int_t^T V_u du, V_T|F_t\right] = \frac{Cov\left[\int_t^T V_u du, V_T|F_t\right]}{\sqrt{Var\left[\int_t^T V_u du|F_t\right] Var\left[V_T|F_t\right]}},$$

where

$$\begin{aligned} Var[V_T|F_t] &= V_t I_{t,T} + J_{t,T} \\ Cov\left[\int_t^T V_u du, V_T|F_t\right] &= V_t Q_{t,T} + R_{t,T}, \end{aligned}$$

where

$$\begin{aligned} I_{t,T} &= \sigma_v^2\left[\left(1+\frac{\sigma_v^2}{k_v}\right)\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right. \\ &\quad \left. +2\left(Te^{-k_v(T-t)} - t + \frac{1}{k_v}\left(e^{-k_v(T-t)} - 1\right)\right)\right] \\ J_{t,T} &= \sigma_v^2\left[\left(1+\frac{\sigma_v^2}{k_v^2}\right)\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right. \\ &\quad \left. +2\theta\left(Te^{-k_v(T-t)} - t + e^{-k_v(T-t)} - 1\right)\right] \end{aligned}$$

$$
\begin{aligned}
Q_{t,T} &= \frac{\sigma_v}{k_v}\left[t+1-(T+1)e^{-k_v(T-t)} - \sigma_v\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right] \\
R_{t,T} &= \frac{\sigma_v\theta}{k_v}\left[t+1-(T+1)e^{-k_v(T-t)} - \sigma_v\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right]
\end{aligned}
$$

The proof of Proposition 1.4.2 is given in Appendix F.2.

The simulation based on the estimates in the literature shows that the correlation is between -0.35 and -0.5 and not negligible. Recall that a noncentral chi-square random variable can be represented as an ordinary chi-square random variable with a random degrees of freedom parameter. In detail, when N is a Poisson random variable with mean $\lambda/2$, consider a random variable χ^2_{d+2N}. Conditioning on N, the variable is an ordinary chi-square distribution. On the other hand, the unconditional distribution is noncentral chi-square distribution. Thus, using this mixture property, we can generate noncentral chi-square random variable by generating Poisson random variable and chi-square random variable, which is in the gamma class. Thus, both $\int_t^T V_u du$ and V_T can be approximated by bivariate gamma distributions up to constant.

Unlike multivariate normal distribution, a bivariate gamma distribution cannot be uniquely identified by two gamma marginal distributions and the correlation. Among various forms of bivariate gamma distributions (see Kotz, Balakrishnan and Johnson (1999)), we choose a bivariate gamma distribution that gives a simple algorithm without putting restrictions on parameters. This algorithm was proposed by Shmeiser and Lal (1982). They discuss different simulation methods for a bivariate gamma distribution. It allows a negative correlation between two variables, while many other algorithms only allow positive correlations. This algorithm is based on a probability mixing. The algorithm results in two marginals $X_1 \sim gamma(\alpha_1, \beta_1)$ and $X_2 \sim gamma(\alpha_2, \beta_2)$ with the correlation ρ_g. Denote the cumulative and inverse cumulative distribution function of the gamma distribution with the shape parameter α and a unit scale parameter as $F_\alpha(u)$ and $F_\alpha^{-1}(u)$, respectively. C is the minimum possible correlation, which occurs when $X_2 = \beta_2 F_{\alpha_2}^{-1}(1 - F_{\alpha_1}(X_1/\beta_1))$. Thus, C is obtained by

$$
\rho_g = \frac{E[F_{\alpha_1}^{-1}(U)F_{\alpha_2}^{-1}(1-U)] - \alpha_1\alpha_2}{\sqrt{\alpha_1\alpha_2}}.
$$

The expected value can be easily calculated numerically using $\int_0^1 F_{\alpha_1}^{-1}(u)F_{\alpha_2}^{-1}(1-u)du$.

Algorithm 1.4.1. step-(1) *Generate* U *from* $Uniform(0,1)$.

step-(2) *If* $U < \rho_g/C$, *go to step-(5).*

step-(3) *Generate* X_1 *from* $gamma(\alpha_1, \beta_1)$

step-(4) *Generate* X_2 *from* $gamma(\alpha_2, \beta_2)$ *and go to step-(9).*

step-(5) *Generate* V *from* $Uniform(0,1)$.

step-(6) *Let* X_1 *be* $\beta_1 F_{\alpha_1}^{-1}(V)$.

step-(7) *If* $\rho_g < 0$, *replace* V *with* $1-V$.

step-(8) *Let* X_2 *be* $\beta_2 F_{\alpha_2}^{-1}(V)$.

step-(9) *Deliver* (X_1, X_2).

The problem with this algorithm is illustrated in Shmeiser and Lal (1982). However, it is not an issue with our parameter setting. For the application of this scheme, we simply let X_1 be $\int_t^T V_u du$ and X_2 be non-central chi-squared random variable in V_T. Then, α_1 and β_1 are obtained by the method of moments, as discussed in 1.4.2. α_2 and β_2 are based on the degrees of freedom, noncentrality parameter and the previously mentioned simulation method for non-central chi-squared random variable. Contrary to the first proposed approximation scheme, α_2 and β_2 change each time. However, the increase in the computation time is not significant and we still have a computational efficiency over the path simulation method. Once $(\int_t^T V_u du, V_T)$ are obtained as illustrated here, the remaining steps are the same as in Section 1.4.2. This simulation scheme improves the accuracy of the approximation, especially for options with a long maturity time. It will be discussed in the simulation study.

1.4.4 Simulation study

We check the performance of our proposed approximation schemes in comparison to a benchmark, the option prices calculated from GBS option pricing without any approximation. SV denotes stochastic volatility without jumps. SVJ denotes stochastic volatility with

	k_v	θ	σ_v	ρ	η_v
Set1	0.019	4.08078	0.22	-0.569	0.01
Set2	0.019	1.933	0.22	-0.569	0
Set3	0.028	0.54365	0.127	-0.53	0
Set4	0.016	0.66	0.077	-0.38	0
Set5	0.023	0.905	0.143	-0.397	0

Table 1.1: Parameters for SV model

jumps. The comparative study is done for SV as well as SVJ described in the previous sections. As mentioned earlier, the volatility paths are generated by Brigo and Alfonsi (2005) and the benchmark option prices are obtained by the conditional Monte Carlo. Before the actual simulation study, we can expect much longer computational time for brute force method than for our approximation scheme. To get a vector of $\int_t^T V_u du$, the benchmark method need to simulate n times $(T-t)/\Delta$ data points, where $T-t$ is a time to maturity, Δ is a discretization interval and n is the Monte Carlo iteration (more than 1,000,000). Contrarily, our approximation scheme only needs to simulate n^* (about 5,000) data points from the gamma distribution for the first proposed scheme or $3n^*$ data points for our second proposed scheme.

We choose a discretization interval $\Delta = 0.01$ for the benchmark, which is often used in the literature for small discretization bias. For option price calculation, we assume $r = 0.01\%$ (daily interest rate) and $S_t = 1000$ throughout simulations. Six different times to maturity are considered, $20, 40, 60, 120, 180, 240$ days (equivalently, $1, 2, 3, 6, 9, 12$ months) and 11 different moneyness $\left(\log\left(\frac{S_t}{KB_{t,T}(t,T)}\right)\right)$ from -0.1 to 0.1 to check the robustness. The parameter values for the square-root stochastic volatility models in the comparative study are based on parameter estimates in the previous studies. Set1 is a set of parameters under Q (risk-neutral) measure and Set2 is a set of parameters under P (physical) measure in table III in Eraker (2004). Similarly, Set3 is from Table 1 in Pan (2002). These three sets are estimates from using both returns and options on S&P 500 index. Set4 is from Andersen, Benzoni and Lund (1997), Set5 is from Eraker, Johannes and Polson (2003). These two are estimates from using only returns. When we include jumps in the return process, we need more parameter estimates. SVJ simulation is based on the parameter values in Pan (2002). The parameter estimates in our simualation study are reported in Table 1.1 and Table 1.2,

k_v	θ	σ_v	ρ	η_v	λ	μ_2	μ_3
0.025	0.61	0.12	-0.53	0.012	0.123	0	-0.192

Table 1.2: Parameters for SVJ model

following the convention in Eraker (2004). The return data is scaled by 100 and a unit of time is defined to be one day. For example, if we use Set2, then

$$dV_t = 0.019((1.933)(0.01)(0.01) - V_t)dt + (0.22)(0.01)\sqrt{V}_t dW_t^v$$

will be the equation we use to generate volatilities. Annualized parameters k_v, θ, σ_v (as in Pan (2002)) are converted by dividing by 252 (approximate number of trading days per year). Refer to Singleton (2006) for more on the parameter conversion. More details on model specification and the option pricing formula will be given with MCMC estimation.

Figure 1.1 contains six different quantile-quantile (QQ) plot of our key elements: $\int_t^T V_u du$, V_T and $\int_t^T \sqrt{V_u} dW_u$. Except for several extreme values, our approximation schemes work well. For more careful study, we calculate errors and computational times. Table 1.3 and Table 1.4 contain the mean of log ratios of the benchmark and the approximation under stochastic volatility models without and with jumps. First, we calculate one option price, C_t using the the method by Brigo and Alfonsi (2005) with 100,000 iterations in Monte Carlo simulation. Assume this option price is a correct one and denote it as C_t. On the other hand, the option prices by the approximation method are calculated with 5,000 Monte Carlo iteration. Due to the dimensional reduction, we see that using the smaller number of iteration is justified. The approximated price is denoted as C_t^a. The mean of log ratio, $\epsilon(N)$ is calculated by

$$\epsilon(N) = \frac{1}{N}\sum_{i=1}^{N} \log(C_t^a(i)/C_t).$$

If the approximation is accurate, this quantity is close to 0. In our simulation study, we use $N = 100$. Time represents the average of N ratios of time required with approximation scheme to the time required for the benchmark method.

$$Time^{avg}(N) = \frac{1}{N}\sum_{i=1}^{N}(Time^a(i)/Time),$$

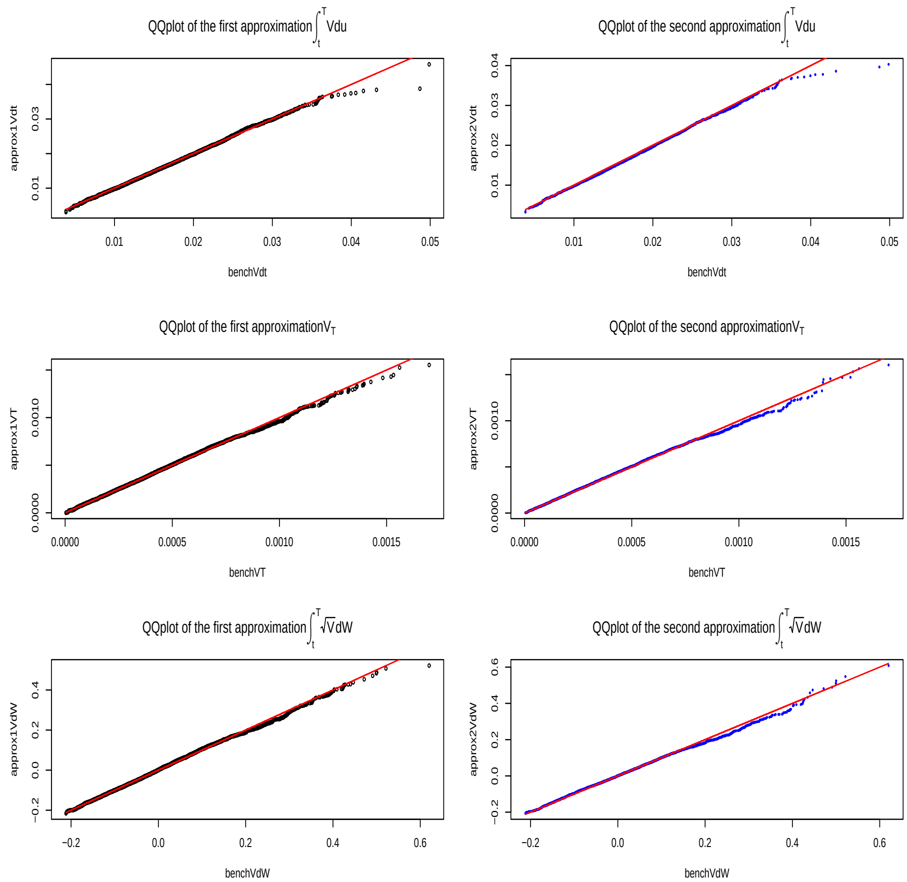

Figure 1.1: The qq plots of $\int_t^T V_u du$, V_T and $\int_t^T \sqrt{V_u} dW_u$. The first approximation is without incorporating the correlation and the second approximation is using the correlation and the bivariate gamma sampling

where $Time^{avg}$ represents the computational time to calculate one option price with the approximation scheme and $Time$ is the time to calculate one option price without approximation.

The result of the first proposed approximation scheme

In both SV and SVJ settings, our simulation scheme performs well as in Table 1.3 and Table 1.4. Absolute mean errors are close to 0. We obtain a significant reduction in computational time without losing accuracy.

moneyness	20 days	40 days	60 days	80 days	180 days	240 days
-0.09	0.085	0.092	0.093	0.095	0.103	0.104
-0.06	0.083	0.085	0.092	0.096	0.099	0.102
-0.03	0.078	0.076	0.079	0.081	0.088	0.098
0	0.071	0.076	0.077	0.086	0.092	0.111
0.03	0.071	0.071	0.083	0.085	0.092	0.122
0.06	0.079	0.083	0.090	0.093	0.099	0.112
0.09	0.078	0.083	0.085	0.105	0.110	0.111
$Time^{avg}$	0.020	0.021	0.021	0.017	0.015	0.010

Table 1.3: Mean of log ratios and the computational times in option prices under SV based on Set2, using the first approximation scheme

moneyness	20 days	40 days	60 days	80 days	180 days	240 days
-0.09	0.086	0.093	0.093	0.095	0.102	0.107
-0.06	0.083	0.086	0.097	0.098	0.099	0.104
-0.03	0.074	0.077	0.079	0.083	0.089	0.098
0	0.072	0.076	0.077	0.091	0.092	0.111
0.03	0.071	0.072	0.083	0.084	0.097	0.124
0.06	0.078	0.085	0.091	0.096	0.108	0.112
0.09	0.077	0.083	0.084	0.107	0.110	0.113
$Time^{avg}$	0.023	0.023	0.022	0.021	0.019	0.018

Table 1.4: Mean of log ratios and the computational times in option prices under SVJ, using the first approximation scheme

The result of the second proposed approximation scheme

moneyness	20 days	40 days	60 days	80 days	180 days	240 days
-0.09	0.079	0.081	0.081	0.083	0.084	0.087
-0.06	0.076	0.079	0.080	0.081	0.084	0.086
-0.03	0.072	0.072	0.073	0.076	0.078	0.079
0	0.071	0.075	0.077	0.081	0.081	0.082
0.03	0.073	0.073	0.076	0.078	0.083	0.082
0.06	0.071	0.073	0.074	0.080	0.080	0.086
0.09	0.074	0.084	0.088	0.090	0.093	0.089
$Time^{avg}$	0.033	0.031	0.031	0.028	0.027	0.024

Table 1.5: Absolute mean errors and the computational times in option prices under SV based on Set2, using the second approximation scheme

moneyness	20 days	40 days	60 days	80 days	180 days	240 days
-0.09	0.079	0.082	0.081	0.085	0.087	0.091
-0.06	0.076	0.079	0.080	0.084	0.087	0.093
-0.03	0.073	0.072	0.075	0.077	0.079	0.089
0	0.071	0.075	0.076	0.083	0.085	0.087
0.03	0.073	0.072	0.074	0.079	0.086	0.087
0.06	0.071	0.076	0.078	0.084	0.086	0.087
0.09	0.074	0.084	0.089	0.092	0.093	0.092
$Time^{avg}$	0.032	0.030	0.030	0.025	0.026	0.025

Table 1.6: Absolute mean errors and the computational times in option prices under SVJ, using the second approximation scheme

In both SV and SVJ settings, our simulation scheme performs better as in Table 1.5 and Table 1.6 than the first proposed approximation scheme. Absolute mean errors are close to 0. The computational time increased slightly because of the parameter adjustment in Poisson random variable generation for V_T. However, the computation is still efficient without loss of accuracy.

1.5 MCMC estimation

The inference of our asset pricing model is about characterizing $p(\Theta, V, J, U|Y, C)$, where Θ is a set of parameters, V is a volatility, J is a jump indicator and U is a jump size and Y, C are observed stock and option prices. It is difficult because $p(\Theta, V, J, U|Y, C)$ is typically high-dimensional and thus the standard sampling fails. Also, in the option pricing models, the parameters and the state variables, such as V, are in a non-analytic form. MCMC methods are well-suited in the setting. It is a unified estimation procedure, simultaneously estimating both parameters and latent variables. Contrary to other methods, applying approximate filters or noisy latent variable proxies, MCMC directly computes the distribution of the latent variables and parameters given the observed prices. Since it avoids any optimization and does the unconditional simulation, it is computationally fast. Some MCMC estimation techniques are based on the previous work by Jacquier, Polson and Rossi (1994) and Jacquier, Polson and Rossi (2004). Among a growing amount of literature in finance using MCMC estimation, work by Eraker (2004) is the closest to our estimation algorithm, since both option and return data are used. The difference is the specification of jump and jump risk premium. Also, details on discretization and specification of conjugate priors are in the chapter by Chib, Nardari and Shephard (2002) and Johannes and Polson (2006). However, there is no precedent who uses the application of MCMC estimation algorithm in our model setting with both option and return data. It will be shown in this section that the approximation bias is minor and it will not affect the MCMC algorithm much. In the current stage, this section serves only as a performance check for the approximation scheme and a rough illustration of MCMC algorithm. Thus, careful study on the choice of priors and diagnostics should follow.

1.5.1 The random-walk Metropolis-Hastings method

In our model setting, some conditional distributions cannot be recognized and conveniently sampled. In this case, the Metropolis-Hastings method can be used. Among several different Metropolis-Hastings methods (refer to Robert and Casella (2004)), we choose to use the random-walk Metropolis-Hastings method in order to tackle this issue. For example, to generate samples for one parameter, say θ from $\pi(\theta)$, it draws a candidate from the following random walk model, $\theta^* = \theta^{(g)} + s\varepsilon$, where s is a tuning parameter and ε is a symmetric density function. The random-walk Metropolis-Hastings method is:

(i) Draw θ^*.

(ii) Accept θ^* with probability α, where $\alpha = \min\left[\frac{\pi(\theta^*)}{\pi(\theta^{(g)})}, 1\right]$

In our simulation study, we use the normal distribution for ε and set different tuning parameters and the uniform priors for each Metropolis-Hastings algorithm.

1.5.2 Model and discretization

For estimation purpose, we need to specify our models further. We use exactly the same model as Pan (2002) for comparison purpose.

$$d(\log S_t) = [r_t + \eta_s V_t - \frac{1}{2}V_t]dt + \sqrt{V_t}dW_t^s + (\log S_t - \log S_{t-})dN_t - \mu_1\lambda V_t dt \tag{1.14}$$

$$dV_t = k_v(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_{1t}, \tag{1.15}$$

where r_t is known, $Cov(dW_t^s, dW_{1t}) = \rho dt$, $\log S_t - \log S_{t-} = U_i$ is i.i.d. normal variable with mean μ_J, and variance σ_J. $\mu_1 = \exp(\mu_J + \sigma_J^2/2) - 1$ and $N_T - N_t$ is a Poisson process with the intensity parameter $\lambda \int_t^T V_u du$. The errors are specified as

$$\log(C_t) = \log(GBS) + \delta\varepsilon_t,$$

where C_t is an observed option price, GBS is the option price calculated from GBS option pricing formula and δ is a price error parameter and $\varepsilon_t \sim N(0,1)$. Taking log prevents us from gettig a negative option price. Using the previous results, the GBS option price is calculated as:

$$\begin{aligned} C_t &= E_t[\widetilde{BS}(S_t\xi_{t,T}, (\overline{\sigma_{t,T}})^2)] \\ &= E_t[S_t\xi_{t,T}\Phi(d_1) - KB^*(t,T)\Phi(d_2)], \end{aligned}$$

where

$$\widetilde{BS}(\cdot,\cdot) \text{ is a BS-like option pricing formula}$$

$$d_1 = \tfrac{1}{\overline{\sigma_{t,T}}\sqrt{T-t}}\left[\log(\tfrac{S_t\xi_{t,T}}{KB^*(t,T)}) + \tfrac{1}{2}(\overline{\sigma_{t,T}})^2(T-t)\right]$$

$$d_2 = d_1 - \overline{\sigma_{t,T}}\sqrt{T-t}$$

$$\overline{\sigma_{t,T}} = \sqrt{\frac{(1-\rho^2)\int_t^T V_u du + (N_T - N_t)\sigma_J^2}{T-t}}$$

$$\begin{aligned} &B^*_{t,T} \\ &= \exp\left[\int_t^T \left(-r_u - \frac{1}{2}\eta_v^2 V_u - \mu_2\lambda V_u\right) du + \eta_v \int_t^T \sqrt{V_u} dW_{1u}\right] (\mu_2+1)^{N_T - N_t}. \end{aligned}$$

$$\begin{aligned} &\xi_{t,T} \\ &= \exp\left[-\int_t^T \mu_3\lambda V_u du\right] \exp\left[(\eta_v + \rho)\int_t^T \sqrt{V_u} dW_{1u} - \frac{1}{2}(\eta_v+\rho)\int_t^T V_u du\right] \\ &(\mu_3+1)^{N_T-N_t}. \end{aligned}$$

Assuming the prices are observed at equally spaced, discrete time interval ($t = 1, 2, \ldots, T$ days) and the equation(1.14) and equation(1.15) become

$$\begin{aligned} Y_t &= Y_{t-1} + r_t + (\eta_s - \frac{1}{2})V_t + \sqrt{V_t}\varpi_t^s + U_t J_t - \mu_1 \lambda V_t \\ &= Y_{t-1} + r_t + mV_t + \sqrt{V_t}\varpi_t^s + U_t J_t, \end{aligned}$$

and

$$V_t = k_v\theta + (1 - k_v)V_{t-1} + \sigma_v\sqrt{V_{t-1}}\varpi_t^v,$$

where $Y_t = \log(S_t)$, $m = \eta_s - \frac{1}{2} - \mu_1\lambda$, $\varpi_t^s \sim N(0, 1)$ and $\varpi_t^v \sim N(0, 1)$ with $corr(\varpi_t^s, \varpi_t^v) = \rho$, and a jump indicator $J_t \sim Bernoulli(\lambda V_t)$. η_s is obtained from m with values of the jump parameters, μ_J, σ_J and λ.

1.5.3 MCMC algorithms

Now, we identify parameters that will be estimated. Although the model and the GBS pricing formula have many parameters, the conditional posterior distributions can be simplified with blocks, using conditional independence. Denote all parameters by Θ, where $\Theta = (m, \rho, \mu_J, \sigma_J^2, \lambda, k_v, \theta, \sigma_v, \mu_2, \mu_3, \eta_v, \delta)$. The methods for performing the draws of the parameters are contained in this section. In drawing these parameters, virtually flat priors are used exclusively when a little information of the prior is found. First, recognize the following distributions, which will be components of the posterior distributions.

- $p(C_t|Y_t, V_t, \Theta) \sim Lognormal(\log(GBS), \delta^2)$ with mean $\log(GBS)$ and variance δ^2
- $p(Y_t|Y_{t-1}, V_t, J_t, U_t, \Theta) = p(Y_t|Y_{t-1}, V_t, J_t, U_t, m, \rho)$
  $\sim N(Y_{t-1} + r_t + mV_t + U_t J_t, (1 - \rho^2)V_t)$
- $p(U_t|\Theta) = p(U_t|\mu_J, \sigma_J^2) \sim N(\mu_J, \sigma_J^2)$
- $p(J_t|V_t, \Theta) = p(J_t|V_t, \lambda) \sim Bernoulli$ with $p(J_t = 1|V_t, \lambda) = \lambda V_t$
- $p(V_t|V_{t-1}, \Theta) = p(V_t|V_{t-1}, k_v, \theta, \sigma_v) \sim truncated\ N(k_v\theta + (1 - k_v)V_{t-1}, \sigma_v^2 V_{t-1})$

For the estimation for SV model, we simply skip updating jump related latent variables and set all jump parameters to zero.

Latent Variables (Volatility, Jump Time and Jump Size)

We use the following notations: $Y = \{Y_t\}$, $C = \{C_t\}$, $V = \{V_t\}$, $U = \{U_t\}$, $J = \{J_t\}$. When the prior is missing, assume that we have a uniform prior.

Updating Volatility, V

By Markov property, V_t is only influenced by its neighbor, V_{t+1} and V_{t-1}. Thus,

$$p(V_t|V_{(-t)}, Y, J, U, C, \Theta) \quad = \quad p(V_t|V_{t-1}, V_{t+1}, Y, J, U, C, \Theta)$$

Sample volatility V_t from

$$\begin{aligned}
&p(V_t|V_{t-1}, V_{t+1}, Y, J, U, C, \Theta) \\
&\quad \propto p(V_{t-1}, V_t, V_{t+1}|Y, J, U, C, \Theta) \\
&\quad \propto p(C_t|Y_t, V_t, \Theta)p(Y_t|Y_{t-1}, V_t, J_t, U_t, \Theta)p(U_t|V_t, J_t, \Theta)p(J_t|V_t, \Theta)p(V_t|V_{t-1}, \Theta)p(V_{t+1}|V_t, \Theta)
\end{aligned}$$

Since this distribution is not recognizable, we cannot do direct simulation from the standard distribution. Thus, we use the random-walk Metropolis-Hastings method to sample from it.

Updating Jump Sizes, U

Since the jump sizes do not depend on the option prices directly and the jump sizes are i.i.d. random variable,

$$p(U_t|U_{(-t)}, Y, V, J, C, \Theta) \quad = \quad p(U_t|Y_t, Y_{t-1}, V_t, J_t, \Theta)$$

Sample jump size U_t from

$$\begin{aligned}
p(U_t|Y_t, Y_{t-1}, V_t, J_t, \Theta) \quad &\propto \quad p(Y_t|Y_{t-1}, J_t, U_t, V_t, \Theta)p(U_t|J_t, V_t, \Theta) \\
&\propto \quad \exp\left[-\frac{(Y_t - Y_{t-1} - r_t - mV_t - U_t J_t)^2}{2(1-\rho^2)V_t}\right] \exp\left[-\frac{(U_t - \mu_J)^2}{2\sigma_J^2}\right]
\end{aligned}$$

$$\begin{aligned}
&\propto \exp\left[-\frac{(U_t-\mu_J^*)^2}{2(\sigma_J^*)^2}\right] \\
&\sim N(\mu_J^*,(\sigma_J^*)^2)
\end{aligned}$$

First, consider the case when $J_t = 0$.

$$p(U_t|Y_t, Y_{t-1}, V_t, J_t = 0, \Theta) \quad \propto \quad \exp\left[-\frac{(U_t-\mu_J)^2}{2\sigma_J^2}\right],$$

which implies that $\mu_J^* = \mu_J$,$(\sigma_J^*)^2 = \sigma_J^2$.

Second, consider the case when $J_t = 1$.

$$\begin{aligned}
p(U_t|Y_t, Y_{t-1}, V_t, J_t = 1, \Theta) \quad &\propto \quad \exp\left[-\frac{(Y_t - Y_{t-1} - r_t - mV_t - U_t)^2}{2(1-\rho^2)V_t}\right]\exp\left[-\frac{(U_t-\mu_J)^2}{2\sigma_J^2}\right] \\
&\propto \quad \exp\left[-\frac{(U_t - (Y_t - Y_{t-1} - r_t - mV_t))^2}{2(1-\rho^2)V_t}\right]\exp\left[-\frac{(U_t-\mu_J)^2}{2\sigma_J^2}\right],
\end{aligned}$$

which implies that $\mu_J^* = (\sigma_J^*)^2\left[\frac{Y_t-Y_{t-1}-r_t-mV_t}{(1-\rho^2)V_t} + \frac{\mu_J}{\sigma_J^2}\right]$, $(\sigma_J^*)^2 = \left[\frac{1}{(1-\rho^2)V_t} + \frac{1}{\sigma_J^2}\right]^{-1}$. Thus, we can generalize that the posterior distribution is

$$U_t|Y_t, Y_{t-1}, V_t, J_t, \Theta \quad \sim \quad N(\mu_J^*, (\sigma_J^*)^2),$$

where

$$\begin{aligned}
\mu_J^* \quad &= \quad (\sigma_J^*)^2\left[\frac{(Y_t - Y_{t-1} - r_t - mV_t)J_t}{(1-\rho^2)V_t} + \frac{\mu_J}{\sigma_J^2}\right] \\
(\sigma_J^*)^2 \quad &= \quad \left[\frac{J_t}{(1-\rho^2)V_t} + \frac{1}{\sigma_J^2}\right]^{-1}
\end{aligned}$$

Updating Jump Indicator, J

Since the jump times do not depend on the option prices directly,

$$p(J_t|J_{(-t)}, Y, V, U, C, \Theta) \quad = \quad p(J_t|Y_t, Y_{t-1}, V_t, U_t, \Theta)$$

The jump indicator J_t only takes two values, 0 and 1. The following two formula provide the Bernoulli probability. Due to the lack of knowledge of its normalizing constant, we need to update both.

$$\begin{aligned} p(J_t = 1|Y_t, Y_{t-1}, V_t, U_t, \Theta) &\propto p(Y_t|Y_{t-1}, V_t, J_t = 1, U_t, \Theta)p(J_t = 1|V_t, \Theta) \\ &\propto \lambda V_t p(Y_t|Y_{t-1}, V_t, J_t = 1, U_t, \Theta) \\ &\propto \lambda V_t \exp\left[-\frac{(Y_t - Y_{t-1} - r_t - mV_t - U_t)^2}{2(1-\rho)^2 V_t}\right] \end{aligned}$$

$$\begin{aligned} p(J_t = 0|Y_t, Y_{t-1}, V_t, U_t, \Theta) &\propto p(Y_t|Y_{t-1}, V_t, J_t = 0, U_t, \Theta)p(J_t = 0|V_t, \Theta) \\ &\propto (1 - \lambda V_t)p(Y_t|Y_{t-1}, V_t, J_t = 0, U_t, \Theta) \\ &\propto (1 - \lambda V_t) \exp\left[-\frac{(Y_t - Y_{t-1} - r_t - mV_t)^2}{2(1-\rho^2) V_t}\right] \end{aligned}$$

Parameters in Return Process

Updating m

The conditional posterior distribution of drift parameter in stock return can be obtained from the jump-adjusted stock prices, conditioning on jump times and sizes. That is,

$$\widetilde{Y_t} = \frac{Y_t - Y_{t-1} - r_t - U_t J_t}{\sqrt{V_t}} = m\sqrt{V_t} + \varpi_t^s.$$

Assuming that m has a prior distribution, $\mathrm{N}(\mu_m,\, \sigma_m^2)$,

$$\begin{aligned} p(m|\widetilde{Y}, V) &\propto p(\widetilde{Y}|V, m)p(m) \\ &\propto \exp\left[-\frac{(\widetilde{Y} - m\sqrt{V})^T(\widetilde{Y} - m\sqrt{V})}{2(1-\rho^2)}\right] \exp\left[-\frac{(m-\mu_m)^2}{2\sigma_m^2}\right] \\ &\propto \exp\left[-\frac{(m-\widehat{m})^2\sqrt{V}^T\sqrt{V}}{2(1-\rho^2)}\right] \exp\left[-\frac{(m-\mu_m)^2}{2\sigma_m^2}\right], \end{aligned}$$

where $\widetilde{Y}$ is an $(n\times1)$ vector, $(\widetilde{Y_1}, \widetilde{Y_2}, \ldots, \widetilde{Y_n})^T$ and $\sqrt{V}$ is an $(n\times1)$ vector, $(\sqrt{V_1}, \sqrt{V_2}, \ldots, \sqrt{V_n})^T$.

Thus, the drift term in stock return has a posterior distribution

$$m|Y, V, J, U, \Theta \sim N(\mu_m^*, (\sigma_m^*)^2),$$

where $\widehat{m}$ is an OLS estimator of m, $\frac{\sqrt{V}^T \widetilde{Y}}{\sqrt{V}^T \sqrt{V}}$

$$\begin{aligned}
\mu_m^* &= (\sigma_m^*)^2 \left[\frac{\mu_m}{\sigma_m^2} + \frac{\sqrt{V}^T \widetilde{Y}}{(1-\rho^2)} \right], \\
(\sigma_m^*)^2 &= \left[\frac{\sqrt{V}^T \sqrt{V}}{(1-\rho^2)} + \frac{1}{\sigma_m^2} \right]^{-1}.
\end{aligned}$$

Updating ρ

We sample ρ from

$$p(\rho|Y, V, C, \Theta) \quad \propto \quad p(\rho) \prod p(C_t|Y_t, V_t, \rho, \Theta) p(Y_t|Y_{t-1}, V_t, J_t, U_t, m, \rho),$$

where U is $(n \times 1)$ vector, $(U_1, U_2, \ldots, U_n)$. It is not a recognizable distribution. Thus, we use the random-walk Metropolis-Hastings method.

Updating Jump Parameters, μ_J and σ_J^2

Assume that jump size parameters, μ_J and σ_J^2, have a prior, the normal-inverse gamma distribution with parameters, (a, A, b, B). That is, $\mu_J|\sigma_J^2 \sim N(a, A\sigma_J^2)$ and $\sigma_J^2 \sim IG(b, B)$.

$$\begin{aligned}
&p(\mu_J, \sigma_J^2|U) \\
&\quad \propto p(U|\mu_J, \sigma_J^2) p(\mu_J, \sigma_J^2) \\
&\quad \propto (\sigma_J^2)^{-\frac{n}{2}} \exp\left[-\frac{\sum (U_t - \mu_J)^2}{2\sigma_J^2} \right] (A\sigma_J^2)^{-\frac{1}{2}} \exp\left[-\frac{(\mu_J - a)^2}{2A\sigma_J^2} \right] \frac{B^b}{\Gamma(b)} (\sigma_J^2)^{-b-1} \exp\left[-\frac{B}{\sigma_J^2} \right] \\
&\quad \propto (\sigma_J^2)^{-(\frac{n}{2}+b)-1} \exp\left[-\frac{B + \sum (U_t - \overline{U})^2/2}{\sigma_J^2} \right] (\sigma_J^2)^{-\frac{1}{2}} \exp\left[-\frac{n(\mu_J - \overline{U})^2}{2\sigma_J^2} \right] \exp\left[-\frac{(\mu_J - a)^2}{2A\sigma_J^2} \right],
\end{aligned}$$

where $\overline{U} = \sum U_t/n$. Thus, the posterior distribution is $\mu_J, \sigma_J^2|U \sim N\text{-}IG(a^*, A^*, b^*, B^*)$, where

$$b^* \quad = \quad b + \frac{n}{2}$$

$$\begin{aligned} B^* &= B + \frac{\sum (U_t - \overline{U})^2}{2} \\ a^* &= (\sum U_t + a)(n + \frac{1}{A})^{-1} \\ A^* &= (n + \frac{1}{A})^{-1}. \end{aligned}$$

We first sample $\sigma_J^2|U$ from $IG(b^*, B^*)$. Then, $\mu_J|\sigma_J^2, U$ from $N(a^*, A^*\sigma_J^2)$

Updating λ

$$p(\lambda|V, J, U, \Theta) \propto p(J|V, \lambda)p(\lambda),$$

which is not a recognizable distribution. Thus, we use the random-walk Metropolis-Hastings method.

Parameters in Volatility Process

Updating k, θ

We sample k, θ from

$$p(k, \theta|Y, V, C, \Theta) \quad \propto \quad p(k, \theta) \prod_{t=1}^{n} p(C_t|Y_t, V_t, \Theta) p(V_t|V_{t-1}, k, \theta, \sigma_v),$$

which is not a recognizable distribution. Thus, we use the random-walk Metropolis-Hastings method.

Updating σ_v

We sample σ_v^2 from

$$p(\sigma_v^2|Y, V, C, \Theta) \quad \propto \quad p(\sigma_v^2) \prod_{t=1}^{n} p(C_t|Y_t, V_t, \Theta) p(V_t|V_{t-1}, k, \theta, \sigma_v),$$

which is not a recognizable distribution. Thus, we use the random-walk Metropolis-Hastings method.

Parameters in GBS Formula and in Pricing Error

We use the random-walk Metropolis-Hastings method for all following parameters.

Updating μ_2

$$p(\mu_2|V,Y,U,J,C,\Theta) \propto p(\mu_2)\prod_{t=1}^{n} p(C_t|Y_t,V_t,\Theta)$$

Updating μ_3

$$p(\mu_3|V,Y,U,J,C,\Theta) \propto p(\mu_3)\prod_{t=1}^{n} p(C_t|Y_t,V_t,\Theta)$$

Updating η_v

$$p(\eta_v|V,Y,U,J,C,\Theta) \propto p(\eta_v)\prod_{t=1}^{n} p(C_t|Y_t,V_t,\Theta)$$

Updating δ^2

By assuming IG(f,F) prior for δ^2,

$$\begin{aligned} p(\delta^2|V,Y,C,\Theta) &\propto p(\delta^2)\prod_{t=1}^{n} p(C_t|Y_t,V_t,\Theta,\delta^2) \\ &\propto (\delta^2)^{-n/2}\exp\left(-\frac{\sum(\log(C_t)-\log(GBS))^2}{2\delta^2}\right)(\delta^2)^{-f-1}\exp(-F/\delta^2), \end{aligned}$$

The posterior distribution is $\delta^2|V,Y,C,\Theta \sim IG(f^*,F^*)$, where

$$\begin{aligned} f^* &= f + n/2 \\ F^* &= F + \frac{\sum(\log(C_t)-\log(GBS))^2}{2}. \end{aligned}$$

1.5.4 MCMC convergence diagnosis

In frequentist analysis, estimated parameters and associated standard errors are examined. In that setting, convergence assessment involves checking that the sequence has converged to a single point. In MCMC estimation, the interest is estimating posterior distributions of model parameters rather than individual parameter values and standard errors. Thus, the convergence assessment of MCMC estimation involves checking the sequence has converged to the posterior distribution. It is simple to check the convergence when the posterior distribution has a standard form. However, convergence to an unknown joint posterior cannot be proved. There are some diagnostic tests that are developed to identify MCMC output that has not converged to a stationary distribution. It is important to use these diagnostic analysis since using option data introduced many non-standard posterior distributions in MCMC algorithm. Here are several methods that we considered. We will only report Gelman and Rubin's statistic for MCMC estimation with approximation since it is sufficient. For computing algorithm, we use coda R package developed by Plummer, Best, Cowles and Vines(2007).

Gelman and Rubin's method

Gelman and Rubin (1992) proposed diagnostic as a univariate statistic, referred to as the potential scale reduction factor (PSRF), for assessing convergence of individual model parameters. Calculation of this statistic is based on the last n samples in each of m parallel chains. In particular, the PSRF is calculated as

$$PSRF = \sqrt{\frac{n-1}{n} + \frac{(m+1)B}{mnW}}$$

where B/n is the between-chain variance and W is the within-chain variance. As chains converge to a common target distribution and traverse said distribution, the between-chain variability should become small relative to the within-chain variability and yield a PSRF that is close to 1. Conversely, PSRF values larger than 1 indicate non-convergence. A corrected scale reduction factor (CSRF) was subsequently proposed to account for sampling variability in the estimate of the true variance for the parameter of interest and is computed as

$$CSRF = PSRF\sqrt{\frac{df+3}{df+1}}$$

where df is a method of moments estimate of the degrees of freedom, based on a t-approximation in the posterior inference.

Geweke's method

The diagnostic of Geweke (1992) is univariate in nature and applicable to a single chain. Convergence is assessed by comparing the sample mean in an early segment of the chain $\{x_{1,j} : j = 1, 2, \ldots, n_1\}$ to the mean in a later segment $\{x_{2,j} : j = 1, 2, \ldots, n_2\}$. Geweke originally suggested that the comparison be between the the first $n_1 = 0.1n$ and last $n_2 = 0.5n$ samples in the chain, although the diagnostic can be applied with other choices. However, inference based on the proposed diagnostic is only valid if the two segments can be considered independent. Thus, the chosen segments should not overlap and be far enough apart so as to satisfy the independence assumption. The statistic upon which this diagnostic is based has the general form

$$z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\widehat{S_1}(0)/n_1 + \widehat{S_2}(0)/n_2}}$$

where the variance estimate $\widehat{S}(0)$ is calculated as the spectral density at frequency zero to account for serial correlation in the sampler output. If the two segments are from the same stationary distribution, the limiting distribution for this statistic is a standard normal. Thus, a frequentist p-value can be computed for this statistic as a measure of evidence against the two sequences being from a common stationary distribution.

1.5.5 Simulation study

To check the MCMC algorithm, stock and option prices with length 2000 are generated. True values are taken from Pan (2002). In option data generation, moneyness and maturity time are randomly selected from $(-0.1, 0.1)$ and $(10, 180)$ days and any approximation methods are not used. The simulation is designed by mimicking the data filtering, using a subset of a whole data set in Eraker (2004). We assume that there is only one option price per a day. The result using the number of iteration of MCMC algorithm equals to $100,000$ is reported in this paper. For comparative study, we choose the second approximation scheme with bivariate gamma distribution.

Based on the posterior means and standard deviations in Table 8 and Table 9, we can see that the approximation scheme does not destroy MCMC results. We have a huge computational time reduction when the approximation method is used in MCMC algorithm for option calculation. In general, the results seem good. However, there are several unstable parameters. We calculated the potential scale reduction factors (PSRF) by Gelman and

Rubins method. All factors are smaller than 1. This is an evidence parameter estimates converge to our final estimate.

	prior	tuning parameter
V_t	NA	0.005
m	N(3,0.5)	NA
ρ	Uniform[-0.8,-0.3]	0.005
μ_J, σ_J^2	N(0, 0.01), IG(0.2, 0.2)	NA
λ	Uniform[10, 15]	0.005
k_v	Uniform[0, 0.05]	0.005
θ	Uniform[0.2, 0.8]	0.02
σ_v	Uniform[0, 0.3]	0.005
μ_2	Uniform[0,0.2]	0.02
μ_3	Uniform[-0.2,0.2]	0.005
η_v	Uniform[0, 0.05]	0.005
δ^2	IG(0.001, 0.1)	NA

Table 1.7: Priors and the tuning parameters

parameters	true value	posterior mean	posterior standard deviation
k_v	0.025	0.031	0.104
θ	0.610	0.585	0.071
σ_v	0.120	0.141	0.090
ρ	-0.530	-0.711	0.312
η_v	0.012	0.018	0.056
λ	12.3	11.0	0.570
m	3.1	2.0	0.073
μ_J	-0.008	-0.011	0.018
σ_J^2	0.1497^2	0.037	0.061
μ_2	0.1	0.08	0.022
μ_3	-0.192	-0.203	0.112
δ^2	0.005	0.032	0.078

Table 1.8: Parameters for simulation and posterior mean and standard deviations when we did not use any approximation

parameters	true value	posterior mean	posterior standard deviation	PSRF
k_v	0.025	0.033	0.103	0.32
θ	0.610	0.594	0.072	0.24
σ_v	0.120	0.138	0.092	0.40
ρ	-0.530	-0.71	0.311	0.91
η_v	0.012	0.017	0.052	0.42
λ	0.123	0.092	0.104	0.46
m	-0.486	-0.520	0.050	0.21
μ_J	-0.009	-0.012	0.013	0.17
σ_J^2	0.045	0.037	0.063	0.28
μ_2	0	-0.008	0.012	0.29
μ_3	-0.192	-0.203	0.109	0.60
δ^2	0.005	0.046	0.088	0.57

Table 1.9: Parameters for simulation and posterior means, standard deviations and potential scale reduction factor (PSRF) by Gelman and Rubins method when we used the second approximation

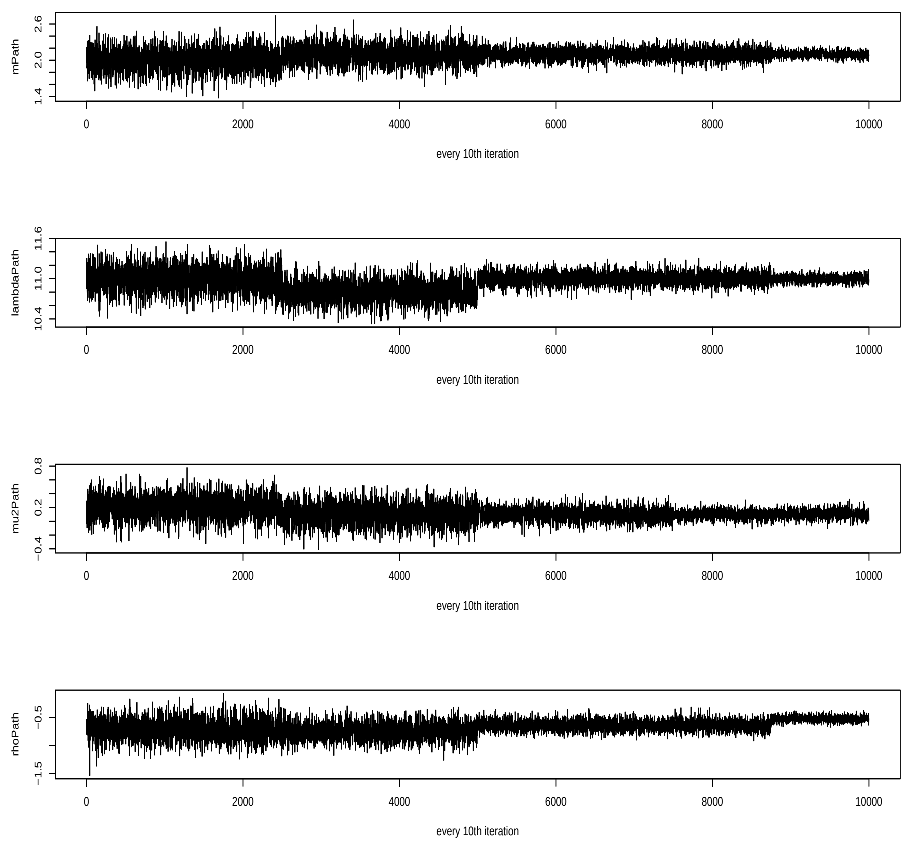

Figure 1.2: Trace plots of four parameters after MCMC estimation step

CHAPTER 2

Why are Black-Scholes implied volatilities not good volatility predictors?

2.1 Introduction

"The volatility implied in an option's price is widely regarded as the option market's forecast of future return volatility over the remaining life of the relevant option" (Christensen and Prabhala (1998)). Surprisingly enough, most of the literature which asks whether Black-Scholes implied volatility (BSIV) predicts future volatility and whether it does so efficiently considers this question as tantamount to know whether the option markets process volatility information efficiently. We argue in this chapter that two questions are not equivalent for two reasons.

First, saying that the option price conveys some efficient forecasting information about future volatility does not mean that the efficient volatility forecast should be the Black-Scholes implied volatility (BSIV). After all, another function of the option price may provide a better predictor. We will in particular document an important Jensen effect, which is due to convexity of BS price with respect to the underlying stock price. We call it a gamma effect. The gamma effect is generally overlooked by the literature on the relation between implied and realized volatility.

Second, also due to Jensen effects, the definition of an optimal predictor depends on the loss function at stake to elicit a predictor. As already stressed by Christoffersen and Jacobs (2004), if one keeps in mind that volatility forecasting has held much attention especially for the purpose of derivative pricing, the relevant loss function should rather be based on option pricing errors than on Mincer-Zarnowitz type regression (Mincer and Zarnowitz

(1969)) about realized volatility. As already used in the context of volatility forecasting by Andersen and Bollerslev (1998), Mincer-Zarnowitz type regression amounts to measure the quality of a forecast of an arbitrary variable, z_t, by the R^2 in a regression of the ex-post realized value of z_t, on its forecast value (and a constant). Even though we briefly consider such regressions to compare volatility forecasts, we rather put the main emphasis of the improvement of option pricing accuracy provided by a proper account for the gamma effect.

The two above arguments are the main motivation of the study in this chapter. We aim at devising a new way of taking advantage of BSIV, in spite of perverse Jensen effects. The naive view of BSIV as "the option market's forecast of future return volatility" may be justified, especially for at-the-money options, by the following argument (see also Renault and Touzi (1996)). Under a stochastic volatility model where the correlation between the asset returns and the volatility is zero, an European call option is priced as

$$C_t = E^Q[BS(\overline{V_{t,T}})|F_t] = E_t^Q[BS(\overline{V_{t,T}})], \tag{2.1}$$

where

$$\overline{V_{t,T}} = \frac{1}{T-t}\int_t^T V_u du, \tag{2.2}$$

Q represents the risk-neutral measure and F_t is the information at time t, including the asset price at t, S_t. Thus, if we may admit that BS option price is, for at-the-money option close to be linear with respect to squared volatility, we deduce that the BSIV should be a good proxy of the volatility forecast $E_t^Q(\overline{V_{t,T}})$. Of course, this forecast is optimal only for the risk-neutral probability measure and may be biased for the historical probability measure P, as documented by Chernov (2007). The difference between P and Q and its impact on volatility forecasting are not our focus of interest in this chapter, and we always do as if the two probability measures were equal. It would actually be possible to merge Chernov's contribution on the impact of volatility risk premium ($P \neq Q$) with our methodology devised to accommodate the gamma effect. Beyond the difference between P and Q, it is generally believed that only the Jensen effect due to the non-linearity of BS pricing formula with respect to squared volatility is responsible for biased forecasts. Surprisingly, people seem to overlook a much more severe Jensen effect, which is due to convexity of BS price with respect to the underlying stock price, the gamma effect.

We are going to show that, due to this gamma effect, the above justifications of the use of BSIV forecast does not work anymore if there is nonzero correlation between the asset returns and the volatility (leverage effect) is introduced. Since the crash of 1987, the observed implied volatility smiles are often skewed and the nonzero correlation needs to be incorporated for the asymmetry of the distribution. Chernov (2007) shows through a simulation study that the nonzero leverage effect does not affect the empirical linkage between the future volatility and BSIV for at-the-money option. No theoretical result has been established for BSIV as a proxy for the future volatility forecast in presence of leverage effect.

The contribution of this work is twofold. First, we try to explain why the nonzero leverage effect does not destroy the linkage between the future volatility and BSIV for at-the-money option. This simple theoretical framework will explain the empirical results given in Chernov (2007) and justify the use of the implied volatility from at-the-money options as future volatility forecasts. Second, we propose a modified BSIV that works for a wider range of moneyness. Only BSIV from at-the-money works as an unbiased volatility forecast. In the design of this proposed implied volatility, our observation on two kinds of Jensen effects on the option price caused by the nonzero leverage effect is used.

The paper is organized as follows. Section 2.2 argues that the linkage between the BS implied volatility and future volatility is affected by the nonzero leverage effect. Also, it discusses a simple simulation result. Section 2.3 introduces a modified BS option pricing formula with an adjusted stock price. We have the good volatility forecast performance, both in-the-sample and out-of-sample on real data. Section 2.4 studies the volatility forecast performance of the modified BSIV in Mincer-Zarnowitz regression.

2.2 The Black-Scholes implied volatility as a proxy for volatility forecast

In this section, the impact of nonzero leverage effect will be examined through the generalized Black-Scholes (GBS) option pricing formula. As explained in Section 2.1, we do not focus on the bias, possibly related to volatility risk premium. In other words, we can simply refer to the risk-neutral measure. We assume the return process:

$$d(\log S_t) = r_t dt + \sqrt{V_t} dW_t^s$$

and the volatility process:

$$dV_t = f_1(V_t)dt + f_2(V_t)dW_t^v,$$

where $Corr(dW_t^s, dW_t^v) = \rho$ and r_t is a risk-free interest rate. For convenience, we can assume r_t is constant, r. Under this stochastic volatility model setting, the European call option can be priced with GBS option pricing formula. It is given as

$$C_t \quad = \quad E_t^Q[BS(S_t \xi_{t,T}, \overline{\sigma_{t,T}}^2)], \tag{2.3}$$

where

$$\begin{aligned} \xi_{t,T} &= \exp\left(\rho \int_t^T \sqrt{V_u} dW_u^v - \frac{1}{2}\rho^2 \int_t^T V_u du\right) \\ \overline{\sigma_{t,T}}^2 &= \frac{1-\rho^2}{T-t} \int_t^T V_u du. \end{aligned}$$

Contrary to the case of the zero leverage effect, the linkage between the BS implied volatility and the future volatility forecast is not obvious because of two factors in (2.3). $\xi_{t,T}$ and $1-\rho^2$ appear in front of S_t and $\frac{1}{T-t}\int_t^T V_u du$, respectively. In the stochastic volatility setting, the use of the implied volatility as a proxy for the forecast of the future volatility implies the following relationship:

$$\begin{aligned} C_t &= E_t^Q\left[BS\left(S_t\xi_{t,T}, \frac{1-\rho^2}{T-t}\int_t^T V_u du\right)\right] & (2.4)\\ &\approx BS\left[S_t, E_t\left(\frac{1}{T-t}\int_t^T V_u du\right)\right]. & (2.5)\end{aligned}$$

Here is an interpretation from comparison between equation (2.4) and equation (2.5). Note that $(1-\rho^2)$ in front of the average future volatility in the equation (2.4) decreases GBS option price since the Black-Scholes formula is an increasing function of the variance of the underlying asset. On the other hand, the Jensen's effect of $\xi_{t,T}$ makes the GBS option price bigger, since the Black-Scholes formula is convex with respect to the underlying asset price and $E_t^Q[\xi_{t,T}] = 1$. These two effects by nonzero leverage effect work in the opposite directions and can be canceled out, depending on the convexity of Black-Scholes formula with respect to the asset, the gamma effect. This cancelation will justify the use of the implied volatility as the proxy. In other words, using the implied volatility for the volatility forecast is justified if nonzero leverage effect that shows up in two places, $\xi_{t,T}$ and $\overline{\sigma_{t,T}^2}$, does not have an impact in the GBS formula. In Section 2.2.1, we provide the simple theoretical framework that gives a better understanding of the impact of the nonzero leverage effect.

2.2.1 The leverage effect on option prices

To investigate the impact of the leverage effect in GBS option prices, we consider the first derivative of the GBS formula with respect to ρ. We can separate the derivative into two, where each represents the effect associated with $\xi_{t,T}$ and $\overline{\sigma_{t,T}}^2$. Lebesgue's dominated convergence theorem allows us to do differentiation through the expectation. We can obtain the following decomposition of $\frac{\partial C_t}{\partial \rho}$:

Proposition 2.2.1.

$$\begin{aligned}\frac{\partial C_t}{\partial \rho} &= \frac{\partial C_t}{\partial (S_t\xi_{t,T})}\frac{\partial (S_t\xi_{t,T})}{\partial \rho} + \frac{\partial C_t}{\partial \overline{\sigma_{t,T}}}\frac{\partial \overline{\sigma_{t,T}}}{\partial \rho}\\ &= S_t E_t^Q\left[\xi_{t,T}\left(\int_t^T \sqrt{V_u}dW_u^v - \rho\int_t^T V_u du\right)\Phi(d_1)\right] - S_t E_t^Q\left[\xi_{t,T}\phi(d_1)\frac{\rho}{\sqrt{1-\rho^2}}\sqrt{\int_t^T V_u du}\right],\end{aligned}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and probability density function of a standard normal distribution, respectively.

Proof of Proposition 2.2.1

Recall that the GBS option pricing formula is

$$C_t = E_t^Q\left[S_t\xi_{t,T}\Phi(d_1) - K\exp(-r(T-t))\Phi(d_2)\right]$$

where

$$d_1 = \frac{1}{\overline{\sigma_{t,T}}\sqrt{T-t}}\left[\log\left(\frac{S_t\xi_{t,T}}{K\exp(-r(T-t))}\right)\right] + \frac{1}{2}\overline{\sigma_{t,T}}\sqrt{T-t}$$

$$d_2 = \frac{1}{\overline{\sigma_{t,T}}\sqrt{T-t}}\left[\log\left(\frac{S_t\xi_{t,T}}{K\exp(-r(T-t))}\right)\right] - \frac{1}{2}\overline{\sigma_{t,T}}\sqrt{T-t}$$

Let $F = S_t\xi_{t,T}\Phi(d_1) - K\exp(-r(T-t))\Phi(d_2)$ for convenience. Then,

$$\frac{\partial F}{\partial \rho} = \frac{\partial F}{\partial(S_t\xi_{t,T})}\frac{\partial(S_t\xi_{t,T})}{\partial \rho} + \frac{\partial F}{\partial \overline{\sigma_{t,T}}}\frac{\partial \overline{\sigma_{t,T}}}{\partial \rho}. \tag{2.6}$$

Since

$$S_t\xi_{t,T}\phi(d_1) = K\exp(-r(T-t))\phi(d_2),$$

we get

$$\begin{aligned}\frac{\partial F}{\partial(S_t\xi_{t,T})} &= \Phi(d_1) + S_t\xi_{t,T}\phi(d_1)\frac{1}{\overline{\sigma_{t,T}}\sqrt{T-t}}\frac{1}{S_t\xi_{t,T}} - K\exp(-r(T-t))\phi(d_2)\frac{1}{\overline{\sigma_{t,T}}\sqrt{T-t}}\frac{1}{S_t\xi_{t,T}} \\ &= \Phi(d_1), \end{aligned} \tag{2.7}$$

$$\frac{\partial(S_t\xi_{t,T})}{\partial \rho} = S_t\xi_{t,T}\left(\int_t^T V_u dW_u^v - \rho\int_t^T V_u du\right), \tag{2.8}$$

$$\frac{\partial F}{\partial \overline{\sigma_{t,T}}} = \xi_{t,T}\phi(d_1)\sqrt{T-t} \tag{2.9}$$

and

$$\frac{\partial \overline{\sigma_{t,T}}}{\partial \rho} = \rho\sqrt{\frac{\int_t^T V_u du}{(T-t)(1-\rho^2)}}. \tag{2.10}$$

By putting the equations (2.7),(2.8),(2.9) and (2.10) into (2.6), we obtain the proposition. A simulation result based on this proposition will follow in Section 2.2.2.

2.2.2 Simulation study

We conduct simulation studies to see the impact of nonzero leverage effect. We assume the volatility process in Heston (1993):

$$dV_t = k_v(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t.$$

The simulated volatilities will be used to demonstrate the numerical illustration of Proposition 2.2.1 and to generate option prices. We adopt a simulation method suggested by Brigo and Alfonsi (2005). More details on the simulation and parameters are included in Section 1.4.4. Parameters are from Table 1.1 and the results reported in this section are based on Set2. The moneyness x_t is defined as

$$x_t = \log\left(\frac{S_t}{K\exp(-r(T-t))}\right), \tag{2.11}$$

where K is a strike price, t is current time and T is a maturity time.

The numerical presentation of Proposition 2.2.1 is in Figure 2.1 and Figure 2.2. Figure 2.1 shows how two Jensen effects caused by the nonzero leverage effect change. Except for the option with the longest maturity time, 12 months, the gamma effect is so big that two effects with opposite directions are canceled out for at-the-money options. Figure 2.2 implies the same. Total impact from nonzero leverage effect on the option price becomes close to zero for at-the-money options. Figure 2.3 is to see the actual BSIV values with the change of moneyness and maturity time and compare them with the future volatility. Figure 2.3 has

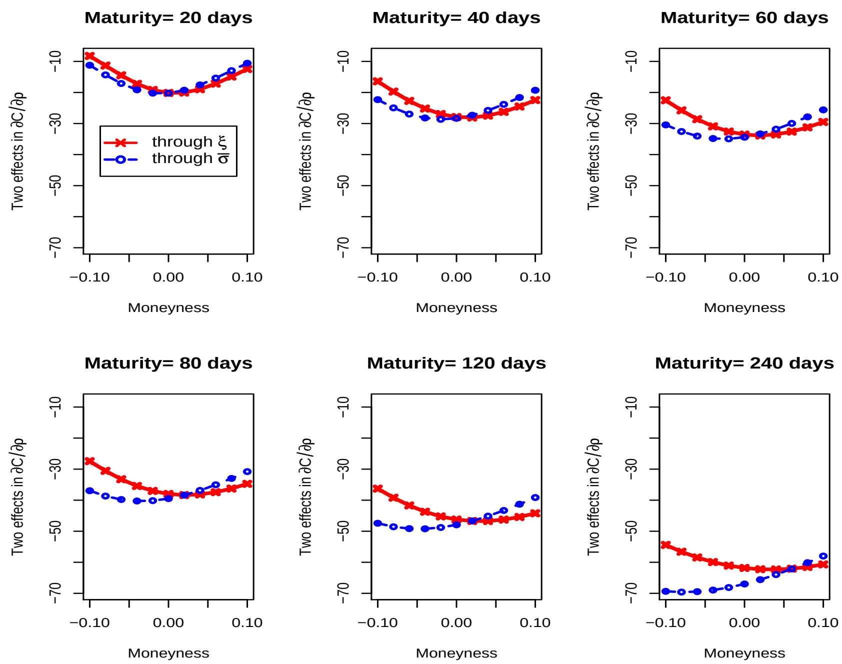

Figure 2.1: Decomposition of $\partial C_t/\partial\rho$

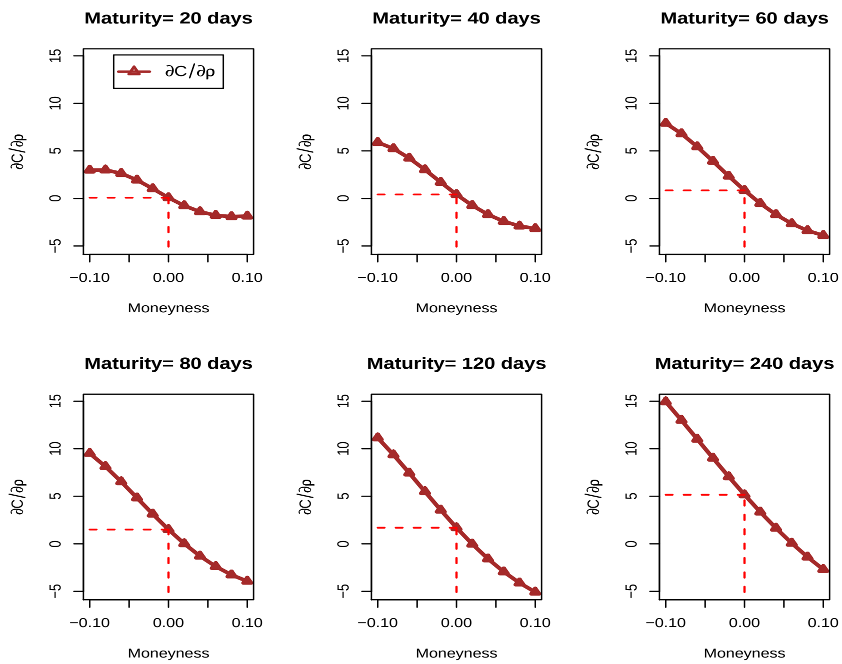

Figure 2.2: The change of $\partial C_t/\partial\rho$

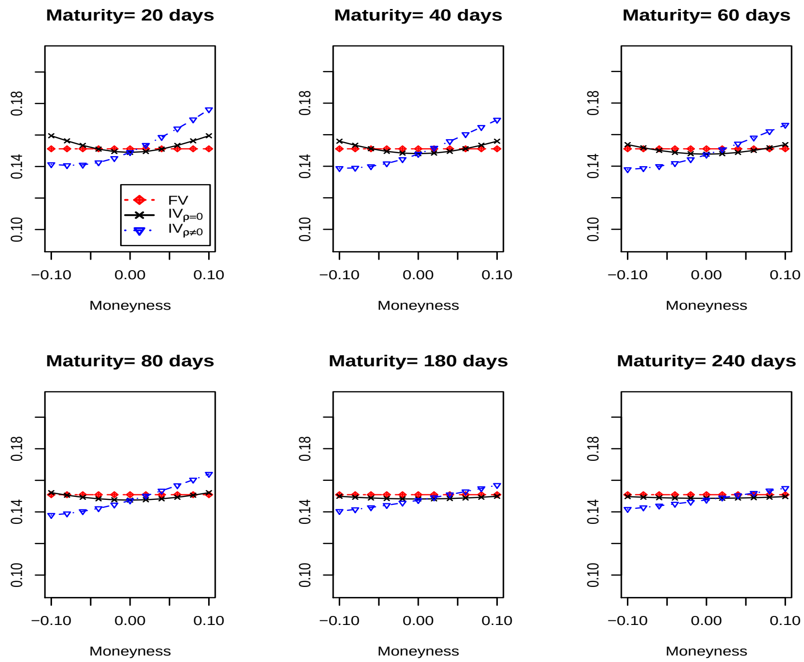

Figure 2.3: BSIV as a proxy for the forecast of the future volatility

three different lines. The first line (FV) is the averaged integrated volatility generated from the above volatility process. The second line ($IV_{\rho=0}$) is BSIV from the simulated option prices using GBS option pricing formula when $\rho = 0$. The third line ($IV_{\rho\neq0}$) is similarly BSIV when $\rho \neq 0$, which gives a further interpretation. First, focusing on at-the-money options that are often used for the forecast of the future volatility, we observe that the change of the leverage effect does not affect the option prices with the short maturity. Thus, we conclude that BSIV from at-the-money options with short maturity can be considered as a reasonable proxy for the forecast of the future volatility. Although the leverage effect has an impact on the at-the-money options with long maturity, it is shown in the simulation study that it remains minor with the different parameter values and does not cause an extra bias. As for the in-the-money and out-of-money options, two effects from the leverage effect don't get canceled and option prices calculated from GBS option pricing formula under zero and nonzero leverage effect does not coincide and neither of them are close to the forecast of the future volatility because of the smile effect. However, Figure 2.3 confirms that the randomness in volatility and the smile effect disappear as the maturity increases. Two implied volatilities and the forecast of the future volatility do not differ much.

2.3 The modified Black-Scholes option pricing formula

From the simulation study in Section 2.2, we learn that BSIV from at-the-money option with the short maturity time is a good proxy for the future volatility, since the impact of the nonzero leverage effect is canceled for at-the-money option. However, this does not hold for in-the-money and out-of-money options, since there is the remaining leverage effect and the smile effect. Thus, to infer the right volatility forecasts from options, especially from in-the-money and out-of-money options, we propose the modified BS option pricing formula, where the stock price in the BS option pricing formula is replaced with the implied stock price. The implied stock price is obtained by inverting BS option pricing formula given all other information including the volatility, σ. This approach was introduced in Longstaff (1995) and followed by Garcia, Luger and Renault (2004). They estimate both volatility and the implied stock price simultaneously. We propose a slightly different approach. We plug BSIV from at-the-money option, which is free from the leverage effect, into the place for volatility in BS option pricing formula and obtain a new version of the observed implied stock price by inverting BS option pricing formula. Based on the result in Section 2.2, the implied stock price captures all the impact of the nonzero leverage effect (skewness) and the smile effect with respect to the moneyness. We fit a model for the observed implied stock price and use the estimate of the implied stock price in the BS option pricing formula. We call this the modified BS option pricing formula. In other words, the modified BS option pricing formula consists of three steps:

step-(i) For each day, obtain implied volatility from at-the-money. If it does not exist, get the volatility from near-the-money option. We denote it as $\sigma^2_{imp,x=0}$, where $x = 0$ means the moneyness is 0, that is, at-the-money.

step-(ii) Obtain the observed implied stock prices S_t^* through

$$C_t = BS(S_t^*, \sigma^2_{imp,m=0}).$$

Due to the nonlinearity of the stock price in BS formula, the inversion is done through a numerical method, such as Newton-Ralphson method.

step-(iii) For forecast, first get an accurate $\widehat{S_t^*}$, the estimate of S_t^* and plug it with the BS implied volatility into BS formula.

Furthermore, we can think about a modified BS implied volatility, $\sigma_{imp,mod}$. The modified implied volatility is obtained by inverting

$$C_t = BS(\widehat{S_t^*}, \sigma^2_{imp,mod}),$$

where $\widehat{S_t^*}$ is the estimated adjusted stock price. If the modified option pricing formula works well and we can successfully capture the leverage effect and the smile effect in $\widehat{S_t^*}$, $\sigma^2_{imp,mod}$ from options with any moneyness is free from any effect and is a good proxy for the future volatility. This will be investigated more in Section 2.4.

S_t^* is a function of moneyness and the leverage effect, given the fixed maturity time. It is reasonable to assume to have a constant leverage effect since most conventional stochastic volatility models have a constant leverage effect. The time to maturity is given as 1 month in most volatility forecast literature. Therefore, S_t^* becomes the function of one parameter, the moneyness. We know from GBS formula that the relationship between S_t^* and the moneyness will not be linear. To capture the non-linearity, we fit the polynomial regression of degree 2 and degree 3.

When we estimate the coefficients of the polynomial regression, we use different loss functions. This is based on Christoffersen and Jacobs (2004). They emphasize the critical impact of the choice of loss functions in the estimation and the evaluation, although different loss functions at the estimation and evaluation stages are widely used in the literature. For example, if we fit the regular ordinary least square (OLS) regression, then the use of the mean-squared errors in option prices to evaluate the performance can lead to a wrong conclusion. We use three different loss functions in the estimation of the polynomial regression.

More specifically, our polynomial regression is:

$$\log(S_i^*/S_i) = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon$$

where X is moneyness. The estimation is akin to solving

$$\widehat{\beta} = \arg\min_{\beta} L(\beta), \tag{2.12}$$

where $L(\beta)$ is a loss function.

The first loss function is the mean residual sum of squares, $MRSS(\beta) = \frac{1}{n}\sum_{i=1}^{n}(R_i - R_i(\beta))^2$, where $R_i = \log(S_i^*/S_i)$ and $R_i(\beta) = \log(S_i^*(\beta)/S_i)$. Solving (2.12) is the common OLS. Two additional loss functions we consider are the following: the mean-squared dollar errors

$$\$MSE(\beta) \equiv \frac{1}{n}\sum_{i=1}^{n}(C_i - C_i(S_i^*(\beta)))^2$$

and the relative mean-squared errors

$$\%MSE(\beta) \equiv \frac{1}{n}\sum_{i=1}^{n}((C_i - C_i(S_i^*(\beta)))/C_i)^2.$$

where C_i is option data, $C_i(S_i^*(\beta))$ is the model option price.

2.3.1 Simulation study

In this section, we check the accurate and robust estimation of the adjusted stock price using a polynomial regression with simulation study. We consider three different leverage effect, -0.7, -0.5, -0.3, which is reasonable in the index option pricing literature. Moneyness is fixed to 21 points. As for the maturity time, 20,40, 60, 80, 180, 240 days are considered. This way, we generate 21 option prices for each maturity time with one fixed leverage effect. Here, the option price is purely a function of moneyness and so is the adjusted stock price. Following step-(i), step-(ii), step-(iii) described above, we get S_i^*, the adjusted stock prices. The result of polynomial regressions are illustrated in Figure 2.4, Figure 2.5 and Figure 2.6. The R-squares are close to 99% for all the combination of 6 maturity times and 3 ρ's. Using a 2-degree polynomial regression for the estimation of $\log(S_t^*/S_t)$ seems reasonable. Three different loss functions give very close results in the simulation study.

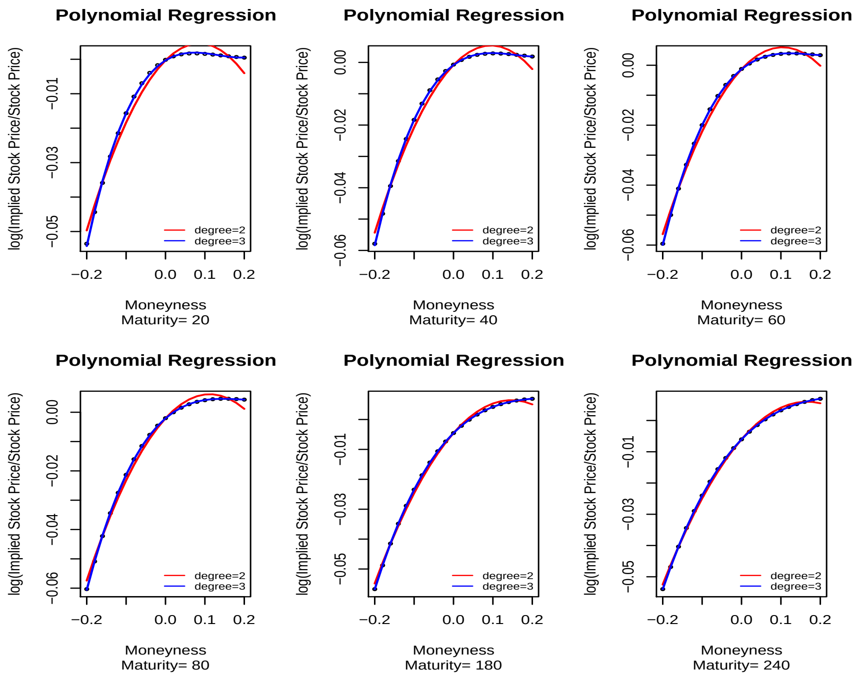

Figure 2.4: The Polynomial Regression when $\rho = -0.7$

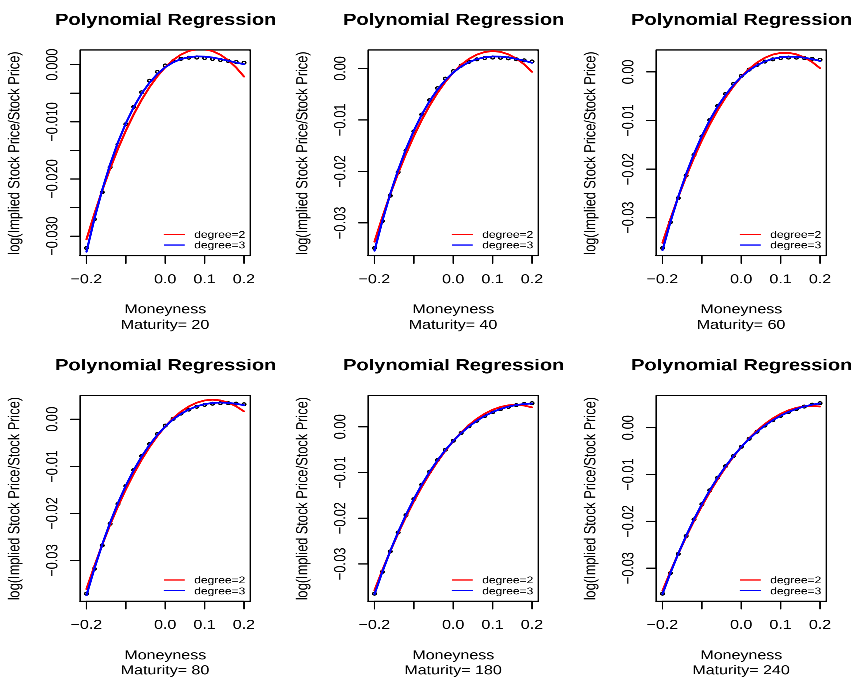

Figure 2.5: The Polynomial Regression when $\rho = -0.5$

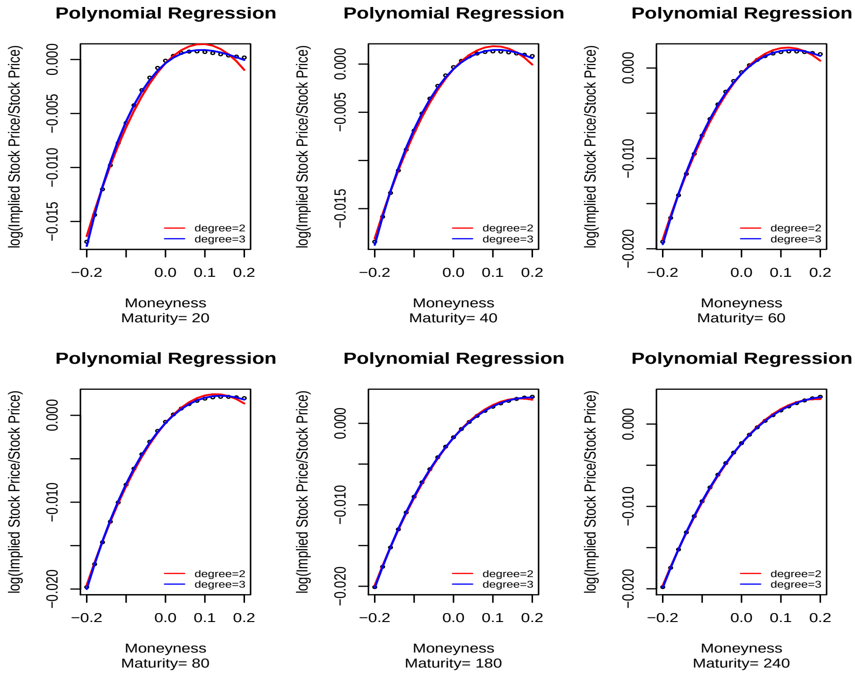

Figure 2.6: The polynomial regression when $\rho = -0.3$

2.3.2 Empirical Study

Our empirical study uses daily closing prices of S&P 500 index call options with a maturity time of 30 calendar days from January 1, 1997 to December 31, 2004. The S&P 500 options have largest trading volume in all options traded on the Chicago Board Options Exchange (CBOE). These options are European style and do not have complication from early exercise. The time to maturity is measured as the number of calendar days from the trade date to the Thursday immediately proceding the Friday when the option expires. We use the closing price of the S&P 500 index as an index level. Criterion for data filtration are adopted from some previous empirical studies, such as Bakshi, Cao and Chen (1997). First, general arbitrage violations are eliminated, since the violation leads to a negative implied volatility.

$$C \geq \max(0, S - K)$$

should be satisfied. Second, very deep out-of-money and very deep in-the-money options are excluded because they are not actively traded and have liquidity-related biases. If the absolute value of the moneyness is greater 0.08, then the option price is discarded. Third, quotes less than 3/8 are eliminated. Fourth, options are chosen not to have overlaps.

Our data filter yields a final daily sample of 1,898 observations.

	$x \leq -0.05$	$-0.05 < x \leq -0.02$	$-0.02 < x < 0.02$	$0.02 \leq x < 0.05$	$x \geq 0.05$
1997($n = 275$)	29.09%	18.91%	11.27%	21.09%	19.64%
1998($n = 262$)	32.44%	14.89%	12.97%	19.85%	19.85%
1999($n = 269$)	30.86%	20.82%	13.01%	22.68%	12.64%
2000($n = 231$)	30.74%	22.08%	17.75%	17.75%	11.69%
2001($n = 155$)	27.74%	25.81%	14.19%	18.71%	13.55%
2002($n = 177$)	32.20%	19.77%	18.64%	18.08%	11.30%
2003($n = 218$)	32.57%	19.27%	12.84%	22.02%	13.30%
2004($n = 311$)	31.19%	20.90%	14.47%	19.93%	13.50%

Table 2.1: The percentage of observations in each group and year (1997-2004). x is moneyness as defined in (2.11)

The risk-free interest rates are obtained from zero-coupon data by matching with the closest maturity time to option expiration. The data is not adjusted for dividend, following the approach in Christensen and Prabhala (1998). They claim that the reduction in call option values will lead to the reduction in implied volatility and the changes are almost constant. More careful study on this adjustment can be done using the future daily dividends (refer to Bakshi, Cao and Chen (1997)). All data in the empirical study are provided by Wharton Research Data Services.

The first empirical study is a continuation of our simulation study that checks the accuracy and the robustness of the estimation of the adjusted stock price using a polynomial regression. The difference is that we only use options with 20-day (equivalently, one month) maturity time. For the calculation of $\sigma_{imp,m=0}$, when at-the-money option is not available, find the option price that has the closest at-the-money option. We observe that using a whole data set does not give any useful information. To deal with noisy market data, we split the data set by one year. Eight sub-data sets are ready for in-sample and out-of-sample forecast. Figure 2.7 is an indication that putting a naive BS implied volatility from at-the-money option and the market stock price into BS option price formula will not give an accurate option price.

We fit polynomial regression for $\log(S^*/S)$ with three different loss functions described

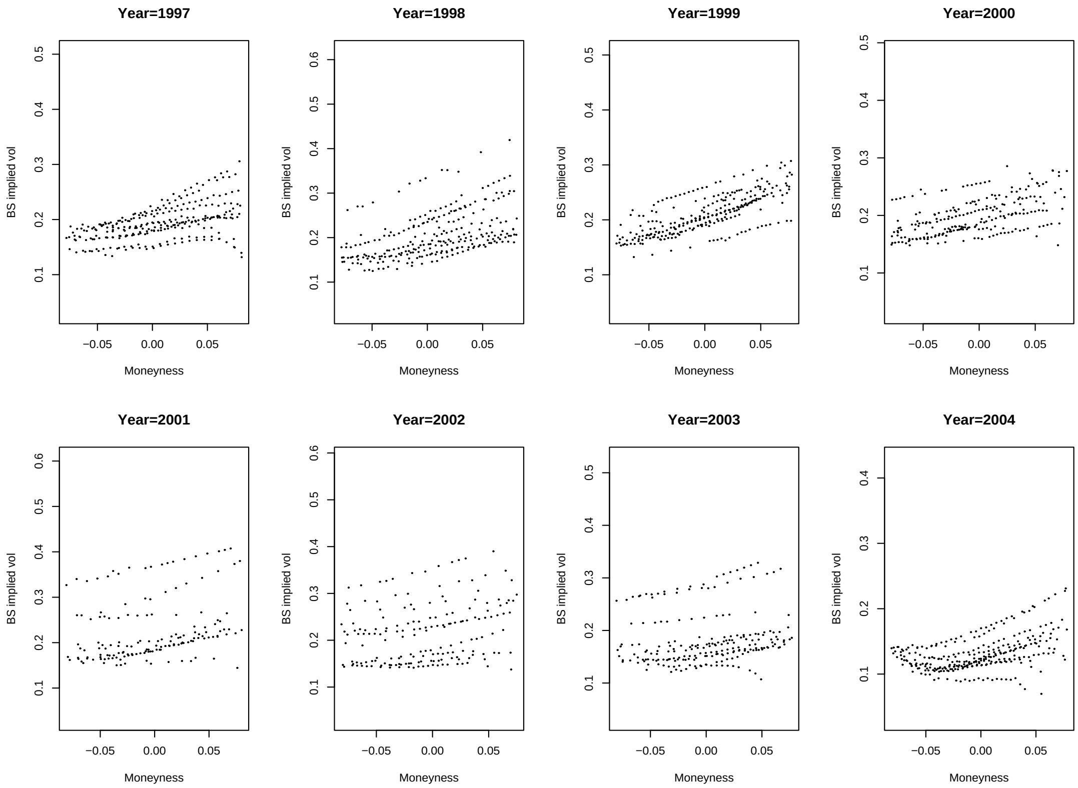

Figure 2.7: BSIV based on the market data during 1997-2004. The plots demonstrate the violation of constant volatility and a proper adjustment is needed

in Section 2.3.1. Figure 2.8 shows that depending on loss functions, the results are different. Especially, OLS estimation gives straight lines that do not capture the curvature in the data.

Market data has noise and we do not see as a perfect polynomial fit as in the simulation. Since we do not know the exact relationship between $\log(S^*/S)$ and the moneyness, we also try the nonparametric kernel regression. For details of this estimation method, refer to Appendix G. Figure 2.9 displays the adjusted stock prices and the nonparametric regression fit. For each year, the nonparametric regression fit seems to be good. However, some years have a different pattern and the out-of-sample is questionable. More careful study with out-of-sample test will follow.

Table 2.2 and Table 2.3 report results of in-sample and out-of-sample forecast during

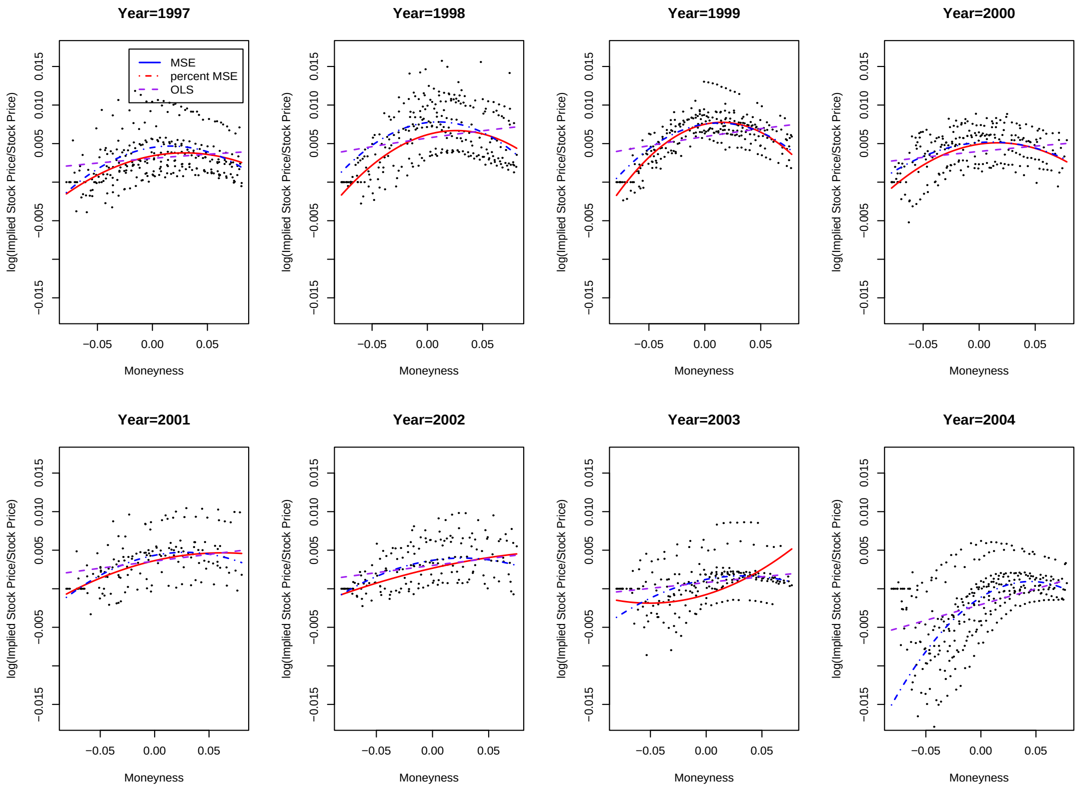

Figure 2.8: The observed $\log(S^*/S)$ and the fitted polynomial regression of $\log(S^*/S)$ using three different loss functions using the market data during 1997-2004

$1997-2004$. The columns represent the estimation methods and loss functions. The rows represent errors for evaluation purposes. All four cases of the modified BS option pricing, we have a much better performance than the traditional BS option pricing formula. Also, we see that the results for polynomial regression confirms the claim in Christoffersen and Jacobs (2004). That is, the choice of loss functions has a critical role in estimation. The strange phenomenon of polynomial regression fitting with loss function $MSE and %MSE in 2004 is still under investigation.

		modified BS				BS
		PR with %MSE	PR with $MSE	PR with MRSS	NP	
1997	%-errors	0.0056	0.0062	0.0076	0.0054	0.0103
	$-errors	2.5490	2.4012	2.7637	2.5278	3.5082
1998	%-errors	0.0061	0.0114	0.0136	0.0108	0.0297
	$-errors	5.5369	5.0463	6.5843	6.4432	8.1379
1999	%-errors	0.0024	0.0034	0.0080	0.0056	0.0976
	$-errors	2.418	2.3813	3.9456	2.781	4.036
2000	%-errors	0.059	0.0071	0.0086	0.0068	0.0138
	$-errors	3.2157	3.1732	3.7758	3.2055	4.5721
2001	%-errors	0.0028	0.0030	0.0041	0.0029	0.0068
	$-errors	3.5678	3.4200	3.6533	3.451	4.7379
2002	%-errors	0.0035	0.0041	0.0094	0.0041	0.0168
	$-errors	2.4765	2.2550	2.3819	2.3960	3.6518
2003	%-errors	0.0064	0.0085	0.0094	0.0076	0.0124
	$-errors	2.5992	1.5161	1.516	1.6549	2.426
2004	%-errors	1.000	0.0600	0.0712	0.0615	0.0989
	$-errors	1198.6328	2.6394	3.2038	2.874	4.573

Table 2.2: The in-sample result during year $1997 - 2004$. Only errors in option prices are reported for the comparison to BS pricing. For the second-order polynomial regression (PR), we use three different loss functions (%MSE,$MSE, MRSS). NP represents nonparametric kernel regression. We compare four different modified BS results with traditional BS pricing

		modified BS				BS
		PR with %MSE	PR with $MSE	PR with MRSS	NP	
1998	%-errors	0.0095	0.0075	0.011	0.0082	0.0335
	$-errors	10.4206	8.7134	10.7488	9.5489	13.2046
1999	%-errors	0.0033	0.0043	0.0078	0.0025	0.0174
	$-errors	2.9486	2.3945	3.9297	2.3480	3.5068
2000	%-errors	0.0085	0.0105	0.0138	0.0088	0.0359
	$-errors	6.3274	6.3857	6.7231	6.6104	8.047
2001	%-errors	0.0035	0.0051	0.0052	0.0042	0.0873
	$-errors	3.5200	3.5916	3.6241	3.5417	4.975
2002	%-errors	0.0039	0.0047	0.0067	0.0050	0.0104
	$-errors	2.4501	2.4131	2.4599	2.6479	2.889
2003	%-errors	0.0145	0.0253	0.0211	0.0204	0.0352
	$-errors	3.3590	3.4053	3.4049	4.2061	5.002
2004	%-errors	0.1126	0.1419	0.1775	0.1483	0.1976
	$-errors	4.8777	3.5877	3.5768	3.7662	5.490

Table 2.3: The out-of-sample result during year $1997-2004$. Only errors in option prices are reported for the comparison to BS pricing. For the second-order polynomial regression (PR), we use three different loss functions (%MSE,$MSE, MRSS). NP represents nonparametric kernel regression. We compare four different modified BS results with traditional BS pricing

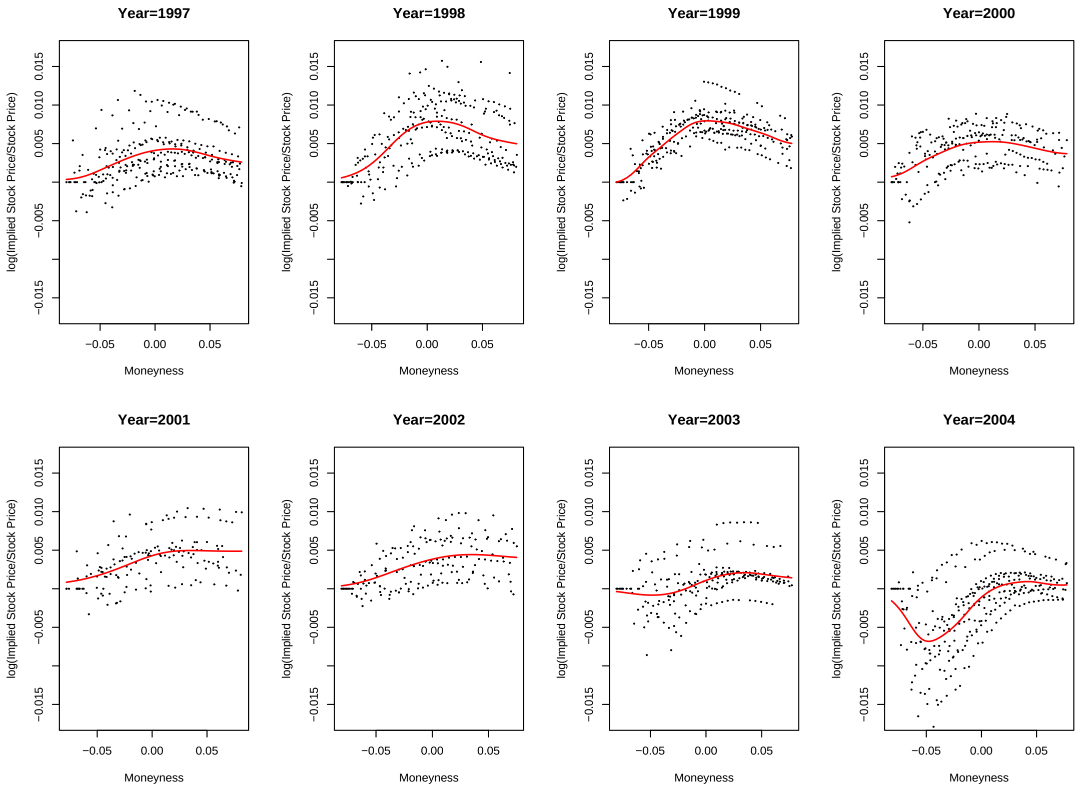

Figure 2.9: The observed $\log(S^*/S)$ and the fitted nonparametric regression of $\log(S^*/S)$ using the market data during 1997-2004

2.4 Volatility forecast with the modified Black-Scholes implied volatility

We expand our study into the modified Black-Scholes implied volatility. If the skewness and the smile effect with respect to moneyness are captured well, the modified BS implied volatilities are close to BS implied volatility from at-the-money option. Figure 2.10 and Figure 2.11 show that some effects are not fully captured.

Before we conclude that there are remaining effects that are not captured by our scheme, we want to see the relationship between the modified BS implied volatility and the realized volatility. We fit Mincer-Zarnowitz type regression of realized volatility. When we estimate the adjusted stock price S_t^*, we use \$MSE loss function. Also, we include options with moneyness x_t between -0.05 and 0.05. We compare this with the relationship between

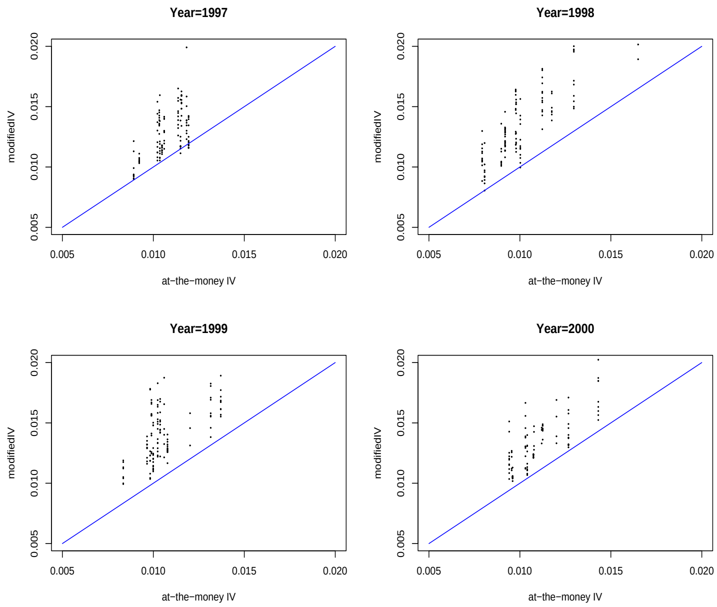

Figure 2.10: Scatter plot of our modified BSIV from all ranges vs. BSIV from at-the-money option during $1997 - 2000$

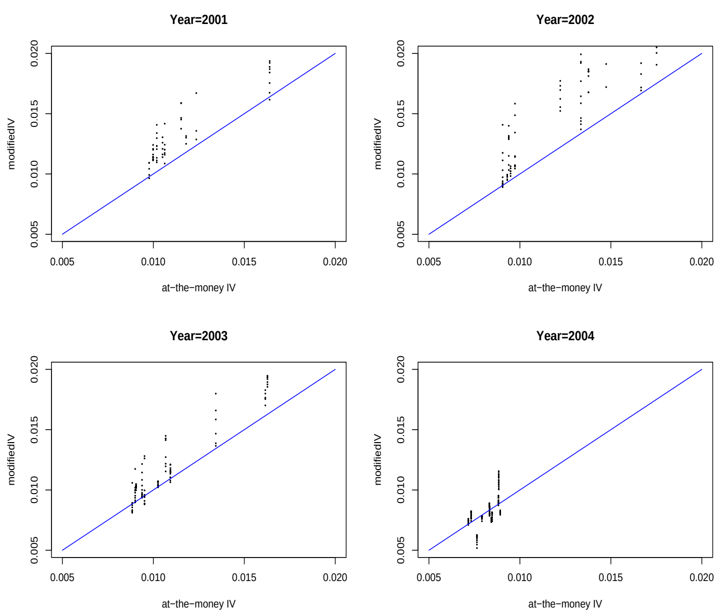

Figure 2.11: Scatter plot of our modified BSIV from all ranges vs. BSIV from at-the-money option during $2001 - 2004$

BS implied volatility from at-the-money option and the realized volatility. The realized volatility is defined as

$$R_{t,t+20} = \sum_{i=1}^{20} \left(\log(\frac{S_{t+i}}{S_{t+i-1}}) \right)^2 .$$

The predictive regression is

$$R_{t,t+20} = a + b\sigma^2,$$

where σ is the BS implied volatility or the modified BS implied volatility. Table 2.4 shows that the volatility forecast with the BS implied volatility from at-the-money option performs better. Our modified implied volatility captured the skewness and smile effect at some degree, but the unsatisfactory results in fitting regression here also indicates the limitation of this approach.

	modified IV			BS IV		
	a	b	R-squared	a	b	R-squared
in-sample(year = 1997)	0.033	0.637	0.441	0.029	0.703	0.627
in-sample(year = 1998)	0.045	0.541	0.428	0.036	0.762	0.613
in-sample(year = 1999)	0.040	0.512	0.426	0.033	0.788	0.691
in-sample(year = 2000)	0.032	0.492	0.471	0.021	0.732	0.703
in-sample(year = 2001)	0.036	0.441	0.385	0.026	0.717	0.702
in-sample(year = 2002)	0.028	0.473	0.400	0.012	0.693	0.598
in-sample(year = 2003)	0.029	0.397	0.419	0.014	0.638	0.614
in-sample(year = 2004)	0.031	0.376	0.424	0.017	0.594	0.562
out-of-sample(year = 1998)	0.046	0.482	0.478	0.041	0.442	0.576
out-of-sample(year = 1999)	0.039	0.421	0.411	0.030	0.543	0.583
out-of-sample(year = 2000)	0.026	0.452	0.478	0.037	0.601	0.588
out-of-sample(year = 2001)	0.036	0.467	0.398	0.026	0.574	0.490
out-of-sample(year = 2002)	0.028	0.451	0.472	0.033	0.573	0.524
out-of-sample(year = 2003)	0.031	0.413	0.376	0.021	0.521	0.497
out-of-sample(year = 2004)	0.036	0.305	0.341	0.024	0.429	0.479

Table 2.4: The predictive regression of the realize volatility and the modified BS implied volatility and the regression of the realized volatility and the traditional BS implied volatility from at-the-money option

APPENDIX A

Proof of Lemma 1.2.1

Note that

$$\begin{aligned} B_t^\pi &= \pi_t \exp\left(\int_0^t r(X_u)du\right) \\ &= \exp\left(\log \pi_t + \int_0^t r(X_u)du\right) \\ &= f(\log \pi_t, t). \end{aligned}$$

Thus, Ito's formula gives

$$\begin{aligned} dB_t^\pi &= \frac{\partial B^\pi}{\partial t}dt + \frac{\partial(B^\pi)}{\partial(\log \pi)}d(\log \pi_t)^c + \frac{1}{2}\frac{\partial^2(B^\pi)}{\partial(\log \pi)^2}d[\log \pi, \log \pi]_t^c + [B_t^\pi - B_{t-}^\pi]dN_t \\ &= r(X_t)B_t^\pi dt + [h(X_t)dt + a(X_t)dW_{1t} + b(X_t)dW_{2t}]B_t^\pi + \frac{1}{2}[a^2(X_t) + b^2(X_t)]B_t^\pi dt \\ &\quad +(B_t^\pi - B_{t-}^\pi)dN_t, \end{aligned}$$

where $(\log \pi_t)^c$ and $[\log \pi, \log \pi]^c$ denote continuous components of $(\log \pi_t)$ and $[\log \pi_t, \log \pi_t]$, respectively.

$(B_t^\pi - B_{t-}^\pi)dN_t = \left(\frac{B_t^\pi}{B_{t-}^\pi} - 1\right) B_{t-}^\pi dN_t = (e^{c_t} - 1)B_{t-}^\pi dN_t$ is compensated by $E[e^{c_t} - 1]B_{t-}^\pi \lambda(X_t)dt$. Thus,

$$\begin{aligned} dB_t^\pi &= \left[r(X_t) + h(X_t) + \frac{1}{2}a^2(X_t) + \frac{1}{2}b^2(X_t) + E[e^{c_t} - 1]B_{t-}^\pi \lambda(X_t)\right] B_t^\pi dt \\ &\quad +[a(X_t)dW_{1t} + b(X_t)dW_{2t}]B_t^\pi + (e^{c_t} - 1)B_{t-}^\pi dN_t - E(e^{c_t} - 1)B_t^\pi \lambda(X_t)dt \end{aligned} \tag{A.1}$$

B_t^π is a local martingale if and only if the drift term in (A.1) is zero. That is,

$$r(X_t) + h(X_t) + \frac{1}{2}a^2(X_t) + \frac{1}{2}b^2(X_t) + E_t(e^{c_t} - 1)\lambda(X_t) = 0.$$

APPENDIX B

Proof of Lemma 1.2.2

Note that

$$\begin{aligned}
{\varphi_t}^{\pi} &= \pi_t S_t \\
&= \exp(\log \pi_t + \log S_t) \\
&= g(\log \pi_t, \log S_t).
\end{aligned}$$

By Ito's formula,

$$\begin{aligned}
&d\varphi_t^{\pi} \\
&\quad = \frac{\partial(\varphi^{\pi})}{\partial(\log \pi)} d(\log \pi_t)^c + \frac{\partial(\varphi^{\pi})}{\partial(\log S)} d(\log S_t)^c + \frac{1}{2}\frac{\partial^2(\varphi^{\pi})}{\partial(\log \pi^2)} d[\log \pi, \log \pi]_t^c \\
&\quad + \frac{1}{2}\frac{\partial^2(\varphi^{\pi})}{\partial(\log S^2)} d[\log S, \log S]_t^c + \frac{\partial^2(\varphi^{\pi})}{\partial(\log \pi)\partial(\log S)} d[\log \pi, \log S]_t^c + [\varphi_t^{\pi} - \varphi_{t-}^{\pi}] dN_t \\
&\quad = [h(X_t)dt + a(X_t)dW_{1t} + b(X_t)dW_{2t}]\varphi_t^{\pi} + [\mu(X_t)dt + \alpha(X_t)dW_{1t} + \beta(X_t)dW_{2t}]\varphi_t^{\pi} \\
&\quad + \frac{1}{2}[a^2(X_t) + b^2(X_t)]\varphi_t^{\pi} dt + \frac{1}{2}[\alpha_t^2 + \beta_t^2]\varphi_t^{\pi} dt + (a(X_t)\alpha(X_t) + b(X_t)\beta(X_t))\varphi_t^{\pi} dt + (\varphi_t^{\pi} - \varphi_{t-}^{\pi})dN_t.
\end{aligned}$$

Note that $(\varphi_t^{\pi} - \varphi_{t-}^{\pi})dN_t = (\frac{\varphi_t^{\pi}}{\varphi_{t-}^{\pi}} - 1)\varphi_{t-}^{\pi} dN_t = (e^{c_t+\gamma_t} - 1)\varphi_{t-}^{\pi} dN_t$ is compensated by $E[e^{c_t+\gamma_t} - 1]\varphi_t \lambda(X_t)$. By the result in Lemma 1.2.1, we obtain

$$\begin{aligned}
&d\varphi_t^{\pi} \\
&\quad = \Big[a(X_t)\alpha(X_t) + b(X_t)\beta(X_t) + \frac{1}{2}(\alpha^2(X_t) + \beta^2(X_t)) + \mu(X_t) - r(X_t) - E\left(e^{c_t} - 1\right)\lambda(X_t) \\
&\quad + E\left(e^{c_t+\gamma_t} - 1\right)\lambda(X_t)\Big]\varphi_t^{\pi} dt + [a(X_t)dW_{1t} + b(X_t)dW_{2t} + \alpha(X_t)dW_{1t} + \beta(X_t)dW_{2t}]\varphi_t^{\pi} \\
&\quad + \left(e^{c_t+\gamma_t} - 1\right)\varphi_t^{\pi} dN_t - \lambda(X_t)E\left(e^{c_t+\gamma_t} - 1\right)\varphi_t^{\pi} dt
\end{aligned}$$

φ_t^{π} is local martingale if and only if the drift term is zero:

$$a(X_t)\alpha(X_t) + b(X_t)\beta(X_t) + \frac{1}{2}(\alpha^2(X_t) + \beta^2(X_t)) + \mu(X_t) - r(X_t) - \lambda(X_t)[E\left(e^{c_t}\right) - E\left(e^{c_t+\gamma_t}\right)] = 0.$$

APPENDIX C

Proof of Proposition 1.2.1

We introduce new notations for simplicity:

$$E_t[\cdot] = E[\cdot|F_t \vee \sigma((X_u)_{t\le u\le T}), (N_u)_{t\le u\le T}]$$

$$Var_t[\cdot] = Var[\cdot|F_t \vee \sigma((X_u)_{t\le u\le T}), (N_u)_{t\le u\le T}]$$

$$Cov_t[\cdot] = Cov[\cdot|F_t \vee \sigma((X_u)_{t\le u\le T}), (N_u)_{t\le u\le T}]$$

$$\begin{aligned} &B^*_{t,T} \\ &\quad = E_t\left[\frac{\pi_T}{\pi_t}\right] \\ &\quad = E_t\left[\exp\left(\int_t^T h(X_u)du + \int_t^T a(X_u)dW_{1u} + \int_t^T b(X_u)dW_{2u} + \sum_{N_t<i\le N_T} c_i\right)\right] \end{aligned}$$

By Lemma 1.2.1 and $E_t\left[\exp\left(\int_t^T b(X_u)dW_{2u} - \frac{1}{2}\int_t^T b^2(X_u)du\right)\right] = 1$,

$$\begin{aligned} &B^*_{t,T} \\ &= \left[\exp\left(\int_t^T (-r(X_u) - \frac{1}{2}a^2(X_u) - \lambda(X_u)E[e^{c_t}-1])du + \int_t^T a(X_u)dW_{1u}\right)\right] E\left[\exp\left(\sum_{N_t<i\le N_T} c_i\right)\right] \\ &= \left[\exp\left(\int_t^T (-r(X_u) - \frac{1}{2}a^2(X_u) - \lambda(X_u)E[e^{c_t}-1])du + \int_t^T a(X_u)dW_{1u}\right)\right] [E(e^{c_t})]^{N_T-N_t} \end{aligned}$$

Thus,

$$B^*_{t,T} = \exp\left[-\int_t^T r(X_u)du\right]\eta_{t,T},$$

with

$$\eta_{t,T} \quad = \quad \exp\left[\int_t^T a(X_u)dW_{1u} - \frac{1}{2}\int_t^T a^2(X_u)du\right]$$

$$\exp\left[(N_T - N_t)\log\left[E\left(e^{c_t}\right)\right] - \left[E\left(e^{c_t}\right) - 1\right]\int_t^T \lambda(X_u)du\right]$$

APPENDIX D

Proof of Proposition 1.2.2

$$\xi_{t,T} = E_t\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t}\right].$$

By using

$$E_t\left[\exp\left(\int_t^T (b(X_u)+\beta(X_u))dW_{1u} - \frac{1}{2}\int_t^T (b(X_u)+\beta(X_u))^2 du\right)\right] \quad = \quad 1$$

and the results in Lemma 1.2.1 and Lemma 1.2.2,

$$\begin{aligned}
\xi_{t,T} \quad &= \quad E_t\Bigg[\exp\Bigg(\int_t^T h(X_u)du + \int_t^T a(X_u)dW_{1u} + \int_t^T b(X_u)dW_{2u} + \sum_{N_t<i\leq N_T} c_i \\
&\qquad \int_t^T \mu(X_u)du + \int_t^T \alpha(X_u)dW_{1u} + \int_t^T \beta(X_u)dW_{2u} + \sum_{N_t<i\leq N_T}\gamma_i\Bigg)\Bigg] \\
&= \quad \exp\left(\int_t^T (a(X_u)+\alpha(X_u))dW_{1u} - \frac{1}{2}\int_t^T (a(X_u)+\alpha(X_u))^2 du - E(e^{c_t+\gamma_t}-1)\int_t^T \lambda(X_u)du\right) \\
&\qquad E(e^{c_t+\gamma_t})^{N_T-N_t}
\end{aligned}$$

APPENDIX E

Proof of Proposition 1.3.1: the GBS option pricing formulas

Computing the price of an European call:

$$C_t = E\left[\frac{\pi_T}{\pi_t} Max[0, S_T - K] | F_t\right]$$

Equivalently,

$$\begin{aligned} \frac{C_t}{S_t} &= E\left[\frac{\pi_T}{\pi_t} Max[0, \frac{S_T}{S_t} - \frac{K}{S_t}] | F_t\right] \\ &= E\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} | F_t\right] - E\left[\frac{\pi_T}{\pi_t}\frac{K}{S_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} | F_t\right] \\ &= \frac{G_t}{S_t} - \frac{H_t}{S_t}, \end{aligned}$$

where

$$\frac{G_t}{S_t} = E\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} | F_t\right]$$

$$\frac{H_t}{S_t} = E\left[\frac{\pi_T}{\pi_t}\frac{K}{S_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} | F_t\right].$$

By the law of iterated expectations,

$$\frac{G_t}{S_t} = E\left[E\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} \mid F_t \vee \sigma((X_u)_{t \leq u \leq T}), (N_u)_{t \leq u \leq T}\right] | F_t\right] \tag{E.1}$$

$$\frac{H_t}{K} = E_t\left[E_t\left[\frac{\pi_T}{\pi_t} I_{[\frac{S_T}{S_t} \geq \frac{K}{S_t}]} \mid F_t \vee \sigma((X_u)_{t \leq u \leq T}), (N_u)_{t \leq u \leq T}\right] | F_t\right]. \tag{E.2}$$

The key of the derivation of GBS option pricing formula is bivariate conditional distribution of $\left[\log(\frac{S_T}{S_t}) - \log(\frac{K}{S_t}), \log(\frac{\pi_T}{\pi_t})\right]$ and $\left[\log(\frac{S_T}{S_t}) - \log(\frac{K}{S_t}), \log(\frac{\pi_T}{\pi_t}) + \log(\frac{S_T}{S_t})\right]$. The following two lemmas on a bivariate normal distribution are useful in the proof.

Lemma E.0.1. *If $[X_1, X_2]^T$ is a bivariate normal vector with $E[X_1, X_2]^T = [m_1, m_2]^T$, $Var[X_1] = {w_1}^2$,$Var[X_2] = {w_2}^2$, and $Cov[X_1, X_2] = \rho w_1 w_2$ then $E[(\exp X_1) I_{[X_2 \geq 0]}] =$*

$E(\exp X_1)\Phi(\frac{m_2+\rho w_1 w_2}{w_2})$, *where* Φ *is the cumulative normal distribution function.*

Lemma E.0.2. *If* $Y = \log(X)$ *has a normal distribution with mean,* μ_Y*, and variance,* σ_Y^2 *then the mean of the lognormal random variable X,* μ_X *is* $\exp(\mu_Y + \sigma_Y^2/2)$*. Thus,* μ_Y *can be written as* $\log(\mu_X) - \sigma_Y^2/2$.

E.1 The first part of the formula: G_t/S_t

By applying Lemma E.0.1 to E.1,

$$\begin{aligned}
&E_t\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t}I_{[\frac{S_T}{S_t}\geq\frac{K}{S_t}]}\right] \\
&\quad = E_t\left[\exp\left(\log(\frac{\pi_T}{\pi_t})+\log(\frac{S_T}{S_t})\right)I_{[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\geq 0]}\right] \\
&\quad = E_t\left[\frac{\pi_T}{\pi_t}\frac{S_T}{S_t}\right]\Phi[\frac{E_t\left[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\right]+Cov_t\left[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t}),\log(\frac{\pi_T}{\pi_t})+\log(\frac{S_T}{S_t})\right]}{\sqrt{Var_t\left[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\right]}}] \\
&= \quad \xi_{t,T}\Phi(d_1).
\end{aligned}$$

The remaining work is to simplify d_1 and express it as $B_{t,T}^*$ and $\xi_{t,T}$. Note the followings:

$$Cov_t\left(\log(\frac{S_T}{S_t}),\log(\frac{\pi_T}{\pi_t})+\log(\frac{S_T}{S_t})\right) = Var_t\left[\log(\frac{S_T}{S_t})\right]+Cov_t\left[\log(\frac{S_T}{S_t}),\log(\frac{\pi_T}{\pi_t})\right] \quad \text{(E.3)}$$

$$\begin{aligned}
E_t[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})] &= E_t\left[\log(\frac{S_T}{S_t})\right]-\log(\frac{K}{S_t}) \\
&= E_t\left[\log(\frac{S_T}{S_t})+\log(\frac{\pi_T}{\pi_t})\right]-E_t\left[\log(\frac{\pi_T}{\pi_t})\right]-\log(\frac{K}{S_t}). \quad \text{(E.4)}
\end{aligned}$$

By Lemma E.0.2,

$$\begin{aligned}
&E_t\left[\log(\frac{S_T}{S_t})+\log(\frac{\pi_T}{\pi_t})\right] \\
&\quad = \log\left[E_t\left(\frac{\pi_T}{\pi_t}\frac{S_T}{S_t}\right)\right]-\frac{1}{2}Var_t\left[\log(\frac{S_T}{S_t})+\log(\frac{\pi_T}{\pi_t})\right] \\
&\quad = \log(\xi_{t,T})-\frac{1}{2}Var_t\left[\log(\frac{S_T}{S_t})\right]-\frac{1}{2}Var_t\left[\log(\frac{\pi_T}{\pi_t})\right]-Cov_t\left[\log(\frac{S_T}{S_t}),\log(\frac{\pi_T}{\pi_t})\right] \quad \text{(E.5)}
\end{aligned}$$

$$
\begin{aligned}
E_t\left[\log(\frac{\pi_T}{\pi_t})\right] &= \log\left[E_t\left(\frac{\pi_T}{\pi_t}\right)\right] - \frac{1}{2}Var_t\left[\log(\frac{\pi_T}{\pi_t})\right] \\
&= \log(B^*(t,T)) - \frac{1}{2}Var_t\left[\log(\frac{\pi_T}{\pi_t})\right]. \qquad \text{(E.6)}
\end{aligned}
$$

Lastly,

$$
\begin{aligned}
Var_t\left[\log(\frac{S_T}{S_t})\right] &= Var_t\left[\int_t^T \mu(X_u)du + \int_t^T \alpha(X_u)dW_{1u} + \beta(X_u)dW_{2u} + \sum_{N_t<i\leq N_T}\gamma_i\right] \\
&= \int_t^T (1-\rho(X_u))V_u du + \sum_{N_t<i\leq N_T} Var(\gamma_i) \qquad \text{(E.7)}
\end{aligned}
$$

Place (E.5) and (E.6) in (E.4) and then combining (E.4) with (E.3) will give the numerator of d_1. Let $Var_t\left[\log(\frac{S_T}{S_t})\right]$ be $\overline{\sigma_{t,T}}$, which is the denominator of d_1. Thus, we obtain

$$
d_1 = \frac{1}{\overline{\sigma_{t,T}}}\left[\log\left(\frac{S_t\xi_{t,T}}{KB^*_{t,T}}\right) + \frac{1}{2}(\overline{\sigma_{t,T}})^2\right].
$$

E.2 The second part of the formula: H_t/S_t

By applying Lemma E.0.1 to (E.2),

$$
\begin{aligned}
E_t\left[\frac{\pi_T}{\pi_t}I_{[\frac{S_T}{S_t}\geq\frac{K}{S_t}]}\right] &= E_t\left[\exp(\log\frac{\pi_T}{\pi_t})I_{[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\geq 0]}\right] \\
&= E_t\left[\frac{\pi_T}{\pi_t}\right]\Phi\left[\frac{E_t\left[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\right] + Cov_t\left[\log(\frac{\pi_T}{\pi_t}),\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\right]}{\sqrt{Var_t\left[\log(\frac{S_T}{S_t})-\log(\frac{K}{S_t})\right]}}\right] \\
&= B^*_{t,T}\Phi(d_2).
\end{aligned}
$$

To simplify d_2 and express it as $B^*_{t,T}$ and $\xi_{t,T}$, the result of the calculation of d_1 can be used.

$$
\begin{aligned}
d_2 &= \frac{1}{\overline{\sigma_{t,T}}}\left[\log\left(\frac{S_t\xi_{t,T}}{KB^*_{t,T}}\right) - \frac{1}{2}(\overline{\sigma_{t,T}})^2\right] \\
&= d_1 - \overline{\sigma_{t,T}}
\end{aligned}
$$

APPENDIX F

Moments of integrated volatility

F.1 Proof of Proposition 1.4.1

The conditional mean is

$$
\begin{aligned}
E\left[\int_t^T V_u du | F_t\right] &= \int_t^T E(V_u|F_t)du \\
&= \int_t^T \left[V_t e^{-k_v(u-t)} + \theta\left(1-e^{-k_v(u-t)}\right)\right] du \\
&= V_t \frac{1}{k_v}\left(1-e^{-k_v(T-t)}\right) + \theta(T-t) - \frac{\theta}{k_v}\left(1-e^{-k_v(T-t)}\right) \\
&= V_t A_{t,T} + B_{t,T}, \qquad \text{(F.1)}
\end{aligned}
$$

where

$$
\begin{aligned}
A_{t,T} &= \frac{1}{k_v}\left(1-e^{-k_v(T-t)}\right) \\
B_{t,T} &= \theta(T-t) - \frac{\theta}{k_v}\left(1-e^{-k_v(T-t)}\right).
\end{aligned}
$$

By Ito's formula,

$$
dE\left[\int_t^T V_u du | F_t\right] = -V_t dt + A_{t,T}\sigma_v \sqrt{V_t} dW_{1t}.
$$

Since

$$
E\left[\int_T^T V_u du | F_T\right] = E\left[\int_t^T V_u du | F_t\right] + \int_t^T (-V_u)du + \int_t^T A_{u,T}\sigma_v\sqrt{V_u}dW_{1u}
$$

and

$$
E\left[\int_T^T V_u du | F_T\right] = 0,
$$

we get

$$\int_t^T V_u du - E\left[\int_t^T V_u du|F_t\right] = \int_t^T A_{u,T}\sigma_v\sqrt{V_u}dW_{1u}$$

Thus, the conditional variance is obtained as

$$\begin{aligned}
Var\left[\int_t^T V_u du|F_t\right] &= E^2\left[\int_t^T V_u du - E\left(\int_t^T V_u du|F_t\right)|F_t\right] \\
&= E^2\left[\int_t^T A_{u,T}\sigma_v\sqrt{V_u}dW_{1u}|F_t\right] \\
&= \int_t^T A_{u,T}^2\sigma_v^2 E(V_u|F_t)du \\
&= \int_t^T A_{u,T}^2\sigma_v^2\left[V_t e^{-k_v(u-t)} + \theta\left(1 - e^{-k_v(u-t)}\right)\right]du \\
&= V_t C_{t,T} + D_{t,T}, \qquad \text{(F.2)}
\end{aligned}$$

where

$$\begin{aligned}
C_{t,T} &= \frac{\sigma_v^2}{k_v^2}\left[\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right] \\
D_{t,T} &= \frac{\sigma_v^2\theta}{k_v^2}\left[(T-t)\left(1+2e^{-k_v(T-t)}\right) + \frac{1}{2k_v}\left(e^{-k_v(T-t)}+5\right)\left(e^{-k_v(T-t)}-1\right)\right]
\end{aligned}$$

F.2 Proof of Proposition 1.4.2

Note that

$$V_T = V_t + k_v\int_t^T V_u du + \sigma_v\int_t^T \sqrt{V_u}dW_{1u}.$$

Thus,

$$\begin{aligned}
&Var[V_T|F_t] \\
&\quad = Var\left[k_v\int_t^T V_u du + \sigma_v\int_t^T \sqrt{V_u}dW_{1u}|F_t\right] \\
&\quad = k_v^2 Var\left[\int_t^T V_u du|F_t\right] + \sigma_v^2 Var\left[\int_t^T \sqrt{V_u}dW_{1u}|F_t\right] \\
&\quad -2k_v\sigma_v Cov\left[\int_t^T V_u du, \int_t^T \sqrt{V_u}dW_{1u}|F_t\right] \qquad \text{(F.3)}
\end{aligned}$$

and

$$
\begin{aligned}
&Cov\left[\int_t^T V_u du, V_T|F_t\right] \\
&\quad = Cov\left[-k_v\int_t^T V_u du + \sigma_v\int_t^T \sqrt{V_u}dW_{1u}, \int_t^T V_u du|F_t\right] \\
&\quad = E\left[-k_v\left(\int_t^T V_u du\right)^2|F_t\right] + \sigma_v E\left[\int_t^T \sqrt{V_u}dW_{1u}\int_t^T V_u|F_t\right] + k_v E^2\left[\int_t^T V_u du|F_t\right] \\
&\quad = -k_v Var\left[\int_t^T V_u du|F_t\right] + \sigma_v E\left[\int_t^T \sqrt{V_u}dW_{1u}\int_t^T V_u du|F_t\right]. \qquad \text{(F.4)}
\end{aligned}
$$

$Var\left[\int_t^T \sqrt{V_u}dW_{1u}|F_t\right]$ is equal to $E\left[\int_t^T V_u du|F_t\right]$, which is already obtained in Section 1.4.2. $Var\left[\int_t^T V_u du|F_t\right]$ is also given in Section 1.4.2. The remaining work is calculation of

$$
Cov\left[\int_t^T V_u du, \int_t^T \sqrt{V_u}dW_{1u}|F_t\right] = E\left[\int_t^T \sqrt{V_u}dW_{1u}\int_t^T V_u du|F_t\right].
$$

By integration by parts,

$$
\begin{aligned}
&E\left[\int_t^T \sqrt{V_u}dW_{1u}\int_t^T V_u du|F_t\right] \\
&\quad = E\left[\int_t^T\left(\int_t^u \sqrt{V_s}dW_{1s}\right)V_u du + \int_t^T\left(\int_t^u V_s ds\right)\sqrt{V_u}dW_{1u}|F_t\right] \\
&\quad = E\left[\int_t^T\left(\int_t^u \sqrt{V_s}dW_{1s}\right)V_u du|F_t\right] \qquad \text{(F.5)} \\
&\quad = E\left[\int_t^T\left(\int_t^u \sqrt{V_s}dW_{1s}\right)\left(V_t + \int_t^u k_v(\theta - V_s)ds + \int_t^u \sigma_v\sqrt{V_s}dW_{1s}\right)du|F_t\right] \\
&\quad = E\left[-k_v\int_t^T\left(\int_t^u \sqrt{V_s}dW_{1s}\int_t^s V_s ds\right)du|F_t\right] + E\left[\sigma_v\int_t^T\left(\int_t^u V_s ds\right)du|F_t\right] \text{(F.6)}
\end{aligned}
$$

By (F.5) and (F.6),

$$
\begin{aligned}
&\int_t^T E\left[\int_t^u \sqrt{V_s}dW_{1s}V_u|F_t\right]du \\
&\quad = -k_v\int_t^T\left[\int_t^u E\left(\int_t^s \sqrt{V_\tau}dW_{1\tau}V_s\right)ds\right]du + \sigma_v\int_t^T E\left(V_s|F_t\right)du.
\end{aligned}
$$

By the result in Section 1.4.2,

$$\int_t^T E\left[\int_t^u \sqrt{V_s}dW_{1s}V_u|F_t\right]du$$
$$= -k_v \int_t^T \left[\int_t^u E\left(\int_t^s \sqrt{V_\tau}dW_{1\tau}V_s|F_t\right)ds\right]du$$
$$\sigma_v \int_t^T V_t\left[\frac{1}{k_v}\left(1-e^{-k_v(u-t)}\right)+\theta(u-t)-\frac{\theta}{k_v}\left(1-e^{-k_v(u-t)}\right)\right]du.$$

It gives the first order ordinary differential equation:

$$\frac{dE\left[\int_t^y \sqrt{V_s}dW_{1s}V_u|F_t\right]}{du} = -k_v E\left[\int_t^u \sqrt{V_s}dW_{1s}V_u|F_t\right]+(V_t-\theta)e^{-k_v(u-t)}.$$

Solving this equation yields:

$$E\left[\int_t^u \sqrt{V_s}dW_{1s}V_u\right] = -e^{-k_v(u-t)}E\left[\int_t^t \sqrt{V_s}dW_sV_u\right]+(V_t-\theta)ue^{-k_v(u-t)}, \quad \text{(F.7)}$$

since (F.7) gives

$$\frac{dE\left[\int_t^y \sqrt{V_s}dW_{1s}V_u|F_t\right]}{du}$$
$$= -k_ve^{-k_v(u-t)}E\left[\int_t^t \sqrt{V_s}dW_{1s}V_u\right]+(V_t-\theta)e^{-k_v(u-t)}-k_v(V_t-\theta)ue^{-k_v(u-t)}$$
$$= -k_v\left[e^{-k_v(u-t)}E\left[\int_t^t \sqrt{V_s}dW_{1s}V_u|F_t\right]+(V_t-\theta)ue^{-k_v(u-t)}\right]+(V_t-\theta)e^{-k_v(u-t)}$$
$$= -k_vE\left[\int_t^u \sqrt{V_s}dW_{1s}V_u|F_t\right]+(V_t-\theta)e^{-kv(u-t)}.$$

Since the first term on the right hand side of (F.7) equals to zero, completing the outside integration operator now yields,

$$Cov\left[\int_t^T V_udu, \int_t^T \sqrt{V_u}dW_{1u}|F_t\right]$$
$$= E\left[\int_t^T\left(\int_t^u \sqrt{V_s}dW_{1s}V_u\right)du|F_t\right]$$
$$= (V_t-\theta)\int_t^T ue^{-k_v(u-t)}du$$

$$= (V_t - \theta)\left[-\frac{T}{k_v}e^{-k_v(T-t)} + \frac{t}{k_v} + \frac{1}{k_v^2}\left(1 - e^{-k_v(T-t)}\right)\right]. \tag{F.8}$$

By plugging (F.1), (F.2) and (F.8) into (F.3), we obtain

$$Var[V_T|F_t] \quad = \quad V_t I_{t,T} + J_{t,T},$$

where

$$\begin{aligned}
I_{t,T} &= \sigma_v^2\left[\left(1+\frac{\sigma_v^2}{k_v}\right)\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right. \\
&\quad \left. +2\left(Te^{-k_v(T-t)} - t + \frac{1}{k_v}\left(e^{-k_v(T-t)} - 1\right)\right)\right] \\
J_{t,T} &= \sigma_v^2\left[\left(1+\frac{\sigma_v^2}{k_v^2}\right)\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right. \\
&\quad \left. +2\theta\left(Te^{-k_v(T-t)} - t + e^{-k_v(T-t)} - 1\right)\right]
\end{aligned}$$

By plugging (F.2) and (F.8) into (F.3) into (F.4), we obtain

$$Cov\left[\int_t^T V_u du, V_T|F_t\right] \quad = \quad V_t Q_{t,T} + R_{t,T},$$

where

$$\begin{aligned}
Q_{t,T} &= \frac{\sigma_v}{k_v}\left[t+1-(T+1)e^{-k_v(T-t)} - \sigma_v\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right] \\
R_{t,T} &= \frac{\sigma_v\theta}{k_v}\left[t+1-(T+1)e^{-k_v(T-t)} - \sigma_v\left(\frac{1}{k_v} - 2(T-t)e^{-k_v(T-t)} - \frac{1}{k_v}e^{-2k_v(T-t)}\right)\right]
\end{aligned}$$

APPENDIX G

Nonparametric estimation

Regression methods model the expected behavior of a dependent variable given a vector of regressors. In our study, the above statement can be formulated as

$$S_i^* = g(x_i, \tau_i, \rho_i) + u_i,$$

where x_i is moneyness, τ_i is time to maturity and ρ_i is the leverage effect. i indicates the i^{th} sample. Under some condition, $E(S^*|x,\tau,\rho)$ is the optimal predictor of S^* given x_i,τ_i and ρ_i. To estimate this conditional expectation, we employ the statistical technique known as nonparametric kernel estimation. Nonparametric kernel regression produces an estimator of the conditional expectation without requiring that $g(\cdot)$ be parameterized by the finite number of parameters. Kernel regression requires few assumptions other than smoothness of the function to be estimated and it is robust to the potential misspecification of any given parametric form of $g(\cdot)$. The choice of the regression type, the kernel type, the bandwidth selection method all can affect the result of the estimation. The type of the regression that is often used is the local constant kernel estimation. It is also known as the Nadaraya-Watson kernel estimation. Denote all explanatory variables as Z. The estimate for the conditional expectation, $\widehat{g}(Z)$ is given as

$$\begin{aligned} \widehat{g}(Z) &= \widehat{E}[S^*|Z] \\ &= \frac{\sum_{i=1}^{n} S_i^* K\left(\frac{Z_i - Z}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{Z_i - Z}{h}\right)}, \end{aligned}$$

where $K\left(\frac{Z_i-Z}{h}\right) = k\left(\frac{x_i-x}{h_1}\right) k\left(\frac{\tau_i-\tau}{h_2}\right) k\left(\frac{\rho_i-\rho}{h_3}\right)$ and k is a kernel function. Intuitively the estimate is given by a weighted average of the observed modified stock prices, S_i^* with more weight given to the stock prices whose characteristics Z_i are closer to the characteristics Z of the stock price to be estimated. Thus, the closer h is to zero, the more peaked is the function around Z_i and thus the greater is the weight given to realizations of the random variable, Z_i.

In our study, we also try the local linear kernel estimation. Three different types of kernels are used:

$$\begin{aligned}\text{the gaussian kernel}: k(\nu) &= \tfrac{1}{2\pi}\exp\left(-\tfrac{1}{2}\nu^2\right)\\ \text{the uniform kernel}: k(\nu) &= 1/2I(|\nu| \leq 1)\\ \text{the epanechnikov kernel}: k(\nu) &= \tfrac{3}{4}(1-\nu^2)I(|\nu| \leq 1),\end{aligned}$$

where I is the indicator function. For the bandwidth selection, the least squares cross-validation is used. Only the result from the gaussian kernel is reported. Refer to Li and Racine (2007) for details.

BIBLIOGRAPHY

Alfonsi, A. (2005). On the discretization schemes for the cir and bessel squared processes. *Monte Carlo Methods and Applications* 11, 355–384.

Andersen, T.G., Benzoni, L., and Lund, J. (1997). An empirical investigation of continuous-time equity return models. *The Journal of Finance* LVII, 1239–1284.

Andersen, T.G. and Bollerslev, T. (1998). Intraday periodicity and volatility persistance in financial markets. *Journal of Empirical Finance* 4, 115–158.

Bakshi, G., Cao, C., and Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003–2049.

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

Bates, D. (2000). Post 87 crash fears in s&p 500 futures options. *Journal of Econometrics* 94, 181–238.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–659.

Brigo, D. and Alfonsi, A. (2005). Credit default swap calibration and derivatives pricing with the ssrd stochastic intensity model. *Finance and Stochastics* 9, 29–42.

Broadie, M. and Kaya, Ö.. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54, 11–32.

Cheng, A., Gallant, R., Ji, C., and Lee, B. (2007). A gaussian approximation scheme for computation of option prices in stochastic volatility models. Working paper, UNC-chapel Hill.

Chernov, M. (2007). On the rold of risk premia in volatility forecasting. *Journal of Business and Economic Statistics* 25, 411–426.

Chib, S., Nardari, F., and Shephard, N. (2002). Markov chain monte carlo methods for stochastic volatility models. *The Journal of Econometrics* 108, 281–316.

Christensen, B.J. and Prabhala, N.R. (1998). The relation between implied and realized volatility. *The Journal of Financial Economics* 50, 125–150.

Christoffersen, P. and Jacobs, K. (2004). The importance of the loss function in option valuation. *Journal of Financial Economics* 63, 3–50.

Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985). A theroy of the term structure of interest rates. *Econometrica* 53, 385–407.

Eraker, B. (2004). Do stock prices and volatility jumps? reconciling evidence from spot and option prices. *Journal of Finance* LIX, 1367–1403.

Eraker, B., Johannes, M.S., and Polson, N.G. (2003). The impact of jumps in returns and volatility. *The Journal of Finance* 53, 1269–1300.

Feller, W. (1971). An introduction to probability theory and its applications, volume ii, second edition, John Wiley and Sons.

Garcia, R., Ghysels, E., and Renault, E. (2007). The econometrics of option pricing. *Handbook of Financial Econometrics*, edited by Yacine Ait-Sahalia and Lars Hansen.

Garcia, R., Lewis, M.A., and Renault, E. (2001). Estimation of objective and risk-neutral distributions based on moments of integrated volatility. Working paper, University of Montreal.

Garcia, R., Luger, R., and Renault, E. (2003). Empirical assessment of an intertemporal option pricing model with latent variables. *Journal of Econometrics* 116, 49–83.

Garcia, R., Luger, R., and Renault, E. (2004). Pricing and hedging options with implied asset prices and volatilities. Working paper, CIRANO.

Gelman, A. and Rubin, D.B. (1992). Interence from iterative simulation using multiple sequences. *Statistical Science* 7, 457–511.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior mometns. *Bayesian Statistics* Edited by Bernado, Berger, Dawid, and Smith.

Glasserman, P. (2003). Monte carlo methods in financial engineering, Springer.

Hansen, L. and Richard, S. (1987). The role of conditioning information in decuding testable restrictions implied by dynamic asset pricing. *Econometrica* 55, 587–613.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6, 327–343.

Jacquier, E., Polson, N., and Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *The Journal of Business and Economic Statistics* 12, 371–417.

Jacquier, E., Polson, N., and Rossi, P. (2004). Bayesian analysis fat-tailed stochastic volatility models with correlated errors. *The Journal of Business and Economic Statistics* 122, 185–212.

Johannes, M. and Polson, N. (2006). Mcmc methods for continuous-time financial econometrics. *Handbook of Financial Econometrics*, edited by Yacine Ait-Sahalia and Lars Hansen.

Kahl, C. and Jackel, P. (2005). Fast strong approximation monte carlo schmes for stochastic volatility models. Working paper, University of Wuppertal.

Kloeden, P. and Platen, E. (1999). Numerical solution of stochastic differential equations, Springer Verlag.

Kotz, S., Balakrishnan, N., and Johnson, N. (1999). Continuous multivariate distributions, Wiley Series.

Li, Q. and Racine, J.S. (2007). Nonparametric econometrics, Princeton University Press.

Longstaff, F.A. (1995). Option pricing and the martingale restriction. *Review of Financial Studies* 8, 1091–1124.

Mincer, J. and Zarnowitz, V. (1969). The evaluation of economic forecasts. *Economic Forecasts and Expectations*, mincer, J.

Naik, V. and Lee, M. (1990). General equilibrium pricing of options on the market portfolio with discontinuous returns. *The Review of Financial Studies* 3, 493–521.

Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.

Protter, P.E. (2004). Stochastic integration and differential equations, Springer.

Renault, E. and Touzi, N. (1996). Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance* 6, 279–302.

Robert, C.P. and Casella, G. (2004). Monte carlo statistical methods, Springer Series in Statistics, Springer, New York.

Roger, L.C.G. (1995). Which model for term-structure of interest rates should one use? *Mathematical Finance* IMA Volume 65, 93–116.

Romano, M. and Touzi, N. (1997). Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance* 7, 399–410.

Shmeiser, B.W. and Lal, R. (1982). Bivariate gamma random vectors. *Operation Research* 30, 355–374.

Singleton, K. (2006). Empirical dynamic asset pricing, Princeton.

Willard, G.A. (1997). Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives* 5, 45–61.