

ESSAYS ON THE SPECIFICATION TESTING FOR DYNAMIC ASSET
PRICING MODELS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jaeho Yun

August 2009

©2009 Jaeho Yun

ESSAYS ON THE SPECIFICATION TESTING FOR DYNAMIC ASSET
PRICING MODELS

Jaeho Yun, Ph.D.

Cornell University 2009

This dissertation consists of three essays on the subjects of specification testing on dynamic asset pricing models.

In the first essay (with Yongmiao Hong), “A Simulation Test for Continuous-Time Models,” we propose a simulation method to implement Hong and Li’s (2005) transition density-based test for continuous-time models. The idea is to simulate a sequence of dynamic probability integral transforms, which is the key ingredient of Hong and Li’s (2005) test. The proposed procedure is generally applicable whether or not the transition density of a continuous-time model has a closed form and is simple and computationally inexpensive. A Monte Carlo study shows that the proposed simulation test has very similar sizes and powers to the original Hong and Li’s (2005) test. Furthermore, the performance of the simulation test is robust to the choice of the number of simulation iterations and the number of discretization steps between adjacent observations.

In the second essay (with Yongmiao Hong), “A Specification Test for Stock Return Models,” we propose a simulation-based specification testing method applicable to stochastic volatility models, based on Hong and Li (2005) and Johannes et al. (2008). We approximate a dynamic probability integral transform in Hong and Li’s (2005) density forecasting test, via the particle filters proposed by Johannes et al.

(2008). With the proposed testing method, we conduct a comprehensive empirical study on some popular stock return models, such as the GARCH and stochastic volatility models, using the S&P 500 index returns. Our empirical analysis shows that all models are misspecified in terms of density forecast. Among models considered, however, the stochastic volatility models perform relatively well in both in- and out-of-sample. We also find that modeling the leverage effect provides a substantial improvement in the log stochastic volatility models. Our value-at-risk performance analysis results also support stochastic volatility models rather than GARCH models.

In the third essay (with Yongmiao Hong), “Option Pricing and Density Forecast Performances of the Affine Jump Diffusion Models: the Role of Time-Varying Jump Risk Premia,” we investigate out-of-sample option pricing and density forecast performances for the affine jump diffusion (AJD) models, using the S&P 500 stock index and the associated option contracts. In particular, we examine the role of time-varying jump risk premia in the AJD specifications. For comparison purposes, nonlinear asymmetric GARCH models are also considered. To evaluate density forecasting performances, we extend Hong and Li’s (2005) specification testing method to be applicable to the famous AJD class of models, whether or not model-implied spot volatilities are available. For either case, we develop (i) the Fourier inversion of the closed-form conditional characteristic function and (ii) the Monte Carlo integration based on the particle filters proposed by Johannes et al. (2008). Our empirical analysis shows strong evidence in favor of time-varying jump risk premia in pricing cross-sectional options over time. However, for density forecasting performances, we could not find an AJD specification that successfully reconcile the dynamics implied by both time-series and options data.

BIOGRAPHICAL SKETCH

Jaeho Yun was born on January 11th, 1971 in Seoul, Republic of Korea. He received his B.A. and M.A. degrees in Economics from Seoul National University in 1994 and 1996, respectively. He joined the Bank of Korea, and served in the Research Department and the Financial System Stability Department. He joined the Doctoral program in Economics at Cornell University in 2004. In 2008, he was awarded a Master of Arts degree, and in August of 2009 he earned the Doctor of Philosophy Degree from Cornell University. Jaeho Yun will resume his career as a senior economist at the Bank of Korea.

To My Parents.

ACKNOWLEDGEMENTS

I would like to thank my chairperson Professor Yongmiao Hong for his support, encouragement and invaluable guidance throughout my graduate career. I have furthermore to thank my committee members including Professors Nicholas Kiefer and Morten Nielsen for their consistent support and stimulating suggestions regarding this thesis. I was really fortunate to have an opportunity to receive excellent training for econometrics from such distinguished scholars.

I am indebted to Professors Shin-haing Kim, In June Kim, and Chon Pyo Lee in Seoul National University. Thanks to them, I could learn economics for the first time of my life, and decide to study it further. I am also grateful to Professors Jennifer Wissink, Francesca Molinari, and Nigar Hashimzade, who all have made my life at Cornell more memorable. In particular, I extend my special gratitude to Professor Tapan Mitra for having given me an opportunity to TA in the department. In fact, the day when I received an e-mail offering a TA scholarship was one of the happiest days at Cornell.

I would like to give my special thanks to the Bank of Korea for the financial support that enabled me to complete this work. I am also grateful to my friends in both Seoul and Ithaca for their encouragement throughout the difficult times. I hope that they will forgive me for not mentioning them individually.

Lastly and most importantly, I wish to thank my parents, Ikhan Yun and Yangja Chi, and my sisters, Jaeyeon and Jaehee for their love and support.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
List of Figures	ix
List of Tables	x
Chapter 1: Introduction	1
References	6
Chapter 2: A Simulation Test for Continuous-Time Models	8
2.1 Introduction	8
2.2 Nonparametric Specification Test	10
2.3 Simulation-based Nonparametric Specification Test	13
2.4 Finite Sample Performances	16
2.4.1 Size	17
2.4.2 Power	20
2.5 Conclusion	23
References	24
Chapter 3: A Specification Test for Stock Return Models	27
3.1 Introduction	27
3.2 The Specification Testing Method	33

3.2.1	The Review of Existing Testing Methods	33
3.2.2	Specification Testing with the Particle Filtering Method . . .	36
3.3	Models	43
3.3.1	Black-Scholes and Pure Jump Models	43
3.3.2	GARCH Models	45
3.3.3	Log Stochastic Volatility Models	46
3.3.4	Square-Root Stochastic Volatility Models	47
3.4	Model Estimation	50
3.4.1	Data	50
3.4.2	Estimation Method	51
3.4.3	Estimation Result	53
3.5	Model Performance Evaluations	62
3.5.1	In-Sample Performance	62
3.5.2	Out-of-Sample Performance	66
3.5.3	Value-at-risk Performance	69
3.6	Conclusion	73
	Appendices	75
	References	79

**Chapter 4: Option Pricing and Density Forecast Performances of the
Affine Jump Diffusion Models: the Role of Time-Varing Jump
Risk Premia** **87**

4.1	Introduction	87
4.2	Models	93
4.2.1	Affine Jump Diffusion Class of Models	93
4.2.2	Nonlinear Asymmetric GARCH	98

4.3	Methodology	100
4.3.1	Dynamic Probability Integral Transform via Inverse Fourier Transform	100
4.3.2	Dynamic Probability Integral Transform via Particle Filtering	103
4.3.3	Option Pricing for the AJD Class of Models	105
4.4	Data and Model Parameter Estimation	107
4.4.1	Data	107
4.4.2	Objective Parameter Estimation	110
4.4.3	Risk-neutral Parameter Estimation	113
4.5	Empirical Results	118
4.5.1	Option Pricing Performance Evaluation	118
4.5.2	Density Forecast Performance Evaluation	129
4.6	Conclusion	135
	Appendices	138
	References	141
	Chapter 5: Conclusion	150
	References	152

LIST OF FIGURES

2.1	The finite sample size performance of the original version $Q(j)$ and simulated version $Q(j)$ statistics for high level of persistent dependence	18
2.2	The finite sample size performance of the original version $Q(j)$ and simulated version $Q(j)$ statistics for low level of persistent dependence	19
2.3	The finite sample power performance of the original version $Q(j)$ and simulated version $Q(j)$ statistics for univariate diffusions	22
3.1	Filtered Model-implied Spot Volatilities	60
3.2	Filtered Jump Probability and Size of Jump-in-return for the SVJ and SVCJ	61
3.3	$Q(j)$ Statistics and Kernel Marginal Densities for In-sample	63
3.4	$Q(j)$ Statistics and Kernel Marginal Densities for Out-of-sample	67
4.1	Historical Volatilities (VIX index)	108
4.2	Model-implied BSIV's for a High-volatility Day (11/28/2001)	125
4.3	Model-implied BSIV's for a Low-volatility Day (3/9/2005)	126
4.4	$Q(j)$ Statistics and Kernel Marginal Densities	132

LIST OF TABLES

3.1	Summary Statistics for Stock Returns	50
3.2	Parameter Estimates for BS and PJ Models	53
3.3	Parameter Estimates for GARCH Models	54
3.4	Parameter Estimates for Log Stochastic Volatility Models	55
3.5	Parameter Estimates for Square Root Stochastic Volatility Models	56
3.6	In- and Out-of-sample Performance Evaluation Statistics $W(p)$	62
3.7	The Kupiec test statistics	71
3.8	The Dynamic Quantile (DQ) test statistics	72
4.1	Parameter Estimates for the NGARCH Models	111
4.2	Objective Parameter Estimates for the AJD Class of Models	112
4.3	Risk-neutral Parameter Estimates for the AJD Models	117
4.4	Option Pricing Errors	121
4.5	Option Pricing Errors by Moneyness-Maturity Categories	122
4.6	$W(p)$ Statistics	131

Chapter 1

Introduction

This dissertation consists of three essays (Chapter 2, 3, and 4, respectively) on the subjects of specification testing on dynamic asset pricing models. Dynamic asset pricing models (discrete- or continuous-time models) are widely used to capture the time-series dynamics of asset prices, such as interest rates, stock prices and foreign exchange rates. Among discrete- and continuous-time models, the rich theories for continuous-time processes often allow one to obtain analytical results that would be unavailable in discrete time (e.g., Cochrane, 2005). For example, the continuous-time affine jump diffusion models provide convenient closed-form option pricing solutions while discrete-time GARCH models generally don't. However, in the past, compared to discrete-time counterparts, it was a challenging task to estimate continuous-time models because continuous-time models should be estimated from a discretely sampled data set, despite their continuous-time nature, and those models often don't provide a closed-form conditional density function needed to apply the maximum likelihood method. Nowadays, with the development of econometric technique, the estimation procedure becomes not as challenging as before. For instance, a lot of simulation-based estimation methods have been proposed in the literature, such as the EMM (efficient method of moment) and the MCMC (Markov chain Monte Carlo), to name a few.¹

In contrast to the vast literature on the estimation of dynamic asset pricing models, there has been relatively little effort on specification analysis for those models. Model misspecification generally yields inconsistent model parameter estimators, which could lead to misleading conclusions on inference and hypothesis

¹For more details on estimation methods, refer to Section 3.4.2.

testing. In this context, Broadie et al. (2007) argue that forcing a misspecified model to fit observed prices is dangerous if the fitted parameters are used to price or hedge other derivatives. Therefore, it is important to develop reliable specification tests for such models.

There has been an increasing interest in testing for dynamic asset pricing models. However, most existing tests are based on the stationary (i.e., marginal) density of the underlying process as described in Chapter 2 and 3. Because the stationary density cannot capture the full dynamics of the underlying process, related tests might lack power against misspecified models which have the same stationary density as the data generating process.

To overcome this drawback, Hong and Li (2005) propose a specification test using the transition density (i.e., the conditional density given the past history), which can capture the full dynamics of the underlying process. Using their testing method, one can evaluate a model performance in terms of density forecast. The basic idea is to construct a sequence of dynamic probability integral transforms (also called the generalized residuals), which are i.i.d. $U[0,1]$ under correct model specification. Hong and Li then test the i.i.d. $U[0,1]$ property by using a nonparametric density estimator. Because of the i.i.d. property of the dynamic probability integral transforms, the finite sample performance of Hong and Li's (2005) is robust to persistent dependence in data, which is not enjoyed by the existing tests. Also, this testing method can be conveniently used to compare relative density forecasting performances of non-nested models since this method compares the relative performance of non-nested models in a unified way by a metric measuring the departure from i.i.d. $U[0,1]$. Despite the various merits of Hong and Li's (2005) method, it is challenging to directly apply the method to some popular models,

such as stochastic volatility models where some latent state variables are involved.

Our dissertation makes methodological contributions to the literature by extending Hong and Li's (2005) method to be applicable to various asset pricing models. In Chapter 2, we propose a simulation method to implement Hong and Li's (2005) transition density-based test for continuous-time models. The idea is to simulate a sequence of dynamic probability integral transforms, which is the key ingredient of Hong and Li's (2005) test. The proposed procedure is generally applicable whether or not the transition density of a continuous-time model has a closed form and is simple and computationally inexpensive. A Monte Carlo study for the size and power performances of the test is also provided in Chapter 2.

Based on the finding in Chapter 2 that a simulation method works well in applying Hong and Li's (2005) testing method, we develop a simulation-based specification testing method applicable to stochastic volatility models in Chapter 3. The presence of latent state variables in the stochastic volatility models makes the dynamic probability integral transform analytically intractable. We circumvent this analytic intractability by employing the particle filters proposed by Johannes et al. (2008).

In Chapter 4, we develop a testing method applicable to the famous affine jump diffusion (AJD) class of models (i.e., square-root stochastic volatility models with various jump specifications). Under the premise that option market is fully integrated with the corresponding spot market, one can use model-implied spot volatilities extracted from option data to forecast densities. Although there is still no closed-form conditional density, the AJD class provides a closed-form conditional characteristic function in many cases. Exploiting this property, we reduce the dynamic probability integral transform procedure to the Fourier inversion of

a known conditional characteristic function. Our method does not incur any discretization bias, enabling one to forecast densities at any forecast time horizon without additional computational cost.

With the proposed testing method, we conduct a comprehensive empirical analysis on the stock return models using S&P500 stock index. Some popular discrete-time stock return models (e.g., GARCH and log stochastic volatility models) are also considered.

In Chapter 3, from both in- and out-of-sample perspective, we analyze density forecasting performances for various models by using time-series data. Our models include various stochastic volatility models as well as famous GARCH models (e.g., GARCH-N, GARCH-T, EGARCH, and GJR). Notably, there has been little comparative analysis on both GARCH and stochastic volatility models. This might be partly because there are only a few specification testing tool which can evaluate both classes on a fair basis. Beside the statistical criterion, we also evaluate each model from an economic criterion. We conduct the VaR (value-at-risk) performance test for each model, which is expected to provide an important practical implication for portfolio allocation or risk assessment. The existing VaR testing methods such as the Kupiec (1995) test and the dynamic quantile test (Engle and Manganelli, 1999) are employed.

In Chapter 4, we extend our analysis to the risk-neutral dynamics implied by the affine jump diffusion (AJD) models in the out-of-sample context. We examine out-of-sample option pricing and density forecast performances for stock return models. The S&P 500 stock index and the associated option contracts are used. We consider the stock return models which provide convenient option pricing methods, such as the AJD models (Heston, 1993; Bates, 2000; Duffie et al., 2000) and

nonlinear asymmetric GARCH models (Duan, 1995). Compared to other studies, we consider more AJD specifications including both constant and time-varying jump risk premia specifications. Particularly, we focus on the role of time-varying jump risk premia in reconciling both options and historical returns data, which is one of the important issues in the time-series consistency of the AJD models. In other words, we assess whether model-implied spot volatilities (filtered from options) for each AJD model are consistent with true volatility level implied by time-series dynamics. The aforementioned Hong and Li's (2005) testing method, extended by our proposed method, will be used to compare model performances.

In sum, the rest of this dissertation is organized as follows. In Chapter 2, we propose a simulation method to implement Hong and Li's (2005) transition density-based test for continuous-time models. In Chapter 3, we introduce simulation-based specification testing method applicable to stochastic volatility models, based on Hong and Li (2005) and Johannes et al. (2008). Also, with the proposed testing method, we conduct a comprehensive empirical time-series study on some popular stock return models, such as the GARCH and stochastic volatility models. In Chapter 4, we investigate out-of-sample option pricing and density forecast performances for the affine jump diffusion (AJD) models. The Fourier inversion method of dynamic probability integral transform for the AJD models is introduced. Finally, Chapter 5 concludes.

REFERENCES

- [1] Bates, D., 2000, “Post-’87 Crash Fears in the S&P500 Futures Option Market,” *Journal of Econometrics*, 94, 181-238.
- [2] Broadie, M., M. Chernov, and M. Johannes, 2007, “Model Specification and Risk Premia: The Evidence from Futures Options,” *Journal of Finance*, 62, 1453-1490.
- [3] Cochrane, J., 2005, *Asset Pricing*, Princeton University Press, Princeton, NJ.
- [4] Duan, J., 1995, “The GARCH Option Pricing Model,” *Mathematical Finance*, 5, 13-32.
- [5] Duffie, D., J. Pan, and K. Singleton, 2000, “Transform Analysis and Asset Pricing for Affine Jump-Diffusions,” *Econometrica*, 68, 1343-1376.
- [6] Engle, R., and S. Manganelli, 1999, “CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles,” Working Paper, UCSD.
- [7] Heston, S., 1993, “A Closed-Form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options,” *Review of Financial Studies*, 6, 327-344.
- [8] Hong, Y., and H. Li, 2005, “Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates,” *Review of Financial Studies*, 18, 37-84.

- [9] Johannes, M., N. Polson, and J. Stroud, 2008, "Learning about Jumps and Stochastic Volatility: Filtering Stochastic Differential Equations with Jumps," *Review of Financial Studies*, forthcoming.
- [10] Kupiec, P. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, 2, 73-84.

Chapter 2

A Simulation Test for Continuous-Time Models

2.1 Introduction

Continuous-time models are widely used in finance to capture the dynamics of economic time-series, such as interest rates, stock prices and foreign exchange rates. In contrast to the vast literature on the estimation of continuous-time models, there has been relatively little effort on specification analysis for continuous-time models. Model misspecification generally yields inconsistent model parameter estimators, which could lead to misleading conclusions on inference and hypothesis testing. Moreover, a misspecified model can yield large errors in derivatives pricing, hedging, and risk management. Therefore, it is important to develop reliable specification tests for continuous time models.

There has been an increasing interest in testing for continuous-time models. This includes Ait-Sahalia (1996), Gao and King (2004), Thompson (2002), Chen and Gao (2005, 2007), Corradi and Swanson (2005, 2006), Li and Tkacz (2006), Li (2007), and Bhardwaj, Corradi and Swanson (2007). Most existing tests are based on the stationary (i.e., marginal) density of the underlying process. The stationary density usually has a closed form and therefore related test statistics are convenient to compute. However, because the stationary density cannot capture the full dynamics of the underlying process, related tests have no power against misspecified models which have the same stationary density as the data generating process. To overcome this drawback, Hong and Li (2005) propose a specification test using the transition density (i.e., the conditional density given the past history), which can capture the full dynamics of the underlying process. The basic idea is to construct a sequence of dynamic probability integral transforms, which will be i.i.d. $U[0,1]$

under correct model specification (Diebold, Gunther and Tay, 1998). Hong and Li then test the i.i.d. $U[0,1]$ property by using a nonparametric density estimator. Because of the i.i.d. property of the probability integral transforms, the finite sample performance of Hong and Li's (2005) is robust to persistent dependence in data, which is not enjoyed by the existing tests. Furthermore, Hong and Li's (2005) is generally applicable to testing both continuous- and discrete-time models, no matter whether they are univariate or multivariate.

An important issue in implementing Hong and Li's (2005) test is the calculation of the dynamic probability integral transforms. When the transition density has a closed form, the probability integral transforms can be calculated via numerical integration of the transition density. However, most continuous-time models have no closed form solution. One has to use Ait-Sahalia's (2002a, 2002b) Hermite polynomial method or the simulation methods of Pedersen (1995), Elerian, Chib and Shephard (2001) and Brandt and Santa-Clara (2002) to approximate the transition density. These methods are computationally expensive, particularly when the sample size is sufficiently large.

In fact, to implement Hong and Li's (2005) test, one needs the dynamic probability integral transforms (i.e., integrals of the transition density) rather than the transition density itself. Based on this observation, we propose a simple and convenient simulation method to implement Hong and Li's (2005) test. The idea is to simulate dynamic probability integral transforms directly which does not require the knowledge of the closed form or the simulation of the transition density. The procedure is generally applicable to test various time-series models because neither closed form solution nor accurate approximation for the transition density is needed. The simulation test is computationally inexpensive. Our Monte Carlo

study shows that the simulation test has very similar size and power performances to Hong and Li's (2005) test using the closed form solution of the transition density (when available). Furthermore, the procedure is robust to the choices of the number of simulation iterations and the number of discretization steps between adjacent observations. We note that Pedersen (1994) first proposed a simulation method to approximate the probability integral transforms. He noted the potential usefulness of the probability integral transforms in diagnostic checking of a diffusion model, but did not propose any test procedure. Our work fills this gap.

In Section 2.2, we describe Hong and Li's (2005) test. In Section 2.3, we propose a simulation method to implement Hong and Li's (2005) test. In Section 1.4, we conduct a simulation study on its finite sample performance. Section 1.5 concludes.

2.2 Nonparametric Specification Test

Hong and Li's (2005) test is generally applicable to both continuous- and discrete-time models, whether they are univariate or multivariate. For simplicity, we consider testing a univariate continuous-time model

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t + dJ_t(X_t, \theta), \quad (2.1)$$

where $\mu(X_t, \theta)$, $\sigma(X_t, \theta)$ and $J_t(X_t, \theta)$ are the drift, diffusion and jump processes respectively, $\theta \in \Theta$ is a finite-dimensional parameter vector, $\Theta \subset \mathbb{R}^p$ is a compact parameter space, W_t is a standard Brownian motion. Throughout, we assume that the process $\{X_t\}$ is time-homogenous and stationary with an unknown transition probability density. Given the specifications on $\mu(X_t, \theta)$, $\sigma(X_t, \theta)$ and $J_t(X_t, \theta)$, the model in (1) fully characterizes a transition density $p(x|X_{t-\Delta}, \theta)$ for the process $\{X_t\}$, where $\Delta > 0$. We say that the continuous-time model in (1) is correctly specified if there exists some unknown parameter $\theta_0 \in \Theta$ such that the model-

implied transition density coincides with the true transition density of $\{X_t\}$. In this case, the continuous-time model can capture the full dynamics of $\{X_t\}$.

Hong and Li (2005) propose a test for model (1) using a discretely observed random sample $\{X_{\tau\Delta}\}_{\tau=1}^n$. For notational simplicity, we assume $\Delta = 1$. To describe this test, we consider the following dynamic probability integral transform:

$$Z_\tau(\theta) = \int_{-\infty}^{X_\tau} p(x|X_{\tau-1}, \theta) dx, \quad \tau = 1, \dots, n. \quad (2.2)$$

When the model in (1) is correctly specified in the sense that there exists some $\theta_0 \in \Theta$ such that $p(x|X_{\tau-1}, \theta_0)$ coincides with the transition density of $\{X_t\}$, then the sequence $\{Z_\tau(\theta_0)\}$ is i.i.d. $U[0, 1]$ (Diebold, Gunther and Tay 1998). The series $\{Z_\tau(\theta)\}$ can be called the "generalized residuals" of the transition density model $p(x|X_{\tau-1}, \theta)$. The i.i.d. $U[0, 1]$ property provides a basis for testing the model. If $\{Z_\tau(\theta)\}$ is not i.i.d. $U[0, 1]$ for all $\theta \in \Theta$, then the model in (1) is not correctly specified.

Hong and Li (2005) measure the distance between a model-implied transition density and the true transition density by comparing a kernel estimator $\hat{g}_j(z_1, z_2)$ for the joint density of the pair $\{Z_t(\theta_0), Z_{t-j}(\theta_0)\}$ with unity, the product of two $U[0, 1]$ densities, where j is a lag order. The kernel estimator of the joint density is, for any integer $j > 0$,

$$\hat{g}_j(z_1, z_2) = (n - j)^{-1} \sum_{\tau=j+1}^n K_h(z_1, \hat{Z}_\tau) K_h(z_2, \hat{Z}_{\tau-j}), \quad (2.3)$$

where $\hat{Z}_\tau = Z_\tau(\hat{\theta})$, $\hat{\theta}$ is any \sqrt{n} -consistent estimator for θ_0 , and $K_h(z_1, \hat{Z}_\tau)$ is a

boundary-modified kernel² defined below. For $x \in [0, 1]$, we define

$$K_h(x, y) \equiv \begin{cases} h^{-1}k\left(\frac{x-y}{h}\right) / \int_{-(x/h)}^1 k(u)du, & \text{if } x \in [0, h), \\ h^{-1}k\left(\frac{x-y}{h}\right), & \text{if } x \in [h, 1-h), \\ h^{-1}k\left(\frac{x-y}{h}\right) / \int_{-1}^{(1-x)/h} k(u)du, & \text{if } x \in [1-h, 1], \end{cases} \quad (2.4)$$

where the kernel $k(\cdot)$ is a prespecified symmetric probability density, and $h \equiv h(n)$

is a bandwidth such that $h \rightarrow 0, nh \rightarrow \infty$ as $n \rightarrow \infty$. One example of $k(\cdot)$ is the quartic kernel $k(u) = \frac{15}{16}(1-u^2)^2 1(|u| \leq 1)$, where $1(\cdot)$ is the indicator function.

We will use this kernel in our simulation study. In practice, the choice of h is more important than the choice of $k(\cdot)$. Like Scott (1992), we choose $h = \widehat{S}_Z n^{-\frac{1}{6}}$, where \widehat{S}_Z is the sample standard deviation of $\{\widehat{Z}_\tau\}_{\tau=1}^n$. This simple bandwidth rule attains the optimal rate for bivariate density estimation.

Hong and Li's (2005) test statistic is based on a properly standardized version of the quadratic form between $\widehat{g}_j(z_1, z_2)$ and 1, the product of two $U[0, 1]$ densities:

$$\widehat{Q}(j) \equiv [(n-j)h \int_0^1 \int_0^1 [\widehat{g}_j(z_1, z_2) - 1]^2 dz_1 dz_2 - hA_h^0] / V_0^{1/2}, \quad (2.5)$$

where the nonstochastic centering and scale factors

$$A_h^0 \equiv \left[(h^{-1} - 2) \int_{-1}^1 k^2(u)du + 2 \int_0^1 \int_{-1}^1 k_b^2(u)du db \right]^2 - 1, \quad (2.6)$$

$$V_0 \equiv 2 \left[\int_{-1}^1 \left[\int_{-1}^1 k(u+v)k(v)dv \right]^2 du \right]^2, \quad (2.7)$$

and $k_b(\cdot) \equiv k(\cdot) / \int_{-1}^b k(v)dv$.

Under correct model specification, $\widehat{Q}(j) \rightarrow^d N(0, 1)$ for any fixed lag order $j > 0$ as $n \rightarrow \infty$. The first lag $j = 1$ is often the most informative and important, but other lags may also reveal useful information on model misspecification. Moreover,

²The modified kernel is used because the standard kernel density estimator produces biased estimates near the boundaries of data due to asymmetric coverage of the data in the boundary regions. The denominators of $K_h(x, y)$ for $x \in [0, h) \cup (1-h, 1]$ ensure that the kernel density estimator is asymptotically unbiased uniformly over the entire support $[0, 1]$ (Hong and Li, 2005).

$cov[\widehat{Q}(i), \widehat{Q}(j)] \xrightarrow{p} 0$ for $i \neq j$ as $n \rightarrow \infty$. This implies that $\widehat{Q}(i)$ and $\widehat{Q}(j)$ are asymptotically independent whenever $i \neq j$. As a result, we can simultaneously use multiple statistics $\{\widehat{Q}(j)\}$ with different lags to examine at which lag(s) the i.i.d. $U[0, 1]$ property is violated. On the other hand, $\widehat{Q}(j) \rightarrow \infty$ in probability as $n \rightarrow \infty$ whenever $\{Z_t(\theta_0), Z_{t-j}(\theta_0)\}$ are not independent or $U[0, 1]$. This ensures that the proposed test has power against model misspecification.

Although the use of the $\widehat{Q}(j)$ statistics with different j 's reveal the information on the lag orders at which there are significant departures from i.i.d. $U[0,1]$, when comparing two different models, it is more convenient to construct a single test statistics. In this regards, Hong, Li, and Zhao (2007) suggest the following portmanteau evaluation test statistic:

$$\widehat{W}(p) = \frac{1}{\sqrt{p}} \sum_{j=1}^p \widehat{Q}(j) \quad (2.8)$$

where p is a lag truncation order. They show that the above statistic also converges to $N(0, 1)$.³

2.3 Simulation-based Nonparametric Specification Test

To implement Hong and Li's (2005) test, we need to calculate the dynamic probability integral transform or generalized residual $Z_\tau(\hat{\theta})$ in (2). When the transition density of a continuous-time model has a closed form, $Z_\tau(\hat{\theta})$ can be calculated via numerical integration. Unfortunately, the transition densities of most continuous-time models have no closed form. In such scenarios, one could use various approximation methods, such as the Hermite polynomial method of Ait-Sahalia (2002a, 2002b), or the simulation method of Pedersen (1995), Elerian, Chib and Shephard (2001), and Brandt and Santa-Clara (2002) to first approximate the transition den-

³For details of the testing method, refer to Hong and Li (2005) or Hong, Li, and Zhao (2007).

sity and then compute the generalized residuals via numerical integration. This may be computationally expensive in practice.

We now develop a simple yet generally applicable simulation method to compute the generalized residuals. The idea is to directly simulate dynamic probability integral transforms rather than the transition density. To avoid confusion, we denote the realizations of the random sample $\{X_\tau\}_{\tau=1}^n$ by $\{x_\tau\}_{\tau=1}^n$. Then a realization of the generalized residual $Z_\tau(\theta)$ is

$$\begin{aligned} z_\tau(\theta) &= \int_{-\infty}^{x_\tau} p(x \mid X_{\tau-1} = x_{\tau-1}, \theta) dx \\ &= \int_{-\infty}^{\infty} 1(x \leq x_\tau) p(x \mid x_{\tau-1}, \theta) dx \\ &= E_\theta[1(X_\tau \leq x_\tau) \mid X_{\tau-1} = x_{\tau-1}], \quad \tau = 1, \dots, n, \end{aligned} \tag{2.9}$$

where $E_\theta(\cdot|\cdot)$ denotes the conditional expectation given $X_{\tau-1} = x_{\tau-1}$ under the model-implied transition density $p(x|X_{\tau-1} = x_{\tau-1}, \theta)$.

Equation (2.12) suggests a simple approach to approximating $z_\tau(\theta)$ by Monte Carlo integration. First, conditional on the observation $X_{\tau-1} = x_{\tau-1}$ at time $\tau - 1$, we use the null continuous-time model in (1) to simulate a sample path for the process $\{X_t\}$ between time $\tau - 1$ and τ , and obtain a simulated observation \tilde{X}_τ at time τ . For this purpose, we should choose an adequate discretization scheme in order to mimic the dynamics of the continuous-time model in (1). The interval between τ and $\tau + 1$ is divided into M subintervals, whose length is $\delta = M^{-1}$. In practice, the Euler scheme and the Milstein scheme are widely used. For the Euler scheme, with a sufficiently small δ , $\tilde{X}_{\tau-1+i\delta}$ can be assumed to follow a conditional normal distribution, given the previous $\tilde{X}_{\tau-1+(i-1)\delta}$. In other words, a random

draw $\tilde{X}_{\tau-1+i\delta}$ is generated recursively under the following distribution:

$$\tilde{X}_{\tau-1+i\delta} | \tilde{X}_{\tau-1+(i-1)\delta} \sim N(\mu(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)\delta, \sigma^2(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)\delta), \quad i = 1, \dots, M, \quad (2.10)$$

with $\tilde{X}_{\tau-1} = x_{\tau-1}$.

Alternatively, we can use the Milstein scheme, which is known to provide a more accurate approximation than the Euler scheme. Suppose $\varepsilon_{\tau-1+i\delta} \sim i.i.d.N(0, 1)$ is drawn for each $i = 1, \dots, M$. Then $\tilde{X}_{\tau-1+i\delta}$ is generated recursively via the following formula:

$$\begin{aligned} \tilde{X}_{\tau-1+i\delta} &= \tilde{X}_{\tau-1+(i-1)\delta} + \mu(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)\delta + \sigma(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)\sqrt{\delta}\varepsilon_{\tau-1+i\delta} \\ &\quad + \frac{1}{2}\sigma(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)\sigma'(\tilde{X}_{\tau-1+(i-1)\delta}, \theta)(\varepsilon_{\tau-1+i\delta}^2 - \delta), \quad i = 1, \dots, M, \end{aligned} \quad (2.11)$$

where $\sigma'(\cdot, \cdot)$ is the partial derivative with respect to the first argument of the function.

For each interval from $\tau-1$ to τ , a simulation path, $\{\tilde{X}_{\tau-1+i\delta}\}_{i=1}^M$, is generated, and then an observation \tilde{X}_τ is obtained. For each given τ , we do so S times, where S is a prespecified number of simulation iterations for Monte Carlo integration. After repeating the procedure S times, we can obtain a simulated independent random sample $\{\tilde{X}_\tau^{(s)}\}_{s=1}^S$ for each $\tau = 2, \dots, n$, conditional on $X_{\tau-1} = x_{\tau-1}$. It follows that we can approximate $z_\tau(\theta)$ by the following sample average

$$\tilde{Z}_\tau(\theta, S, M) = \frac{1}{S} \sum_{s=1}^S 1 \left[\tilde{X}_\tau^{(s)} \leq x_\tau \right]. \quad (2.12)$$

When $M \rightarrow \infty$ and $S \rightarrow \infty$, the sample average $\tilde{Z}_\tau(\theta, S, M)$ converges to $z_\tau(\theta)$ by the uniform law of large numbers. In practice, because θ_0 is unknown, we have to replace it with an estimator $\hat{\theta}$. This results in a sequence of simulated generalized residuals $\{\tilde{Z}_\tau(\hat{\theta}, S, M)\}_{\tau=1}^n$ based on the estimator $\hat{\theta}$. When the estimator $\hat{\theta}$ is \sqrt{n} -

consistent for θ_0 under correct model specification, simulated generalized residual $\tilde{Z}_\tau(\hat{\theta}, S, M)$ will converge to $z_\tau(\theta_0)$ in probability as $M \rightarrow \infty$, $S \rightarrow \infty$, and $n \rightarrow \infty$.

We summarize our simulation procedure to implement Hong and Li's (2005) test:

- Estimate the continuous-time model in (1) using any method that yields a \sqrt{n} -consistent estimator $\hat{\theta}$;
- Compute the simulated generalized residuals $\{\tilde{Z}_\tau = \tilde{Z}_\tau(\hat{\theta}, S, M)\}_{\tau=1}^n$ for some prespecified choices of S and M ;
- Compute Hong and Li's (2005) test statistic $\hat{Q}(j)$ in Equation (2.5) using the simulated sequence of estimated generalized residuals $\{\tilde{Z}_\tau\}_{\tau=1}^n$. We use $\tilde{Q}(j)$ to denote the simulated version of Hong and Li's (2005) test. If $\tilde{Q}(j) > C_\alpha$, the upper tailed $N(0,1)$ critical value at significance level α , then we reject the null hypothesis that the model is correctly specified at level α .

2.4 Finite Sample Performances

We now examine the finite sample performance of the simulated version $\tilde{Q}(j)$ of Hong and Li's (2005) test via a Monte Carlo study. We are interested in how close the performances of the simulated version $\tilde{Q}(j)$ and the original version $\hat{Q}(j)$ of Hong and Li's (2005) are in terms of size and power. Moreover, since the simulated version $\tilde{Q}(j)$ involves the choice of the number (S) of simulation iterations and the number (M) of discretization steps between neighboring observations, we will examine the sensitivity of the simulated version $\tilde{Q}(j)$ to the choices of S, M . For comparison, we adopt the same simulation design as Hong and Li (2005).

2.4.1 Size

To examine the sizes of the tests $\tilde{Q}(j)$ and $\hat{Q}(j)$, we consider a Vasicek (1977) model:

$$dX_t = \kappa(\alpha - X_t)dt + \sigma dW_t \quad (2.13)$$

where α is the long run mean and κ is the speed of mean reversion. The smaller κ is, the stronger the serial dependence in $\{X_t\}$, and consequently, the slower the convergence to the long run mean. Like Hong and Li (2005), we set $(\kappa, \alpha, \sigma^2) = (0.85837, 0.089102, 0.002185)$ and $(0.214592, 0.089102, 0.000546)$ respectively. This generates low and high persistent dependence in data, respectively. It allows us to examine the robustness of the tests to persistence of dependence in data.

For each parameterization, we simulate 1000 data sets of a random sample $\{X_\tau\}_{\tau=1}^n$, with $n = 1000$. For each data set, we estimate a Vasicek model with unknown parameter $\theta = (\kappa, \alpha, \sigma^2)'$ via the maximum likelihood estimation (MLE) method. Because the Vasicek model has a Gaussian closed-form transition density, the computation of the original version $\hat{Q}(j)$ of Hong and Li's (2005) test is feasible. To compute the simulated version $\tilde{Q}(j)$, we have to choose the number (S) of simulation iterations and the number (M) of discretization steps between neighboring observations. To examine the robustness of size and power of the simulated version $\tilde{Q}(j)$, we consider various combinations from $S = 200, 500, 1000$ and $M = 1, 3, 5$. We consider the empirical rejection rates using the asymptotic critical values (1.28 and 1.65) at the 10 and 5% significance levels, respectively.

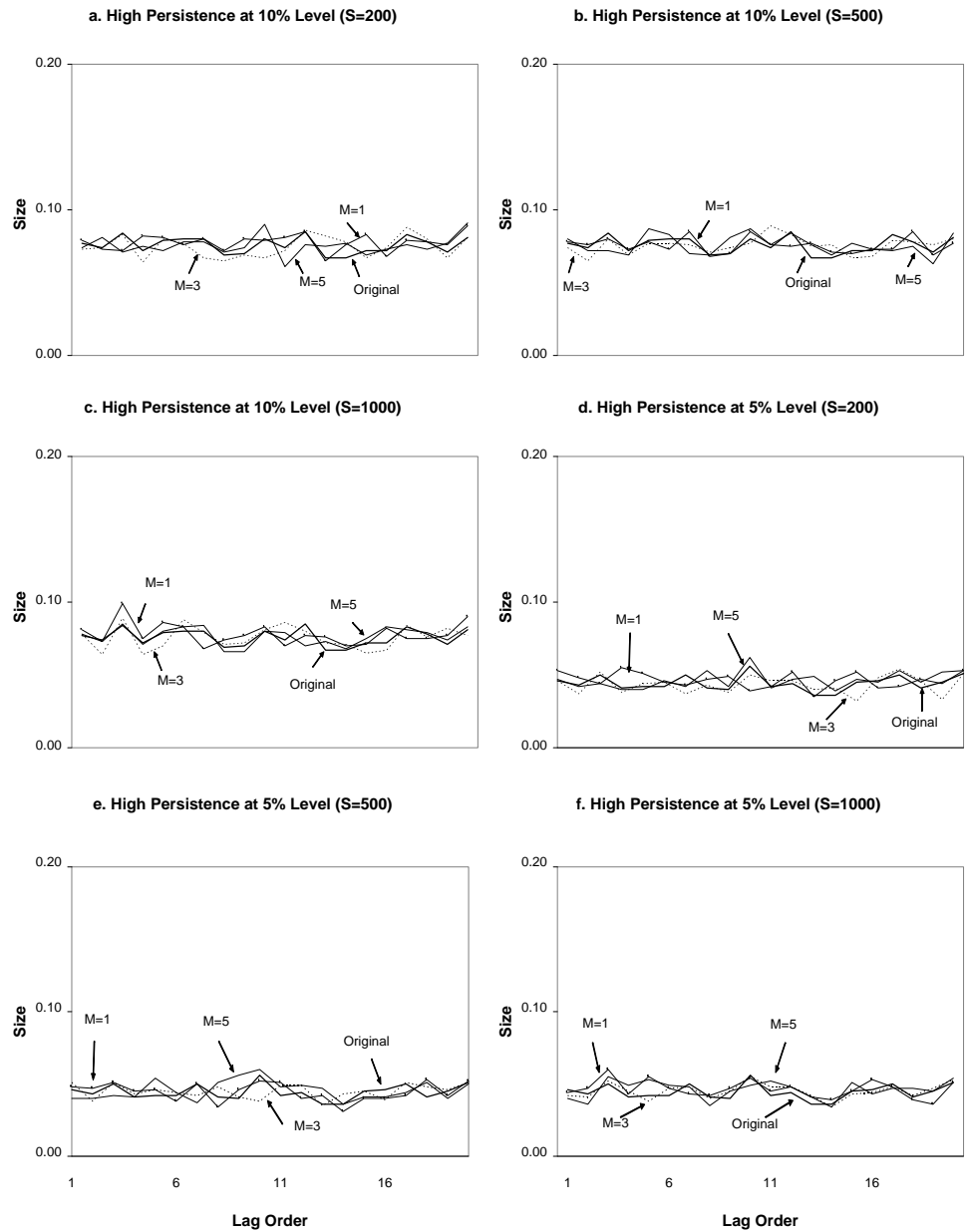


Figure 2.1: The finite sample size performance of the original version $\hat{Q}(j)$ and simulated version $\tilde{Q}(j)$ statistics for high level of persistent dependence

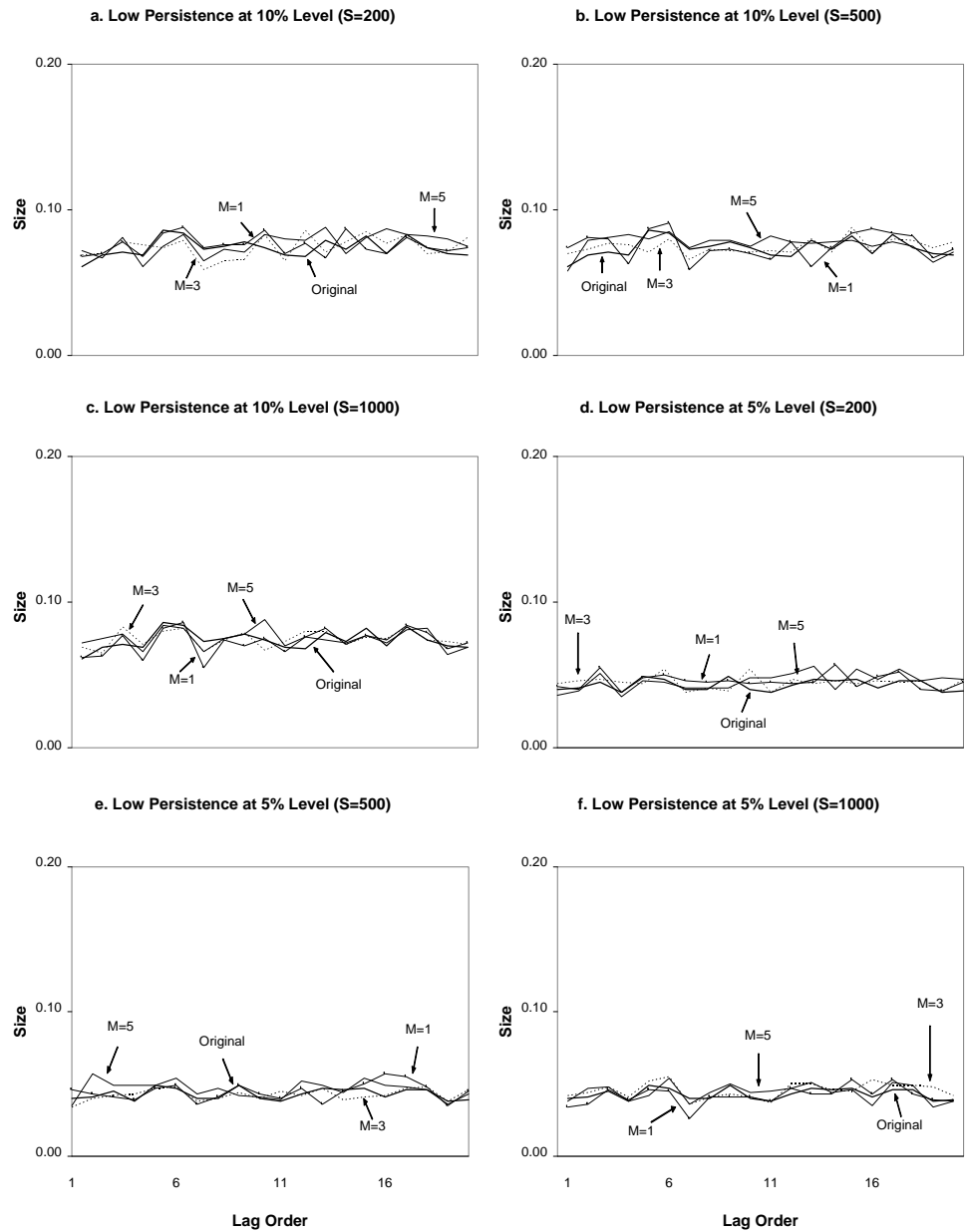


Figure 2.2: The finite sample size performance of the original version $\hat{Q}(j)$ and simulated version $\tilde{Q}(j)$ statistics for low level of persistent dependence

Figure 2.1 and 2.2 report the empirical sizes of the tests under the high and low persistent dependence cases, respectively. Each figure, given the number of simulations (S), provides the empirical sizes of the simulated version $\tilde{Q}(j)$, $j = 1, \dots, 20$, for different choices of M . Also, the sizes of the original version $\hat{Q}(j)$ denoted by "original" are provided. The original $\hat{Q}(j)$ statistics give reasonable sizes in both the high (Figure 2.1) and low (Figure 2.2) persistent dependence cases. In all the cases, the size of the simulated version $\tilde{Q}(j)$ is nearly same as that of $\hat{Q}(j)$ for all the lag orders, whatever combinations of S and M is chosen. Like $\hat{Q}(j)$, the performance of the simulation version $\tilde{Q}(j)$ is not affected by the degree of persistent dependence.

2.4.2 Power

To investigate the powers of $\tilde{Q}(j)$ and $\hat{Q}(j)$, we use four data generating processes (DGPs) considered in Hong and Li (2005):

- DGP 1. The CIR (1985) Model:

$$dX_t = \kappa(\alpha - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad (2.14)$$

where $(\kappa, \alpha, \sigma^2) = (0.89218, 0.090495, 0.032742)$.

- DGP 2. Ahn and Gao's (1999) Inverse-Feller Model:

$$dX_t = X_t[\kappa + (\sigma^2 - \kappa\alpha)X_t]dt + \sigma X_t^{3/2}dW_t, \quad (2.15)$$

where $(\kappa, \alpha, \sigma^2) = (0.181, 15.157, 0.67421)$.

- DGP 3. CKLS (1992) Model:

$$dX_t = \kappa(\alpha - X_t)dt + \sigma X_t^\rho dW_t, \quad (2.16)$$

where $(\kappa, \alpha, \sigma^2, \rho) = (0.0972, 0.0808, 0.52186, 1.46)$.

- DGP 4. Ait-Sahalia's (1996) Nonlinear Drift Model:

$$dX_t = (\alpha_{-1}X_t^{-1} + \alpha_0 + \alpha_1X_t + \alpha_2X_t^2)dt + \sigma X_t^\rho dW_t, \quad (2.17)$$

where $(\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \sigma^2, \rho) = (0.0011, -0.0517, 0.877, -4.604, 0.6475, 1.50)$.

For each of these four alternatives, we generate 500 realizations of a random sample $\{X_\tau\}_{\tau=1}^n$ with size $n = 1000$. For all DGPs 1–4, we simulate data via the Milstein scheme. To reduce the discretization bias, we simulate five observations each day and sample the data at daily frequency. For each data set, we fit a Vasicek model via MLE.

Figure 2.3 reports the powers of both the simulated version $\tilde{Q}(j)$ and original version $\hat{Q}(j)$ of Hong and Li's (2005) test, as a function of lag order j from 1 to 20, at the 5% significance level using asymptotic critical values. Under each DGP, the powers of $\tilde{Q}(j)$ and $\hat{Q}(j)$ are very close for each lag order j , and each combination of the choices of (S, M) in computing $\tilde{Q}(j)$. Even when the number of simulation iterations $S = 200$ (not reported here), the power of $\tilde{Q}(j)$ is very similar to the power of $\hat{Q}(j)$. Both tests have all-round good power against the four alternatives. There is no power loss using the simulated version $\tilde{Q}(j)$.

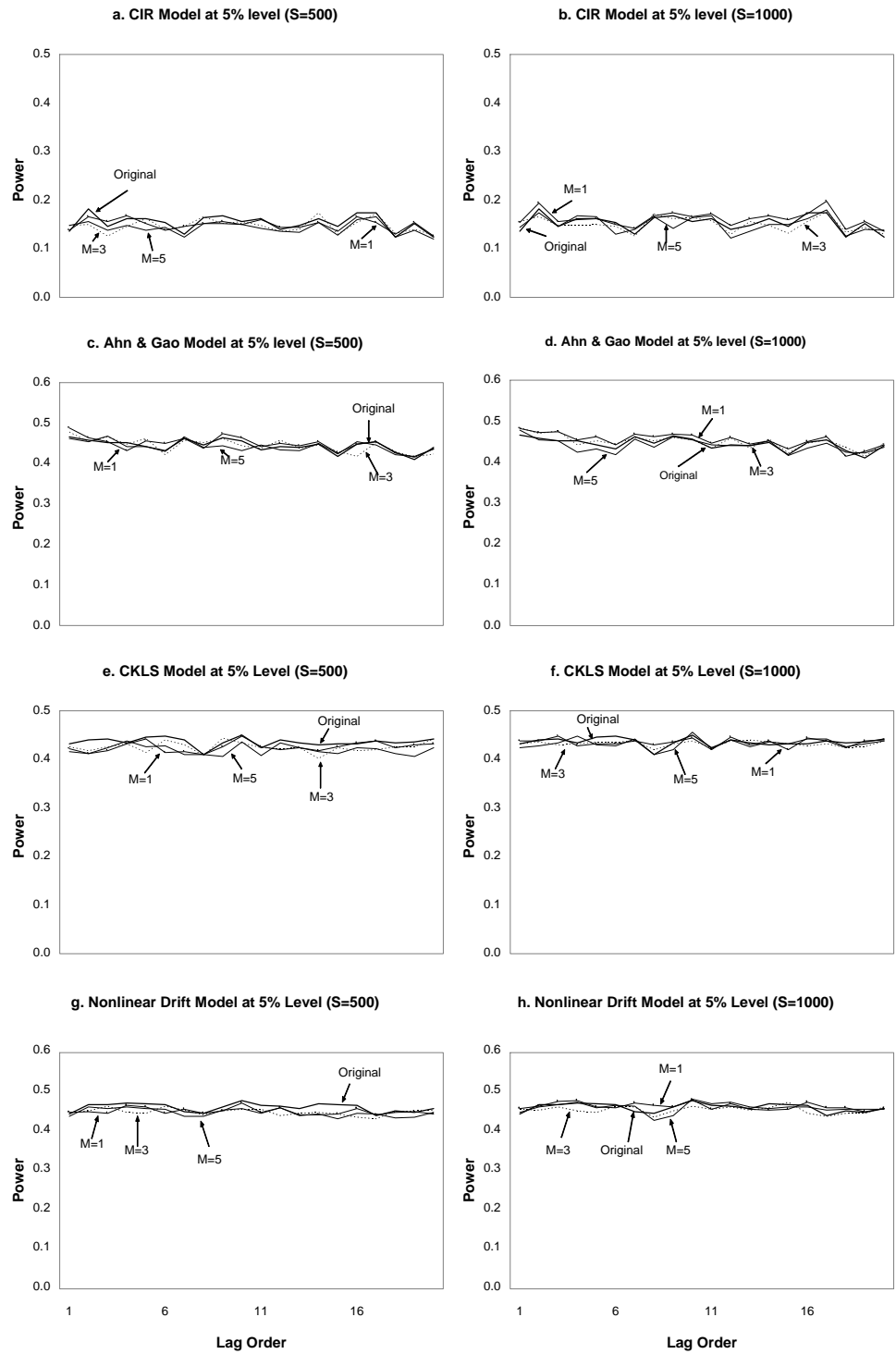


Figure 2.3: The finite sample power performance of the original version $\widehat{Q}(j)$ and simulated version $\widetilde{Q}(j)$ statistics for univariate diffusions

2.5 Conclusion

In this chapter, we propose a convenient simulation method to implement Hong and Li's (2005) test for continuous-time models using discretely sampled data. The idea is to simulate the dynamic probability integrals rather than the transition density. The former is an ingredient of Hong and Li's (2005) test. The proposed simulation test is simple and computationally inexpensive, and is generally applicable to various time-series models whether or not the transition density has a closed form. There is no need to approximate or simulate the transition density. Our Monte Carlo study shows that the proposed simulation test performs, in terms of both size and power, very similarly to the original version of Hong and Li's (2005) test using the closed form solution of the transition density (when available), and the performance of the simulation test is robust to various choices of the number of simulation iterations and the number of discretization steps between adjacent observations.

REFERENCES

- [1] Ahn, D., and B. Gao, 1999, “A Parametric Nonlinear Model of Term Structure Dynamics,” *Review of Financial Studies*, 12, 721-762.
- [2] Ait-Sahalia, Y., 1996, “Testing Continuous-Time Models of the Spot Interest Rate,” *Review of Financial Studies*, 9, 385-426.
- [3] Ait-Sahalia, Y., 2002a, “Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach,” *Econometrica*, 70, 223-262.
- [4] Ait-Sahalia, Y., 2002b, “Closed-Form Expansion for Multivariate Diffusions,” working paper, Princeton University.
- [5] Bhardwaj, G., V. Corradi, V., and N.R. Swanson, 2007, “A Simulation Based Specification Test for Diffusion Processes,” *Journal of Business and Economic Statistics*, forthcoming.
- [6] Brandt, M., and P. Santa-Clara, 2002, “Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets,” *Journal of Financial Economics*, 63, 161-210.
- [7] Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, 1992, “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,” *Journal of Finance*, 47, 1209-1227.
- [8] Chen, S.X. and J. Gao, 2005, “On the Use of the Kernel Method for Specification Tests of Diffusion Models,” Working Paper, University of Western Australia.

- [9] Chen, S.X. and J. Gao, 2007, "An Adaptive Empirical Likelihood Test for Time Series Models," *Journal of Econometrics*, forthcoming.
- [10] Corradi, V. and N.R. Swanson, 2005, "A Bootstrap Specification Test for Diffusion Processes," *Journal of Econometrics*, 124, 117-148.
- [11] Corradi, V. and N.R. Swanson, 2006, "Simulation Based Predictive Density Estimation and Testing for Diffusion Processes," Working Paper, Rutgers University and Warwick University.
- [12] Cox, J.C., J.E. Ingersoll, and S.A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.
- [13] Diebold, F.X., T. Gunther, and A. Tay, 1998, "Evaluating Density Forecasts with Application to Financial Risk Management," *International Economic Review*, 39, 863-883.
- [14] Elerian, O., S. Chib, and N. Shephard, 2001, "Likelihood Inference for Discretely Observed Nonlinear Diffusions," *Econometrica*, 69, 959-993.
- [15] Gao, J., and M.L.King, 2004, "Adaptive Testing in Continuous-Time Diffusion Models," *Econometric Theory*, 20, 844-882.
- [16] Hong, Y., and H. Li, 2005, "Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37-84.
- [17] Li, F., and G. Tkacz, 2006, "A Consistent Bootstrap Test for Conditional Density Functions with Time-Series Data," *Journal of Econometrics*, 92, 101-147.

- [18] Li, F., 2007, "Testing the Parametric Specification of the Diffusion Function in a Diffusion Process," *Econometric Theory*, 23, 221-250.
- [19] Pedersen, A.R., 1994, "Uniform Residuals for Discretely Observed Diffusion Processes," *Technical Report 292*, Department of Statistics, Institute of Mathematics, University of Aarhus.
- [20] Pedersen, A.R., 1995, "A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete observations," *Scandinavian Journal of Statistics*, 22, 55-71.
- [21] Scott, D.W., 1992, *Multivariate Density Estimation, Theory, Practice, and Visualization*, John Wiley and Sons, New York.
- [22] Thompson, S.B., 2002, "Specification Tests for Continuous Time Models," Working Paper, Harvard University.
- [23] Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177-188.

Chapter 3

A Specification Test for Stock Return Models

3.1 Introduction

Stock return volatility models have been widely applied in practice for a variety of purposes, such as option pricing, optimal portfolio allocation, and risk assessment (e.g., Shephard, 2005). Among many alternatives, the ARCH/GARCH class of models (e.g., Engle, 1982; Bollerslev, 1986) and the stochastic volatility models (e.g., Rosenberg, 1972; Clark, 1973; Taylor, 1982) have been of special interest in the related literature. In the past, GARCH-type models were more popular, partly because these models could be easily estimated by the maximum likelihood, while the presence of latent state variables in stochastic volatility models made their estimation challenging. However, with the development of econometric technique, stochastic volatility models have become one of the major topics in the area of financial econometrics. Notably, the stochastic volatility models have developed along with option pricing methods. Heston (1993) has proposed the first rigorous option pricing formula that can be evaluated rapidly for the square root stochastic volatility model (also known as the Heston model). Later, Duffie et al. (2000) provide a general analytic treatment of option valuation problems via the affine-jump diffusion transform analysis. The Heston model belongs to the affine-jump diffusion class as a special case.⁴

In this study, we focus on a specification testing of time-series dynamics implied by a given model (under objective measure) using a statistical criterion. Since stock return models are widely used in practice, the investigation of a correctly

⁴Under the GARCH framework, Duan (1995), Bollerslev and Mikkelsen (1999), Kallsen and Taqqu (1998), and Heston and Nandi (2000) have proposed option pricing methods.

specified model has substantial importance. Ideally, a stock return model should be correctly specified from both (option) pricing and time-series criteria, that is, consistent across both "objective" and "risk-neutral" measures. In this context, Broadie et al. (2007) point out that forcing a misspecified model to fit observed prices is particularly dangerous if the fitted parameters are then used to price or hedge other derivatives. Therefore, there might be a natural question: can a specification testing from time-series criteria provide a meaningful implication to the risk-neutral dynamics, e.g. option pricing performance? In response to this question, Eraker et al. (2003) argue that, due to absolute continuity of the change in measure from an objective to a risk-neutral one, the presence of jumps in returns or volatility under one measure implies their presence under the other and that, in terms of model specification, time-series data should lead to the same conclusion as option price data. In fact, it is more difficult to evaluate the option pricing performance of a given model from statistical criteria, partly because there are currently few theories on the dynamics of option pricing error (e.g., Bates, 2000).

There have been proposed several specification testing methods in the literature; for example, various stationary distribution based tests, the EMM chi-square test (Gallant and Tauchen, 1997), the Bayes factor, and Hong and Li's (2005) nonparametric specification testing method, to name a few.⁵ The stationary distribution based tests (Ait-Sahalia, 1996; Corradi and Swanson, 1996; Bhardwaj et al., 2006) can capture the departure of an empirical stationary distribution from its model-implied counterpart. However, these testing methods are silent about a dynamic feature of the model specification, so they might lack some power against some misspecified models which have the same stationary distribution. Both the EMM chi-square test and the Bayes factor have been widely used in the EMM

⁵All methods listed here will be further detailed in Section 3.2.1.

and MCMC literature, respectively. However, the application of both methods are restricted by the model estimation methods (the EMM and MCMC, respectively); and, furthermore, it is hard to apply these methods to the out-of-sample context, even though the out-of-sample test is a very useful tool to detect an overfitting problem (or data-snooping bias). A more complicated model can always fit a given data set better than simpler models, but it may overfit some idiosyncratic features of the data without capturing the true data-generating process (e.g., Hong et al., 2004).

As introduced in the previous chapter, Hong and Li (2005) have proposed a nonparametric specification test using the transition density, which can capture the full dynamics of the underlying process. Using the Hong and Li statistics, one can evaluate a model performance in terms of density forecasting ability. Because of the i.i.d. property of the probability integral transforms, its finite sample performance is robust to persistent dependence in data, which is not favorable in some other tests. It is also generally applicable when testing both continuous- and discrete-time models. Since this method exploits the full dynamics of the underlying process, it has high power against almost any misspecified model. This feature is in contrast with existing stationary distribution based tests. Furthermore, this method is easily applicable to the out-of-sample test as in Hong et al. (2007) and does not depend on model estimation methods. Despite various merits of Hong and Li's (2005) method, it is challenging to directly apply the method to the famous stochastic volatility class of models. The key step in the method is a dynamic probability integral transform to compute generalized residuals, which requires a closed-form transition density. However, the presence of latent state variables (e.g., stochastic volatility) makes the transform analytically intractable.

Based on Hong and Li (2005) and Johannes et al. (2008), we propose a simulation-based specification testing method, which is applicable to stochastic volatility models. We circumvent the analytic intractability problem in Hong and Li’s (2005) test by using the particle-filtering approach proposed by Johannes et al. (2008). We approximate a dynamic probability integral transform via Monte Carlo integration by help of simulated particles for the latent state variables. In the stochastic volatility literature, the simulation-based dynamic probability integral transform was first used in Kim et al. (1998). Pitt and Shephard (1999) propose the APF (auxiliary particle filters) to make the algorithm more efficient. Later, Johannes et al. (2008) apply the particle filtering based on the APF algorithm to the affine jump-diffusion models. We should mention that our proposed method is based on all these studies.

By using the proposed specification testing method, we conduct a comprehensive empirical study on the stock return models using S&P500 stock index returns, from both in- and out-of-sample context. Our stock return models include various GARCH models (e.g., GARCH-N, GARCH-T, EGARCH, and GJR), to which the Hong and Li’s (2005) original testing method is applicable, as well as various stochastic volatility models (e.g., log stochastic volatility and square-root stochastic volatility models) with either jump-in-return or jump-in-volatility. Broadie et al. (2007) summarize the recent empirical studies, most of which have, however, focused only on the square-root stochastic volatility models (with or without jump). According to their summary, time-series studies unanimously support jump-in-return, but disagree over jump-in-volatility. For instance, Eraker et al. (2003) find strong evidence for the presence of stochastic volatility, jump-in-return, and also jump-in-volatility, while Chernov et al. (1999) find little evidence in support

of jump in volatility. On the other hand, option-based studies disagree over even the importance of jump-in-returns. Bakshi et al. (1997) show a large benefit of jump-in-returns, but Bates (2000) and Eraker (2004) find only marginal benefits of jump-in-returns. Broadie et al. (2007) are supporting the role of jumps in both returns and volatilities. Notably, there has been little comparative analysis on both GARCH and stochastic volatility models.⁶ This is partly due to the fact that there are only a few specification testing tools which can evaluate both classes of models on a fair basis. As noted earlier, the existing specification testing methods depend on the estimation method: for example, the EMM or MCMC method is normally used for stochastic volatility models and the maximum likelihood for the GARCH models. Moreover, there are also few out-of-sample studies. Our study is expected to fill this gap.

We estimate the stochastic volatility models via the Bayesian MCMC and the other models via the maximum likelihood. We run the MCMC algorithm by using the all-purpose Bayesian software WinBUGS.⁷ Meyer and Yu (2001) and Yu (2005) have shown that WinBUGS performs well in estimating the log stochastic volatility model. We find that WinBUGS also performs well in estimating more complicated square-root stochastic volatility models with or without jump.

Although we focus on statistical criteria, the importance of other economic criteria should not be underestimated. For example, some economic criteria, such as pricing and hedging performances for stock return models, have been applied in the literature. There is no guarantee that both statistical and economic criteria would provide the same testing result. Rather, both criteria are expected to provide useful feedback for each other, and one can obtain more robust model selection

⁶Exceptionally, Kim et al. (1998) compare GARCH with the Heston model for foreign exchange rates.

⁷WinBUGS stands for "Windows-version Bayesian Inference using Gibbs Sampling."

result when combining statistical and economic criteria. In this context, we also evaluate each model from an economic criterion. We conduct the VaR (value-at-risk) performance test for each model, which is expected to provide an important practical implication for portfolio allocation or risk assessment. The existing VaR adequacy testing methods such as the Kupiec (1995) test and the dynamic quantile test (Engle and Manganelli, 1999) are employed. In practice, various GARCH-type models (including RiskMetrics) have been widely used in implementing the VaR so far. It would be practically important to compare GARCH models with stochastic volatility models in terms of the VaR performance.

Our empirical analysis shows that all models are misspecified in terms of Hong and Li (2005) statistics. However, the stochastic volatility models perform relatively well in both in- and out-of-sample. Although the introduction of jump-in-return into the model marginally improves the in-sample fit, similar to the results from other existing studies, the stochastic volatility component (rather than jump) appears to have the first order effect in capturing stock return dynamics. Also, we find that modeling the leverage effect provides a substantial improvement in the log stochastic volatility models. On the other hand, although the GARCH-T model (i.e., GARCH with t-distributed innovations) performs as well as stochastic volatility models for in-sample evaluation, it fails to perform consistently well in out-of-sample. Our VaR performance analysis also shows that the stochastic volatility models outperform the GARCH models for both in- and out-of-samples, which implies that, in a practical sense, the stochastic volatility models can be considered as an alternative to the GARCH in the VaR implementation.

The remaining sections are organized as follows. Section 3.2 reviews the existing specification testing methods and then proposes our simulation-based specification

testing method using the particle filtering. Section 3.3 describes the stock return models examined in this study. In Section 3.4, we discuss the data characteristics, estimation methods and estimation results. Section 3.5 presents both in- and out-of-sample performance results for each model, and the value-at-risk performances are also presented. Finally, Section 3.6 concludes.

3.2 The Specification Testing Method

3.2.1 The Review of Existing Testing Methods

We selectively survey existing specification testing methods which have been introduced in the stochastic volatility literature. Here, we outline stationary distribution-based tests, EMM chi-square test, Bayes factor, and generalized residual method.

First, a variety of stationary distribution based tests have been introduced in the literature. Among them, the QQ (quantile-quantile) plot is a widely used informal diagnostic checking tool, usually used as a first step to detect model misspecification (e.g., Kim et al., 1998; Eraker, 2001; Eraker et al., 2003). Using the normalized forecasting errors from a given model, the QQ plot compares an empirical stationary distribution with its model-implied stationary distribution, which is often assumed to be standard normal. If non-normality of normalized errors is detected, one can suspect that there might be a model misspecification. Based on this idea, more formal specification testing tools have been proposed in the literature, which, like the QQ plot, can detect a model misspecification by observing the departure from model-implied stationary distribution. Among such testing methods are Ait-Sahalia's (1996) marginal density test, Corradi and Swanson's (1996) bootstrap specification test, and Bhardwaj et al.'s (2006) simulation-based specification test. However, these stationary distribution-based testing methods might miss some dynamic feature of the model specification, such as serial dependence of

normalized innovations, so they might not have power against some misspecified models which have the same stationary distribution as data generating processes.

Second, Gallant and Tauchen (1997) propose the EMM (Efficient Method of Moments) chi-square test, which is essentially a by-product of the EMM estimation procedure. As a first step, a Hermite polynomial-based, semi-nonparametric (hereafter, SNP) model is fitted to data, where the SNP has a larger number of parameters to be estimated than a parametric model of interest. Then, the parametric model is estimated by minimizing the EMM criterion function, which is a quadratic function of the score vector from a simulated parametric model-implied data process. The inverse of a sample covariance matrix for the score vector estimated from true data (not simulated data) is used as a weighting matrix in the EMM criterion function. To check the adequacy of a parametric model under consideration, the score vector from the SNP likelihood, which is zero by its nature, is compared with the score implied by the parametric counterpart. Under correct specification of the parametric model, both score vectors should be close to each other. For this comparison, one can use the EMM criterion function evaluated at the EMM estimator, which follows chi-square distribution, for an overall test of the overidentifying restrictions. If the test rejects, the individual elements of the score vector are able to provide useful information regarding the dimensions in which the parametric model fails to accommodate the data. These model diagnostics are based on the standard t-statistics of the individual elements of the score vector. This method has been used in much EMM literature such as the stock return studies of Andersen et al. (2002) and Chernov et al. (2003), and the empirical option study of Chernov and Ghysels (2000). However, Gallant and Tauchen (1997) note that, even if no evidence of model misspecification is found

from the test, one cannot still determine whether the model is indeed correctly specified. It's because their procedure lacks in power against a certain direction. Furthermore, this testing method is not applicable to the model estimated by a method different from the EMM and is not appropriate in out-of-sample tests.

Third, the Bayes factor is a popular testing method that assesses the goodness-of-fit of the models that are estimated by the Bayesian MCMC. It has been applied in a variety of MCMC literature (e.g., Kim et al., 1998; Eraker et al., 2003; and Yu, 2005). Johannes and Polson (2006) outline the method as follows. For a finite set of models, $\{M_i\}_{i=1}^M$, one can compute the posterior odds of model i over j . The posterior odds of M_i versus M_j is defined as

$$\frac{p(M_i | Y)}{p(M_j | Y)} = \frac{p(Y | M_i) p(M_i)}{p(Y | M_j) p(M_j)}.$$

The ratio of marginal likelihoods, $p(Y | M_i)/p(Y | M_j)$, is commonly called the Bayes factor. If it is greater than one, the data favors model i over model j and vice versa. One can compute the Bayes factor using the output of MCMC algorithm.⁸ Although the Bayes factor is easy to interpret, and can be used to compare the overall performances of non-nested models, the direct evaluation of Bayes factors is computationally intensive and can be numerically unstable for latent variable models, as argued by Jacquier et al. (2004). Furthermore, this testing method is only applicable to the models estimated by the MCMC.⁹ As another drawback, it provides little information about possible sources of model misspecification. Because of the high computational cost of the Bayes factor, as an alternative, the DIC (the Deviance Information Criterion) method has been

⁸Refer to Chib (1995) for a detailed procedure on how to compute the statistics.

⁹For example, Kim *et al.*(1998) compare the log stochastic volatility model with the GARCH model for various foreign exchange rates. In their study, the GARCH model is estimated by MCMC method, although GARCH could have been estimated by a more efficient and convenient method such as the maximum likelihood.

recently proposed in the MCMC literature.¹⁰

Finally, Kim et al. (1998) develop a model diagnostic checking tool for stochastic volatility models by using "generalized residuals" backed out from squared historical asset returns, which is essentially a dynamic probability integral transform of squared returns. Since volatility is a latent variable in stochastic volatility models, they use the particle filters, thereby computing generalized residuals via Monte Carlo integration. They map the resulting generalized residuals into standard normal random variables through the inverse standard normal distribution function. If a model is correctly specified, the resulting series should be i.i.d. standard normal distributed. To check this property, they suggest the Box-Ljung, unconditional normality, and heteroskedasticity tests. Since they examine the statistical feature of "squared" returns, their method might miss some asymmetry effect (e.g., leverage effect) possibly present in the first moment of returns. Our testing method is similar in principle to their method in that our specification method also uses a dynamic probability integral transform using the particle filters. The difference is that, as will be explained later, we transform stock returns themselves rather than squared returns, into generalized residuals and compute a formal test statistic; that is, the Hong and Li (2005) testing method.

3.2.2 Specification Testing with the Particle Filtering Method

This section introduces the specification testing method applicable to stochastic volatility models. We should mention that our method proposed in this section is based on Hong and Li (2005) and Johannes et al. (2008). Contrary to the GARCH models, stochastic volatility models involve latent state variables such as spot volatility, jump time and jump size. The presence of latent state variables,

¹⁰For details of the DIC, refer to Berg et al. (2004).

particularly stochastic volatility, makes it difficult to directly apply Hong and Li's (2005) testing method because dynamic probability integral transform is analytically intractable. To be specific, the integrand in Equation (3.1) generally has no analytic functional form. Furthermore, without observable model-implied spot volatilities, the transition density of the return at the current date depends on the entire history of past returns as conditioning variables.

To clarify our situation, a specific example will be helpful. Suppose that, as in the log stochastic volatility model, the conditional density of Y_t , the return at time t , depends only on the latent stochastic volatility, h_t , and that the conditional density of h_t depends on both h_{t-1} and Y_{t-1} . Note that the dependence of h_t on Y_{t-1} comes from the leverage effect. The integration problem can then be restated as the following multi-dimensional integration problem in (3.1). For simplicity, the parameter vector θ is suppressed.

$$\begin{aligned}
Z_t &= \int_{-\infty}^{Y_t} p(y | Y^{t-1}) dy \\
&= \int_{-\infty}^{Y_t} \int_0^{\infty} p(y | h_t) p(h_t | Y^{t-1}) dh_t dy \\
&= \int_{-\infty}^{Y_t} \int_0^{\infty} \int_0^{\infty} p(y | h_t) p(h_t | h_{t-1}, Y_{t-1}) p(h_{t-1} | Y^{t-1}) dh_{t-1} dh_t dy.
\end{aligned} \tag{3.1}$$

A general technical difficulty encountered here is to compute the above high-dimensional integration. In particular, among the product terms in (3.1), $p(h_{t-1} | Y^{t-1})$ is analytically intractable, whereas both $p(y | h_t)$ and $p(h_t | h_{t-1}, Y_{t-1})$ may have a closed functional form (e.g. as in the case of the log stochastic volatility model) or, at least, are approximable (e.g. as in the case of the square-root stochastic volatility model). The particle-filtering method can step in at this high-dimensional integration problem. The basic idea is that, if particles for h_{t-1} could be sampled from $p(h_{t-1} | Y^{t-1})$, and h_t could be simulated from each particle repre-

senting h_{t-1} via the transition density, $p(h_t | h_{t-1}, Y_{t-1})$, then we can approximate the high-dimensional integration through the Monte Carlo integration.

Pitt and Shephard (1999) define the particle filters as the class of simulation filters that recursively approximate the filtering random variable $L_t | y^t$ (y^t is a realized past history of returns) by "particles" L_t^1, \dots, L_t^N , with discrete probability mass of π_t^1, \dots, π_t^N . Hence a continuous variable is approximated by a discrete one with random support. These discrete points are viewed as samples from $p(L_t | y^t)$. Normally, all of the π_t^i are assumed to equal $1/N$. Then we require that as $N \rightarrow \infty$, the particles can be used to increasingly approximate well the density of $L_t | y^t$.

For implementing the particle filters, there are two alternatives: the sampling-importance resampling (SIR) algorithm (Gordon et al., 1993) and the auxiliary particle filters (APF) algorithm (Pitt and Shephard, 1999). The SIR algorithm simply uses a likelihood function (or an approximated likelihood function) as a resampling weight for each particle, so it is very simple and easy to code. However, as Pitt and Shephard (1999) argue, using the SIR, one might encounter a well-known "sample impoverishment" problem, which implies a sample degeneration during the resampling stage. Johannes et al. (2008) point out that this problem may be severe during the periods characterized by large movements driven by outliers. Unfortunately, this sometimes occurs to the financial markets (e.g., Black Monday in 1987). Thus, it is inevitable with the SIR to use a large number of particles at the expense of computational efficiency. To remedy this problem, Pitt and Shephard (1999) propose the APF algorithm, which can improve the efficiency of algorithm without incurring too much computational cost. Their algorithm corrects the sample impoverishment problem by re-ordering the algorithm, resampling first and propagating second, incorporating the new observation in both steps. Later,

Johannes et al. (2008) apply the APF algorithm to the square-root stochastic volatility models. Considering all these aspects, we adopt the APF algorithm in this study. We should mention that our approach is based on Johannes et al.'s (2008) optimal filtering method.¹¹ We modify their algorithm to fit our setting, adding to the procedure one more step for dynamic probability integral transform.

The particle-filtering method requires that the functional forms of both likelihood function and latent state variable process are known. While the log stochastic volatility model satisfies these requirements, the square-root stochastic volatility models do not provide exact analytic solutions to them. For the square-root models, we adopt the Euler-Maruyama discretization scheme, whereby we approximate both likelihood function and latent spot volatility process into an appropriate conditionally normal distributed one. Also, we approximate Poisson-distributed jump occurrences into Bernoulli-distributed random variable. As a result, stock return follows the mixture of normal distributions. The Euler-Maruyama discretization scheme is inevitably subject to some level of discretization bias. To reduce this bias, one might be able to augment arbitrarily many artificial data points between fixed sampling intervals. However, Johannes et al. (2008) show by their simulation study that the discretization bias is only modest for the daily frequency, and that, when volatilities are highly persistent, the bias becomes more negligible. Fortunately, they are the main features of stock returns studied in this chapter.

Now we illustrate the dynamic probability integral transform procedure applied to the log stochastic volatility model. In Appendix 3.1, we provide the algorithm for the SVCJ model based on Johannes et al. (2007). The other square-root stochastic volatility models are just a special case of the SVCJ. Note that, unlike

¹¹They argue that their filtering method can be used for estimating latent states, forecasting volatilities and returns, computing likelihood ratios, and parameter estimation. Our paper suggests using the filtering method for dynamic probability integral transform.

the square-root stochastic volatility models, there is no discretization bias in the log stochastic volatility model.

Consider the following log stochastic volatility model:

$$\begin{cases} y_t = \mu + \sqrt{h_t}\varepsilon_t, \\ \Delta \ln h_t = \alpha + \beta \ln h_{t-1} + \sigma_h \rho \varepsilon_{t-1} + \sigma_h \sqrt{1 - \rho^2} \eta_t, \end{cases} \quad (3.2)$$

where $y_t = \Delta \ln S_t$, and $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$. The dynamic probability integral transform for this model works recursively as follows.

- **Step 1:** Suppose that we are given $\{h_{t-1}^{(i)}\}_{i=1}^N$, N particles for spot volatility at time $t - 1$, which have been drawn from the density, $p(h_{t-1} | y^{t-1})$. For each i , draw $\tilde{h}_t^{(i)}$ from the transition density, $p(\Delta \ln h_t | h_{t-1}^{(i)}, y_t)$, which is conditional normal density implied by the second equation in (3.2). With the sampled $\{\tilde{h}_t^{(i)}\}_{i=1}^N$, approximate the generalized residual at time t as in (3.3).

$$\begin{aligned} z_t &= \int_{-\infty}^{y_t} p(y | y^{t-1}) dy \\ &= \int_{-\infty}^{y_t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y | h_t) p(h_t | h_{t-1}, y_{t-1}) p(h_{t-1} | y^{t-1}) dh_{t-1} dh_t dy \\ &\simeq \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} p(y | \tilde{h}_t^{(i)}) dy \\ &= \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} \phi(y | \mu, \tilde{h}_t^{(i)}) dy \end{aligned} \quad (3.3)$$

where $\phi(\cdot | \mu, \sigma^2)$ represents a normal density function with mean μ and variance σ^2 . After computing the generalized residual at time t , we discard $\{\tilde{h}_t^{(i)}\}_{i=1}^N$.

- **Step 2:** Step 2 starts to update the particles from $\{h_{t-1}^{(i)}\}_{i=1}^N$ (a sample from $p(h_{t-1} | y^{t-1})$) to $\{h_t^{(i)}\}_{i=1}^N$ (a sample from $p(h_t | y^t)$). First, compute the auxiliary variable, $\hat{h}_t^{(i)}$ for each i as the following equation. $\hat{h}_t^{(i)}$ is a

conditional expectation of h_t given $h_{t-1}^{(i)}$ and y_{t-1} .¹² This variable plays a role in improving the efficiency of the algorithm.

$$\widehat{h}_t^{(i)} = \exp \left(\ln h_{t-1}^{(i)} + \alpha + \beta \ln h_{t-1}^{(i)} + \rho \sigma_h (y_{t-1} - \mu) / \sqrt{h_{t-1}^{(i)}} + \frac{1}{2} \sigma_h^2 (1 - \rho^2) \right),$$

where the term, $\rho \sigma_h (y_{t-1} - \mu) / \sqrt{h_{t-1}^{(i)}}$, comes from the leverage effect (i.e. dependence of h_t on y_{t-1}), and the term, $\frac{1}{2} \sigma_h^2 (1 - \rho^2)$ comes from the Jensen's inequality adjustment.

Using $\widehat{h}_t^{(i)}$, compute the 1st stage weight for each i :

$$w_t^{(i)} \propto \phi(y_t | \mu, \widehat{h}_t^{(i)}).$$

Using the 1st stage weights, resample the particles, and obtain $\{h_{t-1}^{k(i)}\}_{i=1}^N$.

- **Step 3:** Generate the stochastic volatility at time t , that is, $h_t^{k(i)}$ for each $h_{t-1}^{k(i)}$ by drawing $\widetilde{\eta}_t^{k(i)} \sim N(0, 1)$:

$$\ln h_t^{k(i)} = \ln h_{t-1}^{k(i)} + \alpha + \beta \ln h_{t-1}^{k(i)} + \rho \sigma_h (y_{t-1} - \mu) / \sqrt{h_{t-1}^{k(i)}} + \sigma_h \sqrt{1 - \rho^2} \widetilde{\eta}_{t+1}^{k(i)}.$$

- **Step 4:** Compute the 2nd stage weight for each $k(i)$ ¹³:

$$\pi_t^{k(i)} \propto \frac{p(y_t | h_t^{k(i)})}{p(y_t | \widehat{h}_t^{k(i)})} = \frac{\phi(y_t | \mu, h_t^{k(i)})}{\phi(y_t | \mu, \widehat{h}_t^{k(i)})}.$$

Using the 2nd stage weights, resample the particles, $\{h_t^{k(i)}\}_{i=1}^N$ and obtain the final updated particles, $\{h_t^{(i)}\}_{i=1}^N$, which amounts to the sample drawn from $p(h_t | y^t)$.

¹²In fact, the original APF algorithm by Pitt and Shephard (1999) has a high flexibility in choosing an auxiliary variable. One does not need to stick to our conditional expectation.

¹³Pitt and Shepherd (1999) propose the APF algorithm in the hope that the second-stage weight ($\pi_t^{k(i)}$) are much less variable than for the original SIR method. By introducing the 1st stage reweighting, one can reduce the costs of sampling many times from particles that have very low likelihoods and so will not be resampled at the second stage of the process. This improves the statistical efficiency of the sampling procedure and means that one can reduce the number of particles substantially. For further details, refer to Pitt and Shephard (1999).

- **Step 5:** Go back to Step 1 and compute the generalized residual at time t .

The remaining procedure is to compute the $Q(j)$ statistics as introduced in the previous section. Now, we can summarize the overall procedure for nonparametric specification testing:

- Estimate a model using any method that yields a \sqrt{n} -consistent estimator $\hat{\theta}$;
- Implement the dynamic probability integral transform. When a model involves latent variables like stochastic volatility, compute the generalized residuals via the particle-filtering method outlined in this section. Otherwise, directly apply Hong and Li's (2005) method;
- Compute Hong and Li's (2005) test statistic $\hat{Q}(j)$ or $\hat{W}(p)$ using the sequence of estimated generalized residuals from Step 2. If $\hat{Q}(j)$ (or $\hat{W}(p)$) $> C_\alpha$, the upper tailed $N(0,1)$ critical value at significance level α , then we reject the null hypothesis that the model is correctly specified at level α .

In practice, it is required to choose the number of particles (N). Following Kim, Shephard and Chib (1998), we set N equal to 2,500. We find that, by a separate experiment (not reported here), a smaller value of N such as 1,000 does provide nearly identical results as N equal to 2,500, implying that a convergence has occurred with our chosen number of particles.

It is worth noting that the particle-filtering procedure can be easily adapted to computing the Value-at-Risk (VaR) measure at a given critical value (e.g. $p\%$). One can compute the VaR measure at Step 1 in the above procedure. For instance, to compute $p\%$ VaR measure for the portfolio tracking the stock index, set z_t in (3.3) equal to $0.01p$, then find the value for y_t that makes both sides in (3.3) equal. This y_t gives the maximum loss of the stock index portfolio at $(1 - p)\%$

confidence level. For more complicated portfolio that contains derivatives, one can use a simulation method. Also at Step 1, simulate pairs of return and volatility, that is, $\{\tilde{y}_t^{(i)}, \tilde{h}_t^{(i)}\}_{i=1}^N$, and calculate a portfolio return for each i th particle. Then, by taking p th percentile from the simulated portfolio returns, one can obtain the $p\%$ VaR measure of the portfolio. The VaR performance of stock return models will be presented in Section 3.5.

3.3 Models

This section introduces the stock return models we will analyze in this study. As benchmark models, we consider the Black-Scholes (BS) and Pure Jump (PJ) models. As the GARCH class of models, we study the GARCH-N, GARCH-T (each characterized by the assumption on distribution of error term), EGARCH (Exponential GARCH; Nelson, 1991), and GJR (also known as Threshold GARCH; Glosten et al., 1993). The latter two models explicitly incorporate the leverage effect. We examine two stochastic volatility classes: log stochastic volatility and square-root stochastic volatility models. For the log SV models, we consider two specifications: one is modeling the leverage effect (LSV) and the other is not (LSV0). For the square-root SV class of models, we study three specifications: SV, SVJ, and SVCJ, each of which is characterized by its jump specification as will be detailed below.

3.3.1 Black-Scholes and Pure Jump Models

The Black-Scholes (BS) model (also called the geometric random walk model) in Equation (3.4) assumes a log-normal distribution for stock prices.

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (3.4)$$

where S_t is a stock price at time t , and W_t is a Brownian motion.

Under this specification, the conditional variance of asset return is constant (conditionally homoskedastic). The above process is assumed by the famous Black-Scholes option pricing formula. Because of its simplicity, it has been widely used for modeling asset price dynamics. However, it is well known that the model cannot capture some stylized facts for high-frequency asset returns, such as leptokurtosis, skewness, and conditional heteroskedasticity. In terms of risk-neutral dynamics, this model is known to induce a systematic bias for option pricing, in particular, so-called "volatility smiles" of option-implied volatilities across different strikes. As in Bakshi et al. (1997) and Andersen et al. (2002), we use the BS models as a benchmark.

Merton (1976) proposes the pure jump (PJ) model. It extends the BS model by incorporating a compound Poisson jump process, based on historical evidences about discontinuities in asset returns. Various economic shocks, news announcement, and government interventions in markets might induce large jumps in financial data. The model specification is as follows:

$$\begin{cases} d \ln S_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t + dJ_t, \\ dJ_t = Z_t^S dN_t, N_t \sim Poi(\lambda), Z_t^S \sim N(\mu_S, \sigma_S^2), \end{cases} \quad (3.5)$$

where $Poi(\lambda)$ represents a Poisson distribution with the jump arrival intensity λ .

The added jump component can help to accommodate outliers during market stress and can also induce asymmetry (e.g. leverage effect) in return distribution by introducing nonzero average jump amplitude (μ_S). Andersen et al. (2002) argue that, although this specification can capture some patterns of skewness and leptokurtosis which the BS cannot capture, it can neither account for the volatility clustering of returns nor rationalize the substantial time variations in the level and shape of the implied volatility smile. In the literature, the PJ model has been studied by Press (1967), Jarrow and Rosenfeld (1984), Ball and Torous (1985),

Akgiray and Booth (1986), Das and Uppal (1998), and Das (2002).

3.3.2 GARCH Models

A useful approach to modeling volatility clustering is the ARCH or GARCH class of models (Engle, 1982; Bollerslev, 1986). These models can generate unconditionally leptokurtic distribution for asset returns, while their conditional distributions might be normal. Although the GARCH and stochastic volatility models were introduced almost contemporaneously, the GARCH had been studied a lot more in the past because the maximum likelihood is easily applicable to its estimation. Contrary to stochastic volatility models, the GARCH models are able to provide a known conditional variance at a current date. We consider two different distributional assumptions on innovation terms: the GARCH-N assumes normal distributed innovations and the GARCH-T allows for fatter tails by assuming student-t distribution. Each model is specified as

$$\begin{cases} \Delta \ln S_t = \mu + \sqrt{h_t} \varepsilon_t, \\ h_t = a_0 + a_1 h_{t-1} + a_2 \varepsilon_{t-1}^2, \end{cases} \quad (3.6)$$

where $\varepsilon_t \sim \text{i.i.d. } N(0, 1)$ (for GARCH-N) or $\varepsilon_t \sim \text{i.i.d. } \sqrt{\frac{\nu-2}{\nu}} t(\nu)$ (for GARCH-T).

There have been some studies that aim at finding an implication of the GARCH for continuous-time stochastic models. For example, Nelson (1990) showed that the GARCH models converge to the stochastic volatility counterpart in the limit. However, Corradi (2000) later argued that Nelson's results hold only under his particular discretization scheme, and that other equally reasonable discretization may lead to quite different continuous-time limits for GARCH models. Therefore, special care should be taken in the interpretation of the GARCH model in this context.

Later, some GARCH variants that incorporate the leverage effect have been in-

troduced. Among them, we consider the EGARCH (exponential GARCH; Nelson, 1991) and GJR (also known as threshold GARCH; Glosten et al., 1993). Their conditional variance (h_t) specifications are

$$\begin{cases} \ln h_t = a_0 + a_1 \ln h_{t-1} + a_3 \varepsilon_{t-1} / \sqrt{h_{t-1}} + a_2 \left| \varepsilon_{t-1} / \sqrt{h_{t-1}} \right| & \text{(EGARCH),} \\ h_t = a_0 + a_1 h_{t-1} + a_2 \varepsilon_{t-1}^2 + a_3 \varepsilon_{t-1}^2 1_{[\varepsilon_{t-1} < 0]} & \text{(GJR),} \\ \varepsilon_t \sim \text{i.i.d. } N(0, 1) \text{ for both models.} \end{cases} \quad (3.7)$$

In the above specifications, the parameter a_3 in both EGARCH and GJR plays a role in capturing the leverage effect. If the leverage effect is substantial, a_3 should be significantly negative (positive) in the EGARCH (GJR).

Related to option pricing, Duan (1995) proposes a GARCH framework for option pricing under a local risk-neutrality assumption. Under his framework, option contracts can be priced via Monte-Carlo simulation. Later, Heston and Nandi (2000) propose another variant of GARCH model which provides a closed-form option pricing method up to the Fourier inversion.

3.3.3 Log Stochastic Volatility Models

Unlike the GARCH models, stochastic volatility models assume that spot volatility follows an unobservable stochastic process. Among many other stochastic volatility specifications, the log stochastic volatility model as in Equation (3.8) is so popularly applied that it has been sometimes called *the* stochastic volatility model (e.g., Talor, 2005). In most cases, this log stochastic volatility model has been studied under discrete-time setting. As pointed out by Taylor (2005), this specification permits convenient calculation of moments and allows for any level of unconditional excess kurtosis in returns. The typical model specification is as follows:

$$\begin{cases} \Delta \ln S_t = \mu + \sqrt{h_t} \varepsilon_t, \\ \Delta \ln h_t = \alpha + \beta \ln h_{t-1} + \sigma_h \rho \varepsilon_{t-1} + \sigma_h \sqrt{1 - \rho^2} \eta_t, \end{cases} \quad (3.8)$$

where $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$.

In Equation (3.8), nonzero ρ can accommodate a leverage effect. Many empirical studies have shown that the leverage effect is substantial for stock return dynamics, whereas it is less important for interest rate or foreign exchange rate dynamics. In some log stochastic volatility studies such as Kim, Shephard and Chib (1998) and Chib, Nardari and Shephard (2002), the leverage effect has been sometimes ignored in pursuit of more an efficient MCMC algorithm. Considering this, we address two specifications: we denote by LSV the unrestricted model and denote by LSV0 the restricted model with $\rho = 0$.

In time-series literature, the log stochastic volatility model has been studied by Jacquier et al. (1994), Harvey and Shephard (1996), Andersen et al. (2002), and Yu (2005) for stock returns and Kim et al. (1998) for foreign exchange rates, among others. In spite of its popularity, the log stochastic volatility specification is less convenient for derivative pricing than square-root stochastic volatility models below. Complicated numerical methods are required for pricing derivatives. For this reason, there have rarely been empirical option pricing studies associated with this specification.

3.3.4 Square-Root Stochastic Volatility Models

In contrast to discrete-time log stochastic volatility specification, the square-root stochastic volatility model assumes a continuous-time square-root or CIR-type (Cox, Ingersoll, and Ross, 1985) spot volatility process, possibly with jump components. Following Eraker et al. (2003), we consider three specifications in this class: the SV, SVJ, and SVCJ models.¹⁴ The SV is a plain square-root stochastic volatility model without jump, the SVJ incorporates a jump-in-return component,

¹⁴Eraker et al.(2003) consider another double jump specification called SVIJ. In the SVIJ, both jumps arrive independently. However, they document that the SVIJ does not improve substantially upon the SVCJ.

and the SVCJ allows for contemporaneously arriving jumps in both return and volatility. Equation (3.9) below provides a general specification for those models. Note that the SV is a special case where $N_t = 0$ almost surely and that the SVJ has the restriction that $Z_t^v = 0$ almost surely.

$$\begin{cases} d \ln S_t = (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^S + dJ_t^S, \\ dV_t = \kappa(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t^v + dJ_t^v, \end{cases} \quad (3.9)$$

where $Cov(dW_t^S, dW_t^v) = \rho dt$, $dJ_t^i = Z_t^i dN_t$ for $i = \{S, V\}$, $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$ (jump-in-return size), $Z_t^v \sim \text{i.i.d. } \exp(\mu_v)$ (jump-in-volatility), and $N_t \sim Poi(\lambda)$ (jump timing).

Hull and White (1987) show that the SV is able to explain some anomalies empirically observed in the Black-Scholes implied volatilities (hereafter, BSIV), particularly the "volatility smile" phenomenon. The leverage effect is explicitly incorporated into the model, which can capture the asymmetry in the market-implied BSIV curve (e.g. "volatility smirk"). It is particularly convenient for option pricing, thanks to a closed-form formula for pricing European options, which was first proposed by Heston (1993). For this reason, the SV model has been sometimes called the Heston model. Andersen et al. (2002) and Chernov et al. (2002) find that the log-volatility and SV models provide a nearly identical fit to time-series data. In a similar vein, Benzoni (1998) argues that, in terms of option pricing, there is no qualitative difference between the two models. We will also compare the two models using our proposed testing method.

Andersen et al. (2002), however, show that both log stochastic volatility and SV models cannot adequately capture the fat tails observed in the return distribution. The introduction of jump component into the SV can help to better accommodate this feature. In this context, Andersen et al. (2002) argue that the SVJ¹⁵ provides

¹⁵Alternatively, jump-in-return component can be incorporated into log stochastic volatility

additional flexibility in capturing some important features of equity returns such as skewness and leptokurtosis, and also that the incorporation of a jump component is essential when pricing the options that are close to maturity. The SVJ model provides two sources of nonzero skewness: i) a nonzero mean jump amplitude (μ_S in (3.9)) and ii) negative correlation between the return and volatility shocks (ρ in (3.9)).¹⁶ Most time-series stock return studies (e.g., Andersen et al., 2002; Eraker et al., 2003) have found that the SVJ model markedly improves upon the SV. In particular, Andersen et al. (2002) find no statistical evidence in terms of EMM chi-square test that the SVJ is misspecified. The estimates of jump intensity in the past studies shows that jumps occur very rarely: three to four jumps per year in Andersen et al. (2002) and 1.5 times per year in Eraker et al. (2003), for example.

On the other hand, Bakshi et al. (1997), Bates (2000), and Pan (2002) find that the SVJ are incapable of fully capturing some empirical features observed in stock returns or option prices. Similarly, Eraker et al. (2003) also find empirical evidence that the conditional volatility of returns rapidly increases, a movement difficult to generate using the SVJ, especially during the periods of market stress which are characterized by a short time period with multiple large movements. The introduction of jump-in-volatility in the model can allow volatility to rapidly increase. Eraker et al. (2003) provide statistical evidence that the SVCJ fits the stock return data better than the SVJ.

All the square-root SV class of models belong to the affine jump diffusion (AJD) class specified by Duffie et al. (2000). These models all provide a tractable option pricing formula. This is obviously a great advantage over the log stochastic

specification. However, the empirical result in Andersen, Benzoni, and Lund (2002) shows that both models are very similar in goodness of fit in terms of EMM chi-square criteria. For this reason, we don't consider the log stochastic volatility model with jump.

¹⁶However, Andersen et al.(2002) show that the mean jump amplitude parameter is insignificant, which suggests that the asymmetry is more appropriately captured through ii).

volatility model.

3.4 Model Estimation

3.4.1 Data

We estimate the stock return models using S&P 500 return data from January 1988 to December 2000. Excluding weekends and holidays, we have 3,285 daily observations. We save the data from January 2001 to December 2007 for out-of-sample evaluation, where the number of observations is 1,757.

Table 3.1 presents the summary statistics for continuously compounded stock returns, scaled by 100. Interestingly, the signs of skewness are different across both samples. In the in-sample, the returns exhibit a left skewness with the magnitude equal to -0.46. The left skewness is known to be a typical pattern of equity returns for the stock markets in most developed countries (e.g., Singleton, 2006). However, the sign of skewness turns into the slightly positive (0.07) in the out-of-sample. Thus, it would be interesting to examine out-of-sample performances for the models which explicitly take into account the leverage effect. On the other hand, the kurtosis is much higher than 3 in both samples (8.6 and 5.7, respectively), indicating that unconditional return distribution is far from normal distribution. This implies that the BS model might be far from reality.

Table 3.1: Summary Statistics for Stock Returns

Period	Mean	SD	Skewness	Kurtosis	Max.	Min.
In-sample (1988-2000)	0.051	0.949	-0.455	8.592	4.989	-7.113
Out-of-sample (2001-2007)	0.006	1.069	0.074	5.681	5.574	-5.047

3.4.2 Estimation Method

Among stock return models considered, the BS, PJ, and GARCH models can be easily estimated by the maximum likelihood (ML) method, whereas the estimation of stochastic volatility models are rather challenging. It's because these models do not have a closed-form likelihood function, and stochastic volatility is an unobservable state variable. Recently, there have developed a variety of estimation methods for those models.¹⁷ Among many alternatives, we choose the Bayesian MCMC method. This method is essentially a likelihood-based estimation method, and some previous studies have shown that, in terms of efficiency, likelihood-based inference methods are superior to the moment-based counterparts such as GMM and EMM. For instance, Andersen et al. (1999) provide their Monte Carlo simulation results that the MCMC outperforms the EMM method in estimating log stochastic volatility model.

Among stochastic volatility models, square-root SV models are continuous-time diffusion models, but for the estimation purpose, one should rely on a discretely sampled data set. Following Eraker et al. (2003) and many other related MCMC studies, we adopt the Euler-Maruyama discretization scheme. Although this scheme may create a discretization bias, the bias is known to be quite small with high-frequency data such as daily or even weekly data. In particular, Eraker et al. (2003) present their simulation study result to support this discretization scheme.

We implement the Bayesian MCMC via WinBUGS, which is a recently devel-

¹⁷For example, there are GMM by Melino and Turnbull (1990), and Andersen and Sorensen (1996); QMLE by Harvey, Ruiz, and Shephard (1994), and Harvey and Shephard (1996); EMM by Andersen, Chung, and Sorensen (1999) and Andersen, Benzoni, and Lund (2002); simulation-based maximum likelihood by Danielsson (1994), and Danielsson and Richard (1993); and Markov Chain Monte Carlo (MCMC) method by Jacquier, Polson, and Rossi (1994), Kim, Sephard, and Chib (1998), Eraker (2001), Eraker, Johannes, and Polson (2003) and Yu (2005).

oped all-purpose Bayesian software. Meyer and Yu (2000) and Yu (2005) show that WinBUGS performs well in estimating log stochastic volatility models. Although its computing time is relatively long because its single move Gibbs sampler makes the convergence of chains rather slow, any modification of a model can be easily accommodated with a minor change of the code. We use Meyer and Yu's (2000) WinBUGS code¹⁸ for estimating log stochastic volatility models, and for the square-root SV models, we developed our own codes based on Meyer and Yu. We find that the WinBUGS also performs well in more complicated square-root stochastic volatility specifications equipped with a jump-in-volatility and/or a jump-in-return. Following Eraker et al. (2003) and Yu (2005), we ran the MCMC algorithm for 110,000 iterations, discarding the first 10,000 as a burn-in period to achieve the convergence of the chain.

For the MCMC algorithm, the prior distribution for each parameter should be specified. For the square-root stochastic volatility models, we adopt the same priors as from Eraker et al. (2003): $\mu \sim N(1, 25)$, $\kappa\theta \sim N(0, 1)$, $\kappa \sim N(0, 1)$, $\sigma_v^2 \sim IG(2.5, 0.1)$, $\rho \sim U(-1, 1)$, $\lambda_y = \lambda_v \sim Beta(2, 40)$, $\mu_y \sim N(0, 100)$, $\sigma_y^2 \sim IG(5.0, 20)$, $\mu_v \sim G(20, 10)$, where G , IG , and U refer to the Gamma distribution, the Inverse Gamma distribution, and the standard uniform distribution, respectively. For the log stochastic volatility models, we choose the priors from Yu (2005), which are $\sigma_v^2 \sim IG(2.5, 0.025)$, $\beta^* \sim Beta(20, 1.5)$, where $\beta^* = (\beta + 1)/2$, $\alpha^* \sim N(0, 25)$, where $\alpha^* = \alpha/(1 - \beta)$, and $\rho \sim U(-1, 1)$. Yu (2005) does not estimate μ , since he uses mean-adjusted returns for estimation, so we use the same prior for μ as in the square-root stochastic volatility models.

¹⁸It is available from Jun Yu's webpage (<http://www.mysmu.edu/faculty/yujun>).

3.4.3 Estimation Result

Table 3.2 reports our parameter estimates for the BS and PJ models. In the BS model, the volatility parameter σ is estimated to be 0.95, which reflects an annualized standard deviation of 15.5%. The PJ model reduces the magnitude of σ from 0.95 to 0.61 by introducing a jump component and substantially improves a goodness-of-fit in terms of log likelihood. Interestingly, the jump arrival intensity λ in the PJ is estimated to be 0.26, which implies 64 jumps per year on average. Our estimate is quite different from the estimate from Andersen et al. (2002), who attain 0.059 from the sample period from 1980 through 1996. This marked difference implies that, for the PJ model, the jump arrival rate estimate is very unstable, depending heavily on the sample period chosen.¹⁹ It might be due to the fact that, as pointed out by Singleton (2006), the likelihood function may use the parameters of jump process to compensate for a misspecified volatility process.

Table 3.2: Parameter Estimates for BS and PJ Models

	μ	σ	λ	μ_S	σ_S	Log-likelihood
B S	0.0510 (0.0168)	0.9489 (0.0118)	0	0	0	-4489.0
P J	0.0744 (0.0161)	0.6153 (0.0252)	0.2554 (0.0359)	-0.0916 (0.0732)	1.4274 (0.0750)	-4265.5

Note: The models are specified by $\Delta \ln S_t = \mu + \sigma \varepsilon_t + J_t Z_t^S$, where $J_t \sim \text{i.i.d. } Ber(\lambda)$ and $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$. Standard errors are given in parentheses.

The various GARCH-type models (e.g. GARCH-N, GARCH-T, EGARCH, and GJR) presented in Table 3.3 appear to fit the data much better than the BS and PJ models in terms of log likelihood value, which evidences that a conditional heteroskedasticity is important in modeling stock return dynamics. Among

¹⁹We estimated the PJ model using the sample period from 1986 to 2000 (not reported). We obtained very low λ , which magnitude is very close to Andersen et al. (2003).

the four GARCH-type models, the GARCH-T exhibits a higher likelihood value than the other models. For the GARCH-T, we obtain a relatively low degree of freedom estimate (ν equal to 5.5) indicating that its innovation is far from normal-distributed. The estimation results of the EGARCH and GJR show that the introduction of leverage-effect parameter improves modestly upon the GARCH-N. We obtain highly significant estimates for the leverage-effect coefficients (a_3 for both EGARCH and GJR in Table 3.3) with expected signs (negative for EGARCH and positive for GJR). However, their likelihood values are lower than that of GARCH-T. More flexible distributional assumption on innovation (e.g. student t-distribution) might have improved the in-sample goodness-of-fit for both models.

Table 3.3: Parameter Estimates for GARCH Models

	GARCH-N	GARCH-T	GJR	EGARCH
μ	0.0526 (0.0136)	0.0639 (0.0122)	0.0435 (0.0145)	0.0355 (0.0139)
a_0	0.0029 (0.0009)	0.0028 (0.0014)	0.0120 (0.0015)	-0.0829 (0.0075)
a_1	0.9606 (0.0061)	0.9622 (0.0076)	0.9352 (0.0052)	0.9826 (0.0022)
a_2	0.0384 (0.0059)	0.0355 (0.0071)	0.0135 (0.0066)	-0.0696 (0.0073)
a_3	0	0	0.0740 (0.0091)	0.1043 (0.0097)
ν	-	5.4553 (0.4674)	-	-
Log-likelihood	-4174.0	-4037.6	-4160.4	-4137.7

Note: The GARCH-N, GARCH-T, and GJR models are nested by $\Delta \ln S_t = \mu + \sqrt{h_t} \varepsilon_t$, and $h_t = a_0 + a_1 h_{t-1} + a_2 \varepsilon_{t-1}^2 + a_3 \varepsilon_{t-1}^2 1_{[\varepsilon_{t-1} < 0]}$, where $\varepsilon_t \sim$ i.i.d. $N(0, 1)$ for both GARCH-N and GJR, and $\varepsilon_t \sim$ i.i.d. $\sqrt{\frac{\nu-2}{\nu}} t(\nu)$ for GARCH-T. The conditional variance for the EGARCH is specified by $\ln h_t = a_0 + a_1 \ln h_{t-1} + a_2 \varepsilon_{t-1} / \sqrt{h_{t-1}} + a_3 \left| \varepsilon_{t-1} / \sqrt{h_{t-1}} \right|$. Standard errors are in parentheses.

Next, Table 3.4 and 3.5 present the estimation results for various stochastic volatility models. As noted above, all stochastic volatility models are estimated by the Bayesian MCMC,²⁰ so log likelihood value is not available for these models.²¹ For the two log stochastic volatility models (the LSV and LSV0 in Table 3.4), we attain a highly significant leverage effect estimate (ρ) from the LSV model, which equals -0.47 with standard deviation of 0.06 . Notably, all remaining estimates also differ in magnitude across both models. Like the previous case of PJ model, it seems likely that the remaining parameters in the LSV0 are used to compensate for a misspecification.

Table 3.4: Parameter Estimates for Log Stochastic Volatility Models

	μ	α	β	σ_h	ρ
LSV	0.0410 (0.0129)	-0.0113 (0.0064)	0.0340 (0.0098)	0.1688 (0.0205)	-0.4679 (0.0562)
LSV0	0.0683 (0.0126)	-0.0015 (0.0047)	0.0140 (0.0082)	0.1478 (0.0240)	0

Note: The models are nested by the specification by $\Delta \ln S_t = \mu + \sqrt{h_t} \varepsilon_t$, and $\Delta \ln h_t = \alpha + \beta \ln h_{t-1} + \sigma_h \rho \varepsilon_{t-1} + \sigma_h \sqrt{1 - \rho^2} \eta_t$, where $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$. The estimates correspond to percentage changes in the index value. Standard deviations of posteriors are reported in parentheses.

²⁰In implementing MCMC, it is important to check whether each chain converges to its stationary distribution. Following Yu (2005), we implemented the Heidelberger and Welch (1983) stationary test by using the CODA package from the R statistical software. While all the other models pass the test, some parameters in the SVCJ fail to pass the test. We, however, have found that all parameters are stably estimated. In other words, for every trial of estimating the SVCJ, we always obtain quite similar results.

²¹It is possible to compute a marginal likelihood value when MCMC is used for estimation. However, it is practically complicated to compute the statistics, and in addition, it is known that the marginal likelihood value is unstable when latent variables are involved, which is exactly our case.

Table 3.5: Parameter Estimates for Square Root Stochastic Volatility Models

	SV	SVJ	SVCJ
α	0.0378 (0.0129)	0.0432 (0.0130)	0.0437 (0.0128)
κ	0.0307 (0.0067)	0.0204 (0.0049)	0.0295 (0.0061)
θ	0.8770 (0.0945)	0.8831 (0.1244)	0.6970 (0.0852)
σ_v	0.1631 (0.0155)	0.1333 (0.0121)	0.1339 (0.0126)
ρ	-0.5190 (0.0518)	-0.6129 (0.0525)	-0.5978 (0.0570)
λ	-	0.0125 (0.0056)	0.0051 (0.0020)
μ_S	-	-1.5600 (0.8369)	-3.7220 (1.0668)
σ_S	-	2.0980 (0.3296)	2.0828 (0.4202)
μ_v	-	-	1.6166 (0.3623)

Note: The models are nested by the specification: $\Delta \ln S_t = \alpha + \sqrt{V_t} \varepsilon_t + J_t Z_t^S$, and $V_t = V_{t-1} + \kappa(\theta - V_{t-1}) + \rho \sigma_v (y_t - \alpha) + \sqrt{1 - \rho^2} \sigma_v \sqrt{V_{t-1}} \eta_t + J_{t-1} Z_t^v$, where $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$, $J_t \sim \text{i.i.d. } \text{Ber}(\lambda)$, $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$, and $Z_t^v \sim \text{i.i.d. } \exp(\mu_v)$. The estimates correspond to percentage changes in the index value. Standard deviations of posteriors are reported in parentheses.

Table 3.5 reports the estimation results for square-root stochastic volatility models (i.e., SV, SVJ, and SVCJ). For the diffusion part of each model, we obtain very small magnitude of κ (roughly 0.03), which reflects a slow mean reversion tendency of spot volatility. This shows a highly persistent nature of spot volatilities in stock return dynamics.²² Then, the long-run average volatility parameter, θ , is estimated to be 0.88 for both SV and SVJ models, indicating that annual average standard deviation of stock returns (i.e., $\sqrt{252 \times \theta}$) is around 14.9%. However, for the SVCJ, θ is estimated to be as small as 0.70. This is not surprising because, for the SVCJ model, the average volatility level is determined by the formula, $\theta + \mu_v \lambda / \kappa$, rather than θ alone (e.g., Eraker *et al.*, 2003). By computing an annualized standard deviation from $\theta + \mu_v \lambda / \kappa$, we obtain 15.7% for the SVCJ. All the square-root models' implied annual standard deviations are similar in magnitude to that from the BS (15.5%). Turning to the volatility-of-volatility parameter, σ_v , it is observed that its magnitudes for the SVJ and SVCJ are smaller than for the SV

²²Note that $\kappa = 0$ implies that spot volatility is a random walk (e.g., Singleton, 2006).

(i.e., 0.16 for SV vs. 0.13 for both SVJ and SVCJ). This seems probably because the role of volatility to match the variability of returns is reduced due to the incorporation of jump component in the SVJ and SVCJ.

Compared to the previous time-series studies, we obtain relatively large magnitudes of both volatility-of-volatility parameter, σ_v , and leverage-effect parameter, ρ , for each square-root volatility model. For example, for the estimates of σ_v , Eraker et al. (2003) attain 0.14, 0.10, and 0.08 for SV, SVJ, and SVCJ, respectively, whereas our σ_v is estimated to be 0.16, 0.13, and 0.13, respectively. On the other hand, our estimates of ρ range from -0.52 to -0.61 , while those from Andersen et al. (2002) and Eraker et al. (2003) range between -0.4 and -0.5 . In fact, our estimates are closer in magnitude to the estimates from the previous option pricing studies (e.g., Bakshi et al., 1997; Pan, 2002; Eraker, 2004). For instance, for the SVJ specification, Bakshi et al. (1997) attain -0.57 and 0.15 for ρ and σ_v , respectively, by using options data alone. The estimates for ρ from the past option studies have normally ranged between -0.5 and -0.6 .

Interestingly, Bakshi et al. (1997) and Bates (2000) argue that the magnitudes for ρ and σ_v estimated from options data are too large to be consistent with time-series data. However, from time-series data, we obtain the magnitudes closer to those from option studies by choosing the sample period different from the existing time-series studies. For the past time-series studies, Andersen et al. (2002) and Eraker et al. (2003) include the early eighties in their sample periods. However, our sample period is similar to those used in option pricing studies: most sample periods from the option studies, like ours, begin from the late eighties, due to the availability of options data. Our result shows that it is possible that large magnitudes of ρ and σ_v in the previous options studies might have come from a

different sample period rather than the inconsistency between the spot and options markets.

Our jump component result is roughly consistent with the past time-series studies. The jump arrival intensity parameters, λ , are estimated to be 0.013 for the SVJ and 0.005 for the SVCJ, meaning that the numbers of jump occurrences per year are as small as 3.2 and 1.3, respectively. Additionally, we obtain negative mean jump sizes (μ_S) for both models (-1.6 for SVJ and -3.7 for SVCJ), which implies that, together with negative leverage-effect coefficient ρ from the diffusion part, the jump component can induce a conditional left skewness in returns. For the SVCJ, the estimate of average amplitude in jump-in-volatility, σ_v , is 1.6, which is close to 1.5 from Eraker et al. (2003). However, it should be noted that our jump component parameters are estimated with relatively large standard errors compared to diffusion parameters. As has been documented in the past time-series studies, it is difficult to precisely estimate the jump parameters because jump rarely occurs, and consequently, one would need a long sample period to improve the precision of jump parameter estimation.

As a by-product of particle filters, we can obtain filtered spot volatility paths for the SV, SVJ, and SVCJ, as shown by Figure 3.1.²³ This figure looks similar to Figure 1 in Eraker et al. (2003). What differs is that they obtain the paths from their MCMC posteriors, while we obtain them from particle filters. Our spot volatilities for the SV, SVJ, and SVCJ models are filtered by taking a sample average of simulated spot volatilities (i.e., particles for spot volatilities) at each date, which is equivalent to conditional expectation of spot volatility given a past history of returns. Therefore, our filtered spot volatilities are optimal in the sense of mean-squared error criterion. All four volatility paths, including GARCH-N, show

²³The GARCH-N volatility path is deterministic given past history of returns by its nature.

roughly similar historical trend over time. However, further investigation reveals that, compared to the GARCH-N path, the stochastic volatility paths appear more jagged over time, reflecting their Brownian feature. Among the square-root models' volatility paths, one can observe the occurrences of jump-in-volatility from the SVCJ's filtered volatilities. The particle filters provide further useful information. For instance, also following Eraker et al. (2003), we plot time-series of filtered jump probability and size of jump-in-return for the SVJ and SVCJ as shown in Figure 3.2, which are also by-products from the particle filters. It appears that more jumps are observed in SVJ than in SVCJ, which indicates that a substantial portion of jumps in SVJ are absorbed in a large increase in volatility in SVCJ. It is worth comparing our volatility filtering results with those from Chernov and Ghysels (2000). They filter model-implied volatilities for the SV model (without jump) via the EMM reprojection method, which has been proposed by Gallant and Tauchen (1998), for the period from November 1993 to October 1994. They filter two different spot volatilities for the SV: one is filtered by using historical returns alone and the other by using only options data. They also report the volatility paths for the BS (using options) and for the GARCH (using historical returns). Contrary to our result, their volatility paths show quite different trends over time across different models. Even the same SV models provide strikingly different patterns depending on the information they use (either returns or options). This comparison suggests that one should be cautious about filtering volatilities, since different methods could produce dramatically different outcomes. It seems that, in filtering volatilities, the EMM reprojection method is more sensitive to a model misspecification than the particle filters.

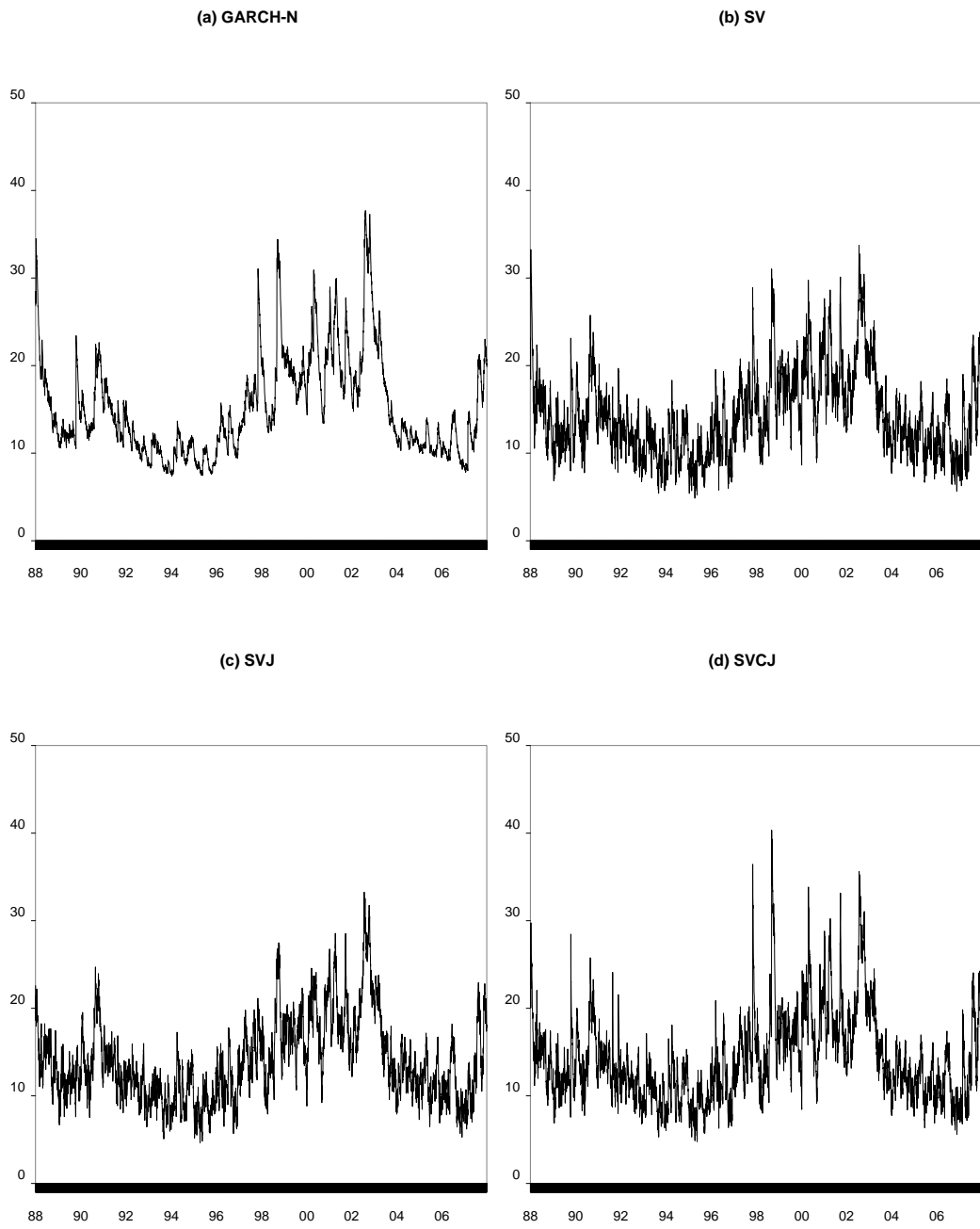


Figure 3.1: Filtered Model-implied Spot Volatilities

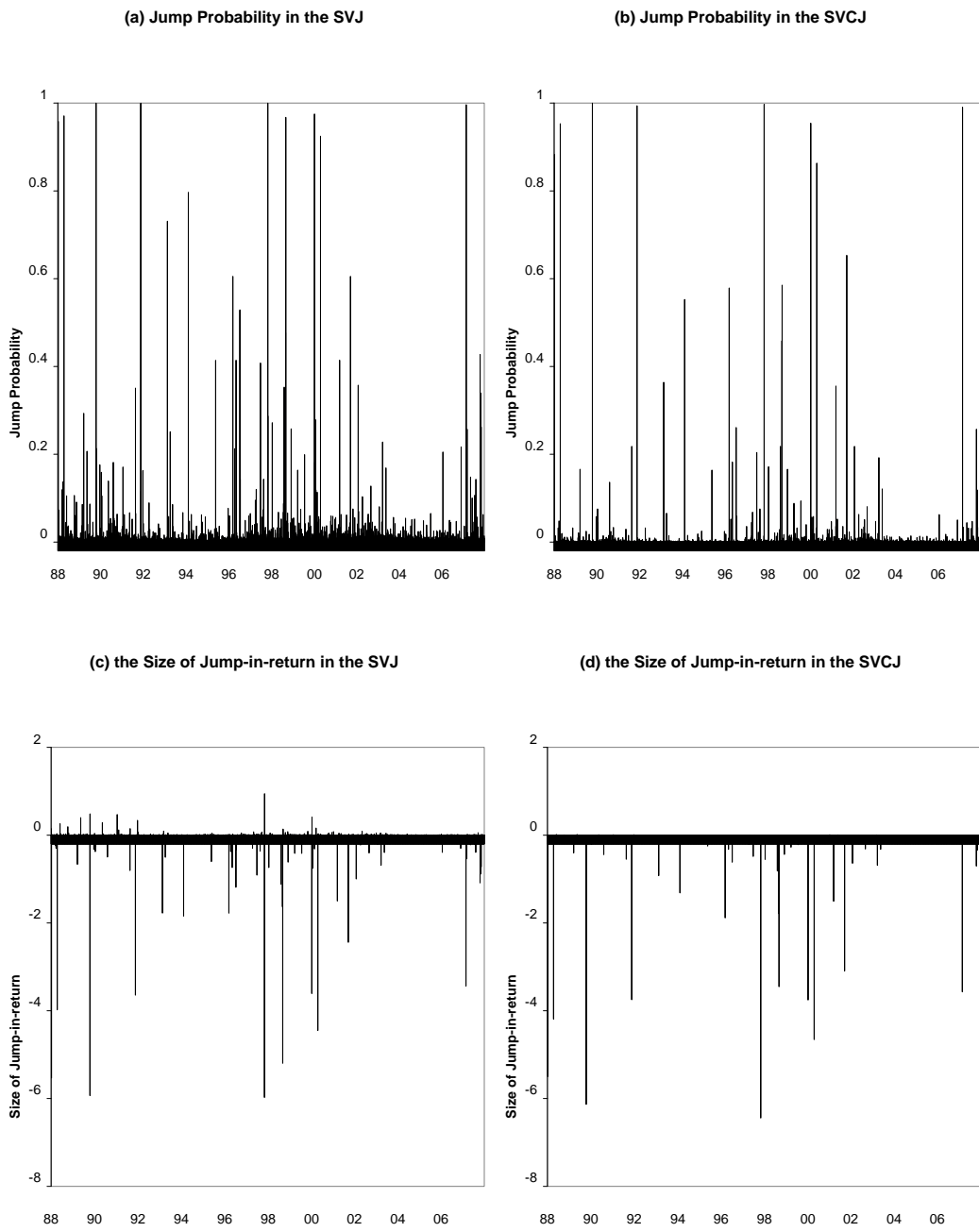


Figure 3.2: Filtered Jump Probability and Size of Jump-in-return for the SVJ and SVCJ

3.5 Model Performance Evaluations

In this section, we evaluate both in- and out-of-sample performances for stock return models. As noted above, we first evaluate each model from a statistical criterion, for which we employ both $\widehat{Q}(j)$ and $\widehat{W}(p)$ statistics introduced in Section 3.2. Besides the statistical criterion, in order to further investigate model performances from a practical viewpoint, we also compare value-at-risk (VaR) performances of the models.

3.5.1 In-Sample Performance

We now examine in-sample performances for all models considered. We first focus on the models with a similar structure and then compare the models across different structures. Figure 3.3 depicts both in-sample $\widehat{Q}(j)$ statistics and kernel marginal densities of generalized residuals for each model. The kernel marginal densities are useful for detecting a departure from stationary $U[0, 1]$ property. Additionally, Table 3.6 reports the corresponding $\widehat{W}(p)$ statistics for both in- and out-of-samples.

Table 3.6: In- and Out-of-sample Performance Evaluation Statistics $\widehat{W}(p)$

	In-sample			Out-of-sample		
	$p = 5$	10	20	$p = 5$	10	20
BS	118.4	164.1	228.2	56.7	78.3	109.4
PJ	23.8	28.7	35.0	49.8	67.8	91.5
GARCH-N	48.5	65.4	90.0	21.2	26.9	36.7
GARCH-T	9.2	8.2	8.3	18.5	22.4	30.2
EGARCH	37.2	49.4	67.6	17.9	22.9	31.8
GJR	41.2	55.1	75.9	17.5	22.0	30.8
LSV	14.6	16.7	21.1	12.2	14.2	18.8
LSV0	23.4	27.9	36.8	16.2	19.5	26.3
SV	11.0	12.9	15.9	11.4	13.2	17.6
SVJ	8.7	9.7	11.1	10.1	12.0	16.5
SVCJ	10.7	12.3	15.1	11.2	13.1	17.7

Note: The statistics are computed by $\widehat{W}(p) = \frac{1}{\sqrt{p}} \sum_{j=1}^p \widehat{Q}(j)$.

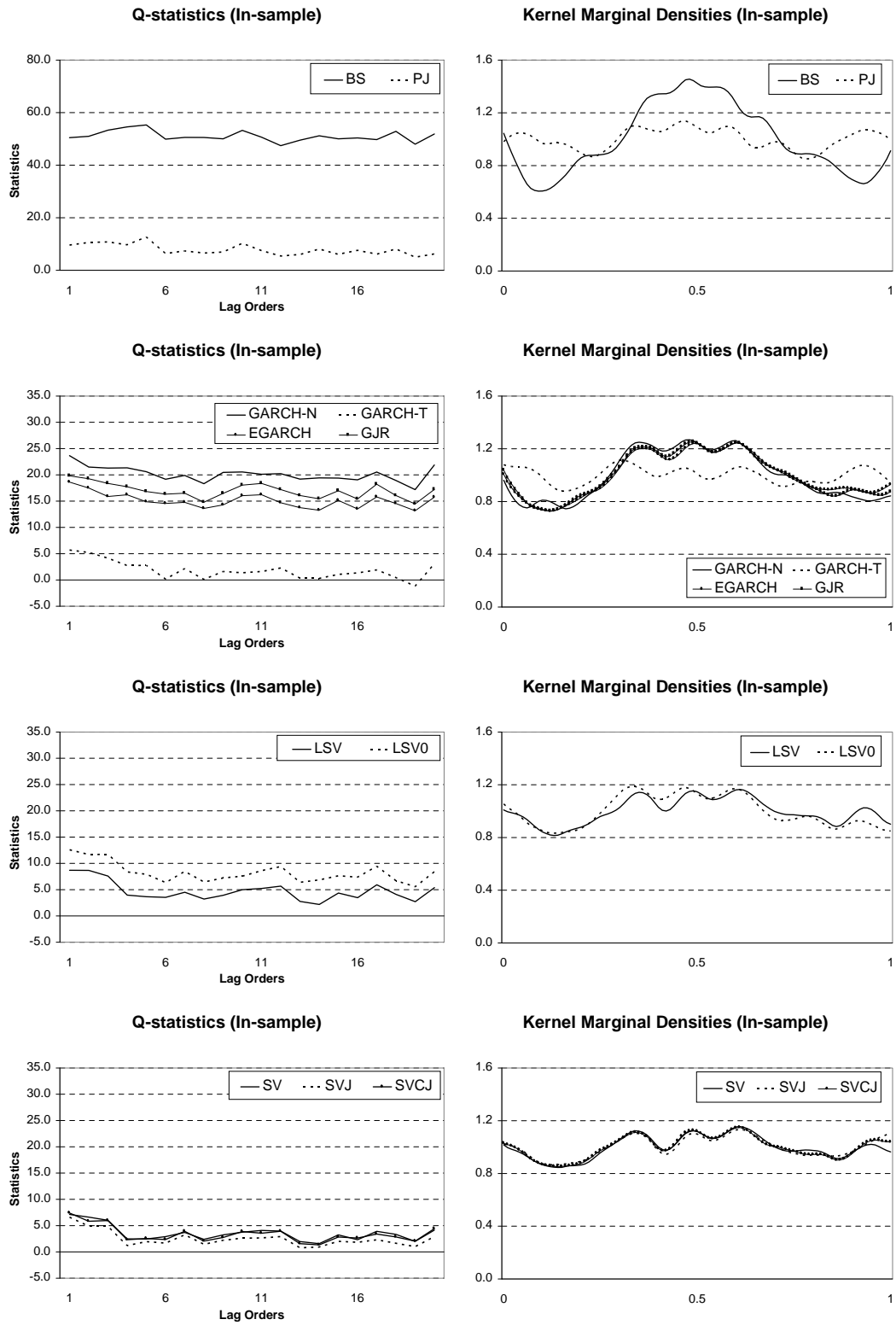


Figure 3.3: $\hat{Q}(j)$ Statistics and Kernel Marginal Densities for In-sample

First, we examine the importance of a pure jump component. The upper left panel in Figure 3.3 compares the BS and PJ models via $\widehat{Q}(j)$ statistic for each lag j . As has been evidenced by the previous studies, the BS model is clearly misspecified. Its $\widehat{Q}(j)$ statistics are approximately 50 at all lags, and $\widehat{W}(5)$ statistic is, consequently, as high as 118.4 (Table 3.6). The inclusion of jump component into the PJ improves substantially upon the BS. For each lag, the $\widehat{Q}(j)$ statistic falls by roughly 40 by help of a jump component.

The second panel in Figure 3.3 shows that GARCH-T outperforms the other GARCH models in terms of $\widehat{Q}(p)$ statistics, implying that it is important in the GARCH to model heavy tails of conditional distribution. The $\widehat{W}(p)$ statistics in Table 3.6 also show that GARCH-T improves, to a large extent, upon the GARCH-N, EGARCH, and GJR. For instance, the $\widehat{W}(5)$ statistic is 9.2 for GARCH-T, whereas the statistics are 48.9, 37.2, and 41.2 for GARCH-N, EGARCH, and GJR, respectively. Although the EGARCH and GJR improve upon the GARCH-N by help of modeling the leverage effect, the result is a worse performance than GARCH-T, at least for in-sample.

Turning to the log stochastic volatility models (i.e., LSV and LSV0) in the third panel of Figure 3.3, we find that modeling the leverage effect contributes a lot to the in-sample fit: for example, the $\widehat{W}(5)$ in Table 3.6 declines from 23.4 (LSV) to 14.6 (LSV0). In some log stochastic volatility studies such as Kim et al. (1998) and Chib et al. (2002), the leverage effect has been sometimes ignored in pursuit of a more efficient MCMC algorithm. Their scheme could possibly be justified for foreign exchange rates or interest rates. However, our result indicates that, at least for stock return models, modeling leverage effect is crucial.

Among the square-root stochastic volatility models, as shown by the fourth

panel of Figure 3.3, the SVJ modestly outperforms the SV and SVCJ. The $\widehat{Q}(j)$ statistics of SVJ are lower at all lags than those of SV and SVCJ. Specifically, the introduction of jump-in-return lowers the $\widehat{W}(5)$ from 11.0 (SV) to 8.7 (SVJ) (Table 3.6). As noted above, the importance of jump-in-return component has been supported by most time-series studies on stock return models (e.g., Andersen et al., 2001; Eraker et al., 2002; Chernov et al., 2003). However, our analysis shows that the incorporation of jump-in-volatility does not make an additional contribution to in-sample performance. The SVCJ seems to perform similarly to the simpler SV without jump. The previous studies have also shown mixed results on the role of jump-in-volatility.

Overall, unfortunately, it appears that all models considered seem to be misspecified from our statistical criterion, though the degrees of misspecification appear to be different across models. As reported in Table 3.6, all the $\widehat{W}(p)$ statistics are greater than eight, although 5% critical value of the normal distribution, which is a limiting distribution of the statistics under the null hypothesis, is 1.65. Our depressing result may be due to the fact that our method has a higher power against the alternatives than the existing methods. For instance, Andersen et al. (2001) find no evidence that the SVJ is misspecified in terms of the EMM criteria. Given our results, one might ask what could be a better model specification. Some studies such as Pan (2002), Chernov et al. (2003), and Eraker (2003) suggest the multifactor latent state variables or the state dependent jump intensity as possible sources that can help to improve the overall goodness-of-fit.

For the overall in-sample performances, it turns out that GARCH-T and SVJ models perform the best. In terms of $\widehat{W}(p)$, the SVJ (8.7) outperforms the GARCH-T (9.2) when $p = 5$. However, at the other lags, GARCH-T exhibits

the lower $\widehat{W}(p)$ than the SVJ. In the overall ranking, these two models are followed by the other square-root stochastic volatility models, such as SV and SVCJ, and then the LSV follows. It sounds rather surprising that GARCH-T performs as well as the stochastic volatility models. Kim et al. (1998) find that, in their foreign exchange rate study, the GARCH-T is slightly better than the SV (without jump) in terms of the marginal likelihood. Our in-sample analysis provides a similar result for stock return models. However, a care must be taken in the interpretation of in-sample evaluation results. A good in-sample performance might be due to so-called "data snooping" bias. The following out-of-sample study will further examine this possibility.

3.5.2 Out-of-Sample Performance

This subsection investigates out-of-sample performance for each model. As noted above, our testing method is very convenient for out-of-sample comparison of nonnested models. One might suspect an overparametrization (also called data snooping bias) if there is a substantial difference between in- and out-of-sample performances. Since it turns out that, from our criterion, all models are misspecified for in-sample fit, we should aim at seeking the least misspecified model which performs consistently well across both in- and out-of-sample periods.

The first panel in Figure 3.4 suggests a potential overfitting present in the PJ model. The distance between the BS and PJ is shorter for out-of-sample evaluation: $\widehat{W}(5)$ statistics in Table 3.6 are 49.8 for PJ and 56.7 for BS. As stated earlier, it seems that the jump intensity parameter, λ , is very unstably estimated, depending heavily on the sample period used.

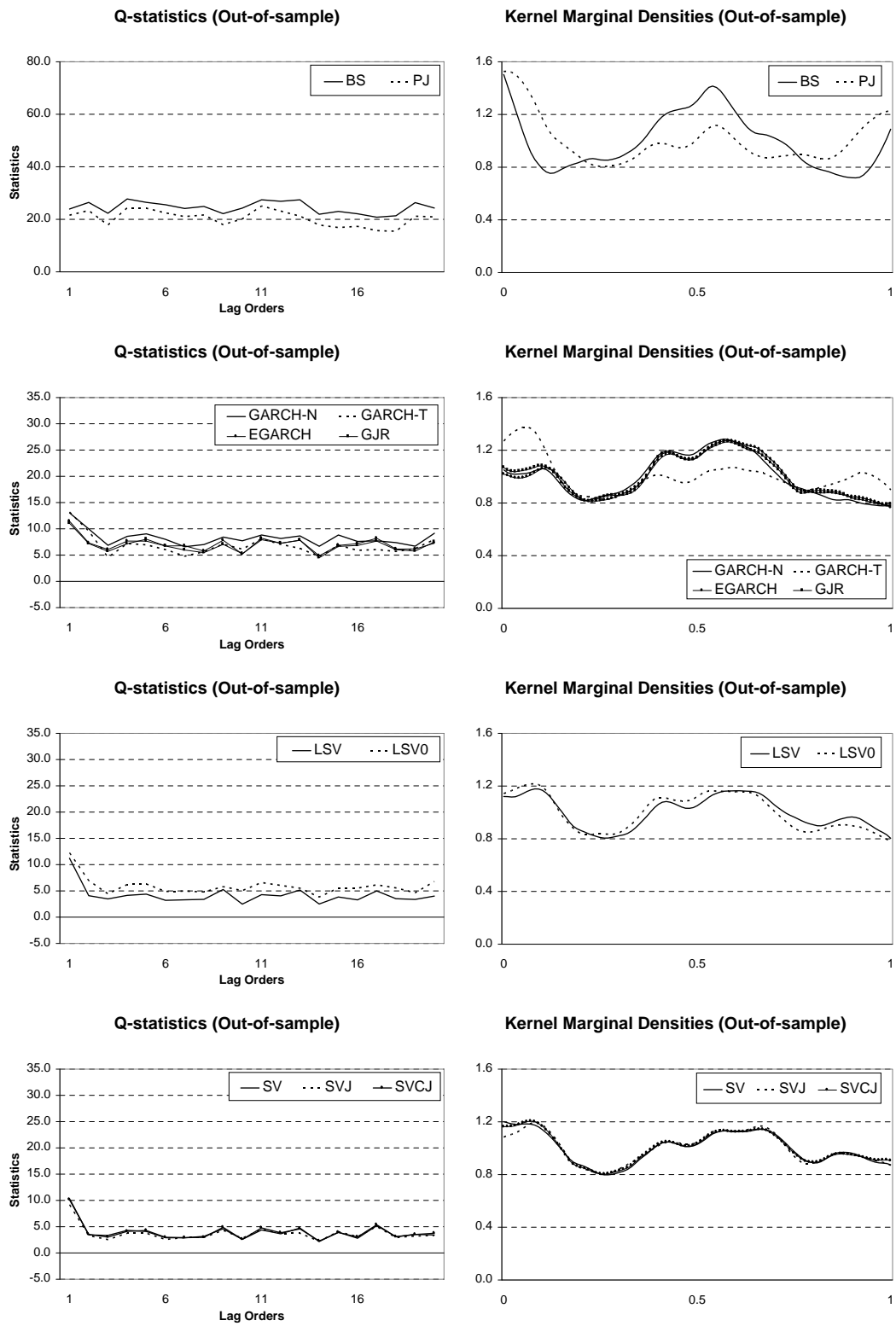


Figure 3.4: $\hat{Q}(j)$ Statistics and Kernel Marginal Densities for Out-of-sample

The performance of GARCH-T model, which has performed the best for in-sample evaluation, also becomes worse for out-of-sample evaluation. The $\widehat{W}(5)$ of GARCH-T (in Table 3.6) is 18.5, much closer to that of GARCH-N (21.2), and now GARCH-T performs slightly worse than EGARCH (17.9) and GJR (17.5). Also, GARCH-T is clearly dominated by stochastic volatility models for out-of-sample performance. This result might cast doubt that the degree of freedom parameter, ν , in GARCH-T might be overparametrized or possibly time-varying. The out-of-sample kernel marginal density of generalized residual in Figure 3.4 shows that GARCH-T underestimates the frequency of left-tail events, while its in-sample marginal density in Figure 3.3 has been close to $U[0, 1]$. This implies that the magnitude of ν is too large to be consistent with the out-of-sample data. It is interesting that the BS and GARCH-N models perform better for out-of-sample evaluation, while the performances of more sophisticated counterparts, which are PJ and GARCH-T, respectively, become deteriorated. This might also provide some evidence of the overfitting problem with the PJ and GARCH-T.

Contrary to the worse performances of the PJ and GARCH-T models, it is notable that stochastic volatility models exhibit stable performances across in- and out-of-sample periods. All the stochastic volatility models, except for the LSV0, perform relatively well for out-of-sample evaluation. Their out-of-sample $\widehat{W}(5)$ statistics in Table 3.6 range from 10.1 (SVJ) to 12.2 (LSV). In the log stochastic volatility models, it is notable that the LSV still outperforms the LSV0 considering that the unconditional skewness for out-of-sample period is positive (0.07) as reported in Table 3.1. Again, our out-of-sample analysis indicates that the leverage effect should not be ignored in modeling stock return stochastic volatility models. Among the square-root stochastic volatility models, there is little difference

across their out-of-sample performances, although the SVJ still provides slightly lower $\widehat{W}(p)$ statistics at all lags than the SV and SVCJ. Similar to the in-sample analysis, the jump-in-volatility (in SVCJ) seems to contribute little to the model performance. In the option pricing study, Bakshi et al. (1997) have argued that the introduction of stochastic volatility term has the first order effect in the hedging performance for option contracts. Eraker (2004) also finds that the jump in return or volatility leads to only a small improvement in option pricing. Our result is consistent with their findings in the time-series context. The stochastic volatility term seems to have the first order effect in terms of density forecast. It might be due to the fact that jumps occur very rarely, so it does play a minor role in shaping a conditional density of stock returns.

3.5.3 Value-at-risk Performance

To further investigate model performances from a practical viewpoint, we compare Value-at-Risk (VaR) performances of the models considered in this study. As noted above, stochastic volatility models are applicable to the VaR implementation through the particle filters. We employ two testing methods to evaluate the VaR performances: the Kupiec (1995) test and the dynamic quantile test proposed by Engle and Manganelli (1999). Before presenting our testing results, we briefly describe each testing method.

First, let us outline the likelihood ratio test proposed by Kupiec (1995). Let N be the number of times the portfolio loss is worse than the true VaR in a sample size T . Then the number of VaR exceptions, N , follows a binomial distribution, that is, $N \sim B(T, p)$, where p is a significance level implied by VaR measure. Ideally, the ratio, N/T , should be very close to the significant level of VaR measure (e.g. 1% or 5%). Hence, the relevant null and alternative hypotheses are $H_0 : N/T = p$

and $H_A : N/T \neq p$, respectively. Kupiec shows that the appropriate likelihood ratio statistic is

$$LR = 2 [\log ((N/T)^N (1 - N/T)^{T-N}) - \log (p^N (1 - p)^{T-N})]. \quad (3.10)$$

This likelihood ratio is asymptotically chi-square distributed with degree of freedom 1 under H_0 . This testing method is very similar to what supervisory authorities require of large banks for backtesting the validity of their internal VaR system. However, this test, though simple, focuses only on the frequency of exceptions, that is, the unconditional mean of the ratio, N/T , but ignores dynamics of the occurrences of exceptions. Note that, if a model is a true VaR model, the occurrences of exception should be independently distributed across time. A clustering of the occurrences reveals that the model might be inadequate for VaR. However, the Kupiec method is silent about this dynamic aspect.

In order to handle this problem, Engle and Manganelli (1999) propose an omnibus test which can examine both unconditional mean and i.i.d. property simultaneously. First, define a random variable "Hit" as follows.

$$Hit_t \equiv I(y_t < -VaR_t) - p, \quad (3.11)$$

where $I(\cdot)$ is an indicator function, y_t is a realized portfolio return at time t , VaR_t is a VaR measure calculated by using the information up to time $t-1$, and p is a significance level for VaR. Note that Hit_t should be independent with any random variable contained in the information set at time $t-1$, I_{t-1} . Then consider the following OLS regression.

$$Hit_t = \delta_0 + \delta_1 Hit_{t-1} + \dots + \delta_l Hit_{t-l} + \delta_{l+1} VaR_t + u_t, \quad (3.12)$$

where

$$u_t \sim \begin{cases} -p & \text{with probability } (1 - p) \\ (1 - p) & \text{with probability } p. \end{cases}$$

It is easy to see that all the OLS coefficients should be jointly zero when a model is true. From the above OLS outcome, Engle and Manganelli (1999) suggest the following test statistic called the dynamic quantile (DQ) test statistic,

$$DQ = \frac{\delta'_{OLS}(X'X)^{-1}\delta_{OLS}}{p(1-p)}, \quad (3.13)$$

where X represents explanatory variables for (3.12) in matrix form, and δ_{OLS} is the corresponding OLS estimator for the coefficient vector, $(\delta_0, \delta_1, \dots, \delta_l, \delta_{l+1})$. Engle and Manganelli (1999) show that the resulting statistic converges to the chi-square distribution with the degree of freedom, $l + 2$, under the null hypothesis that the VaR model is correct. In our analysis, we set l equal to five, and, instead of VaR_t in Equation (3.12), we use the conditional variance inferred from information at time $t - 1$, which has very similar information to VaR_t . In the case of the SV class of models, the particle-filtered stochastic volatility estimates are used.

Table 3.7: The Kupiec test statistics

	In-sample		Out-of-sample	
	1% VaR	5% VaR	1% VaR	5% VaR
BS	5.453 (0.020)	0.240 (0.624)	20.387 (0.000)	7.073 (0.008)
PJ	1.776 (0.183)	0.093 (0.761)	0.050 (0.822)	11.262 (0.001)
GARCH-N	6.886 (0.009)	0.956 (0.328)	1.550 (0.213)	0.000 (0.988)
GARCH-T	0.339 (0.561)	0.263 (0.608)	0.063 (0.802)	2.384 (0.123)
EGARCH	3.387 (0.066)	0.581 (0.446)	0.275 (0.600)	0.023 (0.880)
GJR	3.765 (0.052)	0.148 (0.700)	0.451 (0.502)	0.133 (0.715)
LSV	0.478 (0.489)	0.040 (0.842)	0.008 (0.929)	0.196 (0.658)
LSV0	3.030 (0.082)	0.064 (0.801)	0.142 (0.706)	0.138 (0.710)
SV	0.824 (0.364)	0.001 (0.992)	0.005 (0.945)	2.181 (0.140)
SVJ	0.003 (0.985)	0.009 (0.925)	0.172 (0.678)	0.427 (0.514)
SVCJ	0.061 (0.905)	0.022 (0.883)	0.340 (0.560)	0.628 (0.428)

Note: P-values are in parentheses.

Table 3.8: The Dynamic Quantile (DQ) test statistics

	In-sample		Out-of-sample	
	1% VaR	5% VaR	1% VaR	5% VaR
BS	36.097 (0.000)	44.367 (0.000)	206.457 (0.000)	114.682 (0.000)
PJ	14.697 (0.023)	47.328 (0.000)	10.313 (0.112)	129.072 (0.000)
GARCH-N	58.353 (0.000)	16.260 (0.023)	61.226 (0.000)	15.150 (0.034)
GARCH-T	6.493 (0.484)	14.820 (0.038)	27.848 (0.000)	20.218 (0.005)
EGARCH	12.183 (0.095)	12.825 (0.076)	19.377 (0.007)	5.559 (0.592)
GJR	17.857 (0.013)	9.945 (0.192)	7.692 (0.361)	3.774 (0.805)
LSV	5.711 (0.574)	11.816 (0.107)	10.033 (0.187)	8.665 (0.278)
LSV0	15.362 (0.032)	17.589 (0.014)	6.295 (0.506)	6.518 (0.481)
SV	6.011 (0.538)	8.769 (0.270)	4.792 (0.685)	11.913 (0.103)
SVJ	4.223 (0.754)	10.910 (0.143)	6.013 (0.538)	5.690 (0.576)
SVCJ	3.688 (0.815)	6.935 (0.436)	6.846 (0.445)	7.506 (0.378)

Note: P-values are in parentheses.

Table 3.7 and 3.8 report the Kupiec and the DQ test statistics for both in- and out-of-sample evaluations. As can be expected, the Kupiec test seems to have a low power against the alternative, since it is silent about a dynamic feature. Most models except for the BS and PJ pass the tests for both 1% and 5% VaR cases. One more interesting exception occurs to the GARCH-N for the in-sample 1% VaR test. It produces a too high exception ratio (1.8%, not reported in Table 3.7).

On the other hand, it turns out that many more models fail to satisfy the DQ criterion. Similar to the results from Kupiec test, the BS and PJ also fail to meet the criterion due to their underestimation of true left-tail risk in the "unconditional mean" sense. However, some other models get rejected due to their non-i.i.d. features. In particular, both GARCH-N and GARCH-T models turn out to exhibit a substantial autocorrelation in their "hit" variables between different lags, producing a large magnitude of the DQ statistics. Similarly, the EGARCH and GJR models also fail to satisfy the criterion at some slots in Table 3.8. The interesting finding is that the stochastic volatility-class models, regardless of their

incorporation of jump component, all succeed in passing our VaR adequacy tests under both testing methods. The only exception is the LSV0 model, into which the leverage effect is not modeled. This finding gives some practical implication that stochastic volatility models can be considered as an alternative to the GARCH class of models in implementing VaR.

3.6 Conclusion

In this study, we have proposed a simulation-based specification testing method to be able to evaluate the density forecasting performance of stochastic volatility models, which involve a latent state variable process. Our testing method is based on Hong and Li (2005) and Johannes et al. (2008). Basically, we have extended Hong and Li's (2005) nonparametric specification testing method to be applicable to the stochastic volatility models. The idea is to approximate a dynamic probability integral transform, which is a key step in Hong and Li's (2005) original test via the Monte Carlo integration. To conduct the Monte Carlo integration, we use the particle filtering method proposed by Johannes et al. (2008). Our testing method enables us to compare famous GARCH and stochastic volatility models, which have seldom been pursued in the literature partly due to the fact that there are only a few specification testing tools which can evaluate both classes.

By using both Hong and Li's (2005) original and our own proposed methods, we have conducted a comprehensive empirical study on the stock return models using S&P500 stock index returns, from both in- and out-of-sample context. Our empirical analysis shows that all models considered are, unfortunately, misspecified in terms of density forecast. However, the stochastic volatility models perform relatively well in both in- and out-of-sample. It seems that the stochastic volatility component, rather than jump, appears to have the first order effect in capturing

the stock return dynamics, which reflects rare occurrences of jump events in reality. Also, we find that modeling the leverage effect provides a substantial improvement in the log stochastic volatility models. Some past studies have ignored the leverage effect in pursuit of computational efficiency in the MCMC algorithm. Our result supports the importance of leverage effect, at least for stock return models. On the other hand, although one GARCH model (i.e., GARCH with t-distributed innovation) performs as well as stochastic volatility models for in-sample evaluation, its performance becomes worse in out-of-sample. Besides the statistical test, we have also evaluated the VaR performance for each model from an economic viewpoint. We find that the stochastic volatility models outperform the GARCH models, implying that stochastic volatility models can be a possible alternative to the widely used GARCH models in the VaR implementation.

Focusing on time-series dynamics, we did not investigate a risk-neutral aspect of model specification. To the best of our knowledge, empirical option pricing studies have usually been conducted inside the stochastic volatility class alone. The option pricing performances for both classes of models will be examined in Chapter 4.

APPENDICES

Appendix 3.1: Dynamic Probability Integral Transform via Particle Filtering for the SVCJ Model

This appendix describes the dynamic probability integral transform procedure via particle filtering for the SVCJ model. Our method is based on Johannes et al.'s (2008) optimal filtering method and Pitt and Sherphard's (1999) APF algorithm. The algorithm for the SV and SVJ is a special case when $\lambda = 0$ and when $Z_t^V = 0$, respectively. The procedure is very similar to the procedure for the LSV. However, the incorporation of jump components further complicates the algorithm. We consider the following Euler-discretized and Bernoulli-approximated SVCJ model:

$$\Delta \ln S_t = \alpha + \sqrt{V_t} \varepsilon_t + J_t Z_t^S, \quad (3.14)$$

$$\begin{aligned} V_t &= V_{t-1} + \kappa(\theta - V_{t-1}) + \rho\sigma_v(\Delta \ln S_{t-1} - \alpha - J_{t-1}Z_{t-1}^S) + \\ &\quad \sqrt{1 - \rho^2}\sigma_v\sqrt{V_{t-1}}\eta_t + J_{t-1}Z_t^v, \end{aligned} \quad (3.15)$$

where $(\varepsilon_t \ \eta_t)' \sim \text{i.i.d. } N(0, I_2)$, $J_t \sim \text{i.i.d. } \text{Ber}(\lambda)$, $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$, and $Z_t^v \sim \text{i.i.d. } \exp(\mu_v)$. A latent state variable vector at time t , $(V_t, J_t, Z_t^S, Z_t^v)'$, is denoted by L_t , and also $\Delta \ln S_t$ by y_t . Now we take the following steps recursively.

- **Step 1:** Suppose that we are given N particles at time $t-1$, $\{L_{t-1}^{(i)} \mid y^{t-1}\}_{i=1}^N = \{(V_{t-1}^{(i)}, J_{t-1}^{(i)}, Z_{t-1}^{S,(i)}, Z_{t-1}^{v,(i)})' \mid y^{t-1}\}_{i=1}^N$ which have been drawn from the conditional distribution of latent state vector given a past history of returns, i.e. y^{t-1} . For each i , simulate $\{\eta_t^{(i)}, Z_t^{v,(i)}\}_{i=1}^N$ ²⁴ from their corresponding distributions, then we obtain $\{V_t^{(i)}\}_{i=1}^N$ via Equation (3.15).

²⁴If $J_{t-1}^{(i)} = 0$ for some i , then there is no need to simulate $Z_t^{v,(i)}$. Simply, $Z_t^{v,(i)}$ can be taken to be zero for such i 's.

With $\{V_t^{(i)}\}_{i=1}^N$, we approximate the generalized residual at time t :

$$\begin{aligned}
z_t &= \int_{-\infty}^{y_t} p(y | y^{t-1}) dy \\
&= \int_{-\infty}^{y_t} \int_0^{\infty} \int p(y | V_t) p(V_t | L_{t-1}, y_{t-1}) p(L_{t-1} | y^{t-1}) dL_{t-1} dV_t dy \\
&\simeq \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} p(y | V_t^{(i)}) dy \\
&= \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} p(J_t = 1) p(y | J_t = 1, V_t^{(i)}) + p(J_t = 0) p(y | J_t = 0, V_t^{(i)}) dy \\
&= \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} \lambda \phi(y | \alpha + \mu_S, V_t^{(i)} + \sigma_S^2) + (1 - \lambda) \phi(y | \alpha, V_t^{(i)}) dy \quad (3.16)
\end{aligned}$$

where $\phi(\cdot | \mu, \sigma^2)$ denotes a normal distribution density function with mean μ and variance σ^2 . After computing the generalized residual at time t , the simulated latent variables, $\{Z_t^{S,(i)}, V_t^{(i)}\}_{i=1}^N$, are discarded for the same reason discussed in Section 4.2.

- **Step 2:** This step begins to update the particles from $\{L_{t-1}^{(i)} | y^{t-1}\}_{i=1}^N$ to $\{L_t^{(i)} | y^t\}_{i=1}^N$. First, compute an auxiliary variable, $\widehat{V}_t^{(i)}$ for each i , which is a conditional expectation of volatility at time t , given $V_{t-1}^{(i)}$. In other words, that is

$$\widehat{V}_t^{(i)} = V_{t-1}^{(i)} + \kappa(\theta - V_{t-1}^{(i)}) + \rho\sigma_v(y_{t-1} - \alpha - J_{t-1}^{(i)} Z_{t-1}^{S,(i)}) + J_{t-1}^{(i)} \mu_v.$$

Using $\widehat{V}_t^{(i)}$, we evaluate the first stage weighting for each i :

$$w_t^{(i)} \propto (1 - \lambda) \times \phi(y_t | \alpha, \widehat{V}_t^{(i)}) + \lambda \times \phi(y_t | \alpha + \mu_S, \widehat{V}_t^{(i)} + \sigma_S^2).$$

With the weight assigned to each i , we resample the existing particles, and obtain the resampled particles, $\{L_{t-1}^{k(i)}\}_{i=1}^N = \{(V_{t-1}^{k(i)}, J_{t-1}^{k(i)}, Z_{t-1}^{S,k(i)}, Z_{t-1}^{v,k(i)})\}_{i=1}^N$.

- **Step 3:** For each $k(i)$, draw an occurrence of jump-in-return, $J_t^{k(i)}$ from the

binomial distribution with each probability,

$$\begin{aligned} p(J_t^{k(i)} = 1 \mid \widehat{V}_t^{k(i)}, y_t) &\propto \lambda \times p(y_t \mid \widehat{V}_t^{k(i)}, J_t^{k(i)} = 1) \\ &= \lambda \times \phi(y_t \mid \alpha + \mu_S, \widehat{V}_t^{k(i)} + \sigma_S^2) \\ p(J_t^{k(i)} = 0 \mid \widehat{V}_t^{k(i)}, y_t) &\propto (1 - \lambda) \times p(y_t \mid \widehat{V}_t^{k(i)}, J_t^{k(i)} = 0) \\ &= (1 - \lambda) \times \phi(y_t \mid \alpha, \widehat{V}_t^{k(i)}) \end{aligned}$$

Therefore,

$$\begin{aligned} p(J_t^{k(i)} = 1 \mid \widehat{V}_t^{k(i)}, y_t) &= \frac{\lambda \times \phi(y_t \mid \alpha + \mu_S, \widehat{V}_t^{k(i)} + \sigma_S^2)}{\lambda \times \phi(y_t \mid \alpha + \mu_S, \widehat{V}_t^{k(i)} + \sigma_S^2) + (1 - \lambda) \times \phi(y_t \mid \alpha, \widehat{V}_t^{k(i)})}, \end{aligned}$$

and $p(J_t^{k(i)} = 0 \mid \widehat{V}_t^{k(i)}, y_t) = 1 - p(J_t^{k(i)} = 1 \mid \widehat{V}_t^{k(i)}, y_t)$.

- **Step 4:** Draw a jump-in-return size for each $k(i)$ for which $J_t^{k(i)} = 1$. We draw it from normal distribution with mean, $\mu_Z^{S,k(i)}$ and variance, $V_Z^{S,k(i)}$ such that

$$\begin{aligned} \mu_Z^{S,k(i)} &= \sigma_Z^{2,k(i)} \left(\frac{y_t}{\widehat{V}_t^{k(i)}} + \frac{\mu_S}{\sigma_S^2} \right), \\ V_Z^{S,k(i)} &= \left(\frac{1}{\widehat{V}_t^{k(i)}} + \frac{1}{\sigma_S^2} \right)^{-1}. \end{aligned}$$

- **Step 5:** Draw a volatility innovation, $\eta_t^{k(i)}$ from i.i.d. $N(0, 1)$ for each $k(i)$ and draw a jump-in-volatility $Z_t^{v,k(i)}$ from i.i.d. $\exp(\mu_v)$ for each $k(i)$ for which $J_{t-1}^{k(i)} = 1$:

$$\begin{aligned} V_t^{k(i)} &= V_{t-1}^{k(i)} + \kappa(\theta - V_{t-1}^{k(i)}) + \sigma_v \rho (y_{t-1} - \mu - J_{t-1}^{k(i)} Z_{t-1}^{S,k(i)}) + \\ &\quad \sigma_v \sqrt{1 - \rho^2} \sqrt{V_{t-1}^{k(i)}} \eta_t^{k(i)} + J_{t-1}^{k(i)} Z_t^{v,k(i)}. \end{aligned}$$

After this step, we have the particles,

$$\{L_t^{k(i)}\}_{i=1}^N = \{(V_t^{k(i)}, J_t^{k(i)}, Z_t^{S,k(i)}, Z_t^{v,k(i)})\}_{i=1}^N.$$

- **Step 6:** As a final step, resample the particles $L_t^{k(i)}$ with the second stage weights defined as

$$\pi_t^{k(i)} \propto \frac{\lambda \phi(Z_t^{S,k(i)} | \mu_Z^{S,k(i)}, \sigma_Z^{2,k(i)}) \phi(y_t - J_t^{k(i)} Z_t^{S,k(i)} | \alpha, V_t^{k(i)})}{p(J_t^{k(i)} = 1 | \widehat{V}_t^{k(i)}, y_t) \phi(Z_t^{S,k(i)} | \mu_S, \sigma_S^2)}$$

when $J_t^{k(i)} = 1$, and

$$\pi_t^{k(i)} \propto \frac{(1 - \lambda) \phi(y_t - J_t^{k(i)} Z_t^{k(i)} | \alpha, V_t^{k(i)})}{p(J_t^{k(i)} = 0 | \widehat{V}_t^{k(i)}, y_t)}$$

when $J_t^{k(i)} = 0$. Then we obtain the updated particles, $\{L_t^{(i)} | y^t\}_{i=1}^N = \{(V_t^{(i)}, J_t^{(i)}, Z_t^{S,(i)}, Z_t^{v,(i)})' | y^t\}_{i=1}^N$. The particles after the 2nd stage reweighting are approximately same as a sample drawn from $p(L_t | y^t)$.

- **Step 7:** Go back to Step 1, and compute the generalized residual at time $t + 1$.

REFERENCES

- [1] Ahn, D., and B. Gao, 1999, “A Parametric Nonlinear Model of Term Structure Dynamics,” *Review of Financial Studies*, 12, 721-762.
- [2] Ait-Sahalia, Y., 1996, “Testing Continuous-Time Models of the Spot Interest Rate,” *Review of Financial Studies*, 9, 385-426.
- [3] Ait-Sahalia, Y., 2002a, “Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach,” *Econometrica*, 70, 223-262.
- [4] Ait-Sahalia, Y., 2002b, “Closed-Form Expansion for Multivariate Diffusions,” working paper, Princeton University.
- [5] Akgiray, V., and G. Booth, 1986, “Stock Price Processes with Discontinuous Time Paths: An Empirical Examination,” *The Financial Review*, 21, 163-184.
- [6] Andersen, T., L. Benzoni, and J. Lund, 2002, “An Empirical Investigation of Continuous-Time Equity Return Models,” *Journal of Finance*, 57, 1239-1284.
- [7] Andersen, T., H. Chung, and B. Sorensen, 1999, “Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study,” *Journal of Econometrics*, 91, 61-87.
- [8] Bakshi, G., C. Cao, and Z. Chen, 1997, “Empirical Performance of Alternative Option Pricing Models,” *Journal of Finance*, 52, 2003-2049.
- [9] Ball, C.A., and W.T. Torous, 1985, “On Jumps in Common Stock Prices and their Impact on Call Option Pricing,” *Journal of Finance*, 40, 155-173.

- [10] Bates, D., 2000, "Post-'87 Crash Fears in the S&P500 Futures Option Market," *Journal of Econometrics*, 94, 181-238.
- [11] Benzoni, L., 1998, "Pricing Options under Stochastic Volatility: an Econometric Analysis," Working Paper, Northwestern University.
- [12] Berg, A., R. Meyer, and J. Yu, 2004, "Deviance information Criterion for Comparing Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 22, 107-120.
- [13] Bhardwaj, G., V. Corradi, V., and N.R. Swanson, 2007, "A Simulation Based Specification Test for Diffusion Processes," *Journal of Business and Economic Statistics*, forthcoming.
- [14] Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- [15] Bollerslev, T., and H.O. Mikkelsen, 1996, "Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 73, 151-184.
- [16] Brandt, M., and P. Santa-Clara, 2002, "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets," *Journal of Financial Economics*, 63, 161-210.
- [17] Broadie, M., M. Chernov, and M. Johannes, 2007, "Model Specification and Risk Premia: The Evidence from Futures Options," *Journal of Finance*, 62, 1453-1490.

- [18] Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209-1227.
- [19] Chen, S.X. and J. Gao, 2005, "On the Use of the Kernel Method for Specification Tests of Diffusion Models," Working Paper, University of Western Australia.
- [20] Chen, S.X. and J. Gao, 2007, "An Adaptive Empirical Likelihood Test for Time Series Models," *Journal of Econometrics*, forthcoming.
- [21] Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen, 2003, "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225-257.
- [22] Chernov, M. and E. Ghysels, 2000, "Estimating of Stochastic Volatility Models for the Purpose of Option Pricing," *Computational Finance* 1999.
- [23] Chernov, M. and E. Ghysels, 2000, "A Study towards a Unified Approach to the Joint Estimation of the Objective and Risk Neutral Measures for the Purpose of Option Valuation," *Journal of Financial Economics*, 56, 407-458.
- [24] Chib, S., 1995, "Marginal Likelihood from a Gibbs Output," *Journal of the American Statistical Association*, 90, 1313-1321.
- [25] Clark, P.K., 1973, "A subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135-155.
- [26] Corradi, V., 2000, "Reconsidering the Continuous-Time Limit of the GARCH(1,1) Process," *Journal of Econometrics*, 96, 145-153.
- [27] Corradi, V. and N.R. Swanson, 2005, "A Bootstrap Specification Test for Diffusion Processes," *Journal of Econometrics*, 124, 117-148.

- [28] Corradi, V. and N.R. Swanson, 2006, "Simulation Based Predictive Density Estimation and Testing for Diffusion Processes," Working Paper, Rutgers University and Warwick University.
- [29] Cox, J.C., J.E. Ingersoll, and S.A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.
- [30] Das, S.R., 2002, "The Surprise Element: Jump in Interest Rate," *Journal of Econometrics*, 106, 27-65.
- [31] Das, S.R., and R. Uppal, 1998, "International Portfolio Choice with Systemic Risk," Working Paper, Harvard Business School and MIT.
- [32] Diebold, F.X., T. Gunther, and A. Tay, 1998, "Evaluating Density Forecasts with Application to Financial Risk Management," *International Economic Review*, 39, 863-883.
- [33] Duan, J., 1995, "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13-32.
- [34] Duffie, D., J. Pan, and K. Singleton, 2000, "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343-1376.
- [35] Elerian, O., S. Chib, and N. Shephard, 2001, "Likelihood Inference for Discretely Observed Nonlinear Diffusions," *Econometrica*, 69, 959-993.
- [36] Engle, R., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.
- [37] Engle, R., and S. Manganelli, 1999, "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," Working Paper, UCSD.

- [38] Eraker, B., 2004, "Do Stock Prices and Volatility Jump? Reconciling Evidences from Spot and Option Prices," *Journal of Finance*, 59, 1367-1403.
- [39] Eraker, B., M. Johannes, and N. Polson, 2003, "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 53, 1269-1300.
- [40] Gallant, A.R., and G. Tauchen, 1996, "Which Moments to Match?," *Econometric Theory*, 12, 657-681.
- [41] Gallant, A.R., and G. Tauchen, 1998, "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions," *Journal of American Statistical Association*, 93, 10-24.
- [42] Glosten, L., R. Jagannathan, and D. Runkle, 1993, "Relationship between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779-1801.
- [43] Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-344.
- [44] Hong, Y., and H. Li, 2005, "Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37-84.
- [45] Hong, Y., H. Li and F. Zhao, 2004, "Out-of-Sample Performance of Discrete-Time Spot Interest Rate Models," *Journal of Business & Economic Statistics*, 22, 457-473.

- [46] Hong, Y., H. Li and F. Zhao, 2007, "Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates," *Journal of Econometrics*, 141, 736-776.
- [47] Hull, J., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 52, 281-300.
- [48] Jacquier, E., N. Polson, and P. Rossi, 1994, "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economics Statistics*, 12, 70-87.
- [49] Jarrow, R., and E. Rosenfield, 1984, "Jump Risks and the Intertemporal Capital Asset Pricing Model," *Journal of Business*, 57, 337-351.
- [50] Johannes, M., and N. Polson, 2005, "MCMC for Financial Econometrics," in Y. Ait-Sahalia and L. Hansen (Eds.), *Handbook of Financial Econometrics*, Amsterdam: Elsevier-North-Holland.
- [51] Johannes, M., N. Polson, and J. Stroud, 2008, "Optimal Filtering of Jump-Diffusions: Extracting Latent States from Asset Prices," *Review of Financial Studies*, forthcoming.
- [52] Kallsen, J., and M.S. Taqqu, 1998, "Option Pricing in ARCH-type models," *Mathematical Finance*, 8, 13-26.
- [53] Kim, S., N. Shephard, and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economics Studies*, 65, 361-394.
- [54] Kupiec, P. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, 2, 73-84.

- [55] Li, F., and G. Tkacz, 2006, “A Consistent Bootstrap Test for Conditional Density Functions with Time-Series Data,” *Journal of Econometrics*, 92, 101-147.
- [56] Li, F., 2007, “Testing the Parametric Specification of the Diffusion Function in a Diffusion Process,” *Econometric Theory*, 23, 221-250.
- [57] Melino, A., and Turnbull, 1990, “Pricing Foreign Currency Options with Stochastic Volatility,” *Journal of Econometrics*, 45, 239-265.
- [58] Merton, R.C., 1976, “Option Pricing when Underlying Stock Returns are discontinuous,” *Journal of Financial Economics*, 3, 125-144.
- [59] Meyer, R., and J. Yu, 2000, “BUGS for a Bayesian Analysis of Stochastic Volatility Models,” *Econometrics Journal*, 3, 198-215.
- [60] Nelson, D., 1990, “ARCH Models as Diffusion Approximation,” *Journal of Econometrics*, 45, 7-39.
- [61] Nelson, D., 1991, “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 59, 703-708.
- [62] Pan, J., 2002, “The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study,” *Journal of Financial Economics*, 63, 3-50.
- [63] Pedersen, A.R., 1994, “Uniform Residuals for Discretely Observed Diffusion Processes,” *Technical Report 292*, Department of Statistics, Institute of Mathematics, University of Aarhus.
- [64] Pedersen, A.R., 1995, “A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete observations,” *Scandinavian Journal of Statistics*, 22, 55-71.

- [65] Pitt, M., and Shephard, N., 1999, "Filtering via simulation: Auxiliary Particle Filter," *Journal of American Statistical Association*, 94, 590-599.
- [66] Press, S.J., 1967, "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317-335.
- [67] Rosenberg, B., 1972, "The Behavior of Random Variables with Nonstationary Variance and the Distribution of Security Prices," University of California, Berkeley.
- [68] Rosenblatt, M., 1952, "Remarks on a multivariate transformation," *Annals of Mathematical Statistics*, 23, 470-472.
- [69] Scott, D.W., 1992, *Multivariate Density Estimation, Theory, Practice, and Visualization*, John Wiley and Sons, New York.
- [70] Singleton, K., 2006, *Empirical Dynamic Asset Pricing*, Princeton University Press, Princeton.
- [71] Taylor, S., 1982, "Financial Returns Modelled by the Product of Two Stochastic Processes - a Study of Daily Sugar Prices 1961-79," In Anderson, O.D., (ed.), *Time Series Analysis: Theory and Practice. 1* 203-26, Amsterdam: North-Holland.
- [72] Taylor, S., 1986, *Modelling Financial Time Series*, John Wiley and Sons, London.
- [73] Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177-188.
- [74] Yu, J., 2005, "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics*, 127, 165-178.

Chapter 4

Option Pricing and Density Forecast Performances of the Affine Jump Diffusion Models: the Role of Time-Varying Jump Risk Premia

4.1 Introduction

There have been numerous empirical studies on option pricing and time-series consistency for the affine jump diffusion (AJD) stock return models, but, unfortunately, their empirical results have often provided different results affected by estimation strategies and sample periods. The AJD class provides a variety of model specifications, particularly depending on the restrictions imposed upon jump components (for either return or volatility). The popular specifications in the class that have been widely studied in the literature are a square-root stochastic volatility model (SV), an extended model with jump-in-returns (SVJ), and a "double-jump" model with contemporaneous jumps in both return and volatility (SVCJ). Moreover, time-varying jump risk premia can be modeled into the SVJ, for instance, by incorporating a time-varying jump arrival intensity as in Bates (2000) and Pan (2002) (henceforth, SVJtv). Similarly, although rarely studied in the literature, the double-jump model (SVCJ) can also be equipped with time-varying jump risk premia (henceforth, SVCJtv). More complicated two-factor models have been studied by Bates (2000). One of the advantages in the AJD class over other stochastic volatility models (e.g., log stochastic volatility models) is that the AJD class provides an analytically tractable option pricing method.

We investigate option pricing and one-day-ahead density forecast performances for the AJD class of models in out-of-sample context. The one-day forecasting

horizon is chosen since it is typically used in practice to compute a value-at-risk measure. We use the S&P500 index and its associated options, which have been of major interest in the related literature. Compared to other studies, we consider more AJD specifications, including both constant and time-varying jump risk premia specifications, among which we pay attention to how time-varying jump risk premia contribute to model performances. For comparison purposes, we also consider nonlinear asymmetric GARCH (NGARCH) models proposed by Duan (1995), which provide a convenient option pricing method among the GARCH family. For our analysis, we estimate objective and risk-neutral parameters for each model from the in-sample period ranging from 1987 to 2000. Our out-of-sample covers 2001-2007, a relatively longer time-span than in the existing studies, which is expected to reveal how consistently each model can perform under different market conditions.

The past studies have provided mixed results for the option pricing performance of each AJD specification, as summarized by Broadie et al. (2007). Specifically, Bakshi et al. (1997) support jump-in-return, whereas Bates (2000) and Eraker (2004) argue that the benefit from incorporating jump is only modest. Recently, Broadie et al. (2007) show that both SVJ and SVCJ models (with constant jump risk premia) are superior to the simpler SV model in explaining cross-sectional differences in option prices. Comparing the SV and SVJtv models, Pan (2002) argues that time-varying jump risk premia are important in explaining the time-varying volatility smirks, and shows empirical evidence that the SVJtv specification provides a better fit to cross-sectional option data. However, to the best of our knowledge, a comparative analysis between constant and time-varying jump risk premia has rarely been pursued in the literature. One exception is Eraker (2004),

whose empirical results show that time-varying jump risk premia contribute little in fitting options data. We will address the role of time-varying jump risk premia in this study.

To find a better model specification, time-series consistency should also be addressed. Note that the time-series consistency requires that, besides a correct model specification, option market should be fully integrated with spot market, sharing the same price dynamics and market prices of risk (e.g., Pan, 2002). In other words, it requires the consistency between the objective and risk-neutral dynamics implied by a model. In fact, as argued by Bates (2000), there are so many testable implications from the time-series consistency issue that different studies have examined different aspects of this issue, leading to, unavoidably, mixed empirical results in the literature. For instance, Bakshi et al. (1997) and Bates (2000) note that option-based estimates for the volatility of volatility, σ_v , and correlation between the two Brownian increments, ρ , are generally inconsistent with the time-series dynamics of historical returns and volatilities. In contrast, Eraker (2004) argues that, through his joint returns and options study, his "volatility of volatility" estimate is consistent with time-series data, and that the SVCJ model captures the time-series dynamics reasonably well. Similarly, Broadie et al. (2007) also support the SVCJ model based on their diagnostic studies on both historical returns and volatilities. Notably, Pan (2002) argues that time-varying jump risk premia are important in reconciling the dynamics implied by the joint return and option data. In her overall goodness-of-fit test, she could not reject the SVJtv model.

We address the time-series consistency from the perspective of density forecast, pursuing more a practical implication. Density forecast is particularly important

in risk management. For instance, the value-at-risk (VaR), an important risk management tool in the financial industry, keeps track of certain aspect of asset returns' (conditional) distribution. It is needless to say that an accurate risk measurement is crucial in efficient capital allocation for financial institutions. If the time-series consistency is guaranteed for a given stock return model, one might be able to exploit option market information to forecast densities since option prices provide valuable forward-looking information on the future volatility path of asset process. On the contrary, under the separation between the option and corresponding spot markets (probably due to some market microstructural problem), density forecast based on option information may be seriously distorted. In this context, it would be important to examine to what extent option market information could improve density forecasting performance. For this purpose, we evaluate density forecasting performances for our AJD specifications under two different settings. Specifically, we consider return-based AJD models and option-based AJD models. The return-based models, denoted by SV-PF, SVJ-PF, and SVCJ-PF, forecast densities by using a past history of returns alone. The particle filters are used for these models as detailed below. In contrast, the option-based models, denoted by SV-OPT, SVJ-OPT, and SVCJ-OPT (characterized by constant jump risk premia), use option market information to forecast densities. Both SVJtv and SVCJtv with time-varying jump risk premia are also considered option-based models.

We evaluate density forecast performance for each model via Hong and Li's (2005) testing method, which is introduced in Chapter 2. This testing method can be conveniently used to compare non-nested models' relative density forecasting performances. Many stock return models are non-nested and estimated via different methods. In this case, model comparison might be challenging with other

testing methods. In our case, the GARCH and AJD classes are non-nested, and we estimate both classes by the maximum likelihood and Bayesian MCMC method, respectively. However, Hong and Li's (2005) testing method can compare the relative performance of non-nested models in a unified way by a metric measuring the departures of their generalized residuals from i.i.d. $U[0, 1]$.

As noted in Chapter 1, the key step in the Hong and Li's (2005) testing procedure is a dynamic probability integral transform, which requires an integration of a model-implied conditional density. The GARCH-type models generally have a closed-form density without any latent variable. Thus, the testing method is conveniently applicable to this type of models. Unfortunately, the AJD class of models have no closed-form conditional density. Moreover, if a spot volatility is not available, the conditional distribution will require the entire past history of returns as conditioning variables, making it even harder to apply the Hong and Li's (2005) method.

Our study makes methodological contributions to the literature by extending the Hong and Li (2005) testing method to be applicable to the famous AJD class of models, whether model-implied spot volatilities are available or not. First, under the premise that option market is fully integrated with the corresponding spot market, one can back out a spot volatility from a market option price. Although there is still no closed-form conditional density, the AJD class, fortunately, provides a closed-form conditional characteristic function in many cases. Exploiting this property, we reduce the dynamic probability integral transform procedure to the Fourier inversion of a known conditional characteristic function. This method is an extension of the well-known Levy inversion formula used in computing a risk-neutral probability of exercising a European call option for the AJD class (e.g.,

Heston, 1993; Bates, 1996; Duffie et al., 2000). We extend this method to the dynamic probability integral transform under the objective measure. In contrast to other approximation methods, this method does not incur any discretization bias, making it possible to forecast densities at any forecast time horizon without additional computational cost. Next, if option market is segmented from the spot market, due to some market microstructural problem, the only information we can use to forecast densities will be limited to the past history of stock returns. In this case, we use the particle filters based on Johannes et al. (2008), as proposed in Chapter 3, to implement dynamic probability integral transform through the Monte Carlo integration method.

Our empirical analysis shows the following results. First, for pricing option contracts, we find strong evidence in favor of time-varying jump risk premia for the AJD class of models. In particular, the SVCJtv, which has seldom been studied in the literature, are the most successful in fitting option prices. Our two time-varying jump risk premia models (i.e., SVJtv and SVCJtv) dominate the other constant jump risk premia models uniformly across both high- and low-volatility periods. This is in contrast to the empirical results in Bates (2000) and Eraker (2004). The constant jump risk premia models fail to capture the volatility smirk especially in the short-maturity in-the-money category, during the high-volatility period. Next, our out-of-sample density forecast (one-day-ahead) results show that all models are misspecified in terms of Hong and Li's (2005) statistics. However, the return-based AJD models (particularly, SVJ-PF) exhibit relatively good and stable density forecasting performances across both high- and low-volatility periods. The time-varying jump risk premia models (i.e., SVJtv and SVCJtv), which have drastically improved option pricing performance, seem to be inconsistent with

time-series dynamics. Among our AJD models, we could not find a specification that successfully reconcile both time-series and options data across both subsamples.

The rest of this chapter is organized as follows. In Section 4.2, we describe the stock return models considered throughout this chapter. In Section 4.3, we discuss the methodology for both density forecast evaluation and option pricing. In Section 4.4, we describe data, model parameter estimation methods, and estimation results. In Section 4.5, we evaluate option pricing and density forecast performances for each model. Finally, Section 4.6 concludes.

4.2 Models

We now introduce stock return models that we will study: the affine jump diffusion (AJD) class of models and Nonlinear Asymmetric GARCH (NGARCH) models. All models considered provide a convenient method of pricing option contracts.

4.2.1 Affine Jump Diffusion Class of Models

For the AJD class of models, we consider the following stock price and stochastic volatility processes under the objective measure, P :

$$\frac{dS_t}{S_t} = (r_t - \delta_t + \eta_S V_t - \lambda \mu_S) dt + \sqrt{V_t} dW_t^S + Z_t^S dN_t, \quad (4.1)$$

$$dV_t = \kappa(\theta - V_t) dt + \rho \sigma_v \sqrt{V_t} dW_t^S + \sqrt{1 - \rho^2} \sigma_v \sqrt{V_t} dW_t^v + Z_t^v dN_t, \quad (4.2)$$

where r_t is a risk-free interest rate, δ_t is a dividend yield, η_S is a equity premium, W_t^S and W_t^v are uncorrelated Brownian motions, $N_t \sim Poi(\lambda)$ is a Poisson-distributed jump timing with a constant intensity λ , Z_t^S is a jump-in-return size with $\ln(1 + Z_t^S) \sim N(\mu_S, \sigma_S^2)$, and $Z_t^v \sim \exp(\mu_v)$ is a jump-in-volatility size.²⁵

²⁵In this study, we consider the model with contemporaneously arriving jump-in-return and jump-in-volatility. Eraker et al. (2003) examine the model denoted by "SVIJ," where both jumps

We address three popular AJD specifications for the time-series process (under objective measure): a square-root stochastic volatility (SV) model, an extended model with jump-in-returns (the SVJ model), and a double-jump model with contemporaneously arriving jumps in return and volatility (the SVCJ model). The SV and SVJ models are special cases of our AJD specification with the restrictions that $N_t = 0$ for the SV and $Z_t^v = 0$ for the SVJ. Also, note that, for the time-series processes, we restrict the jump arrival intensity under objective measure (say λ_t^P) to be constant λ .

From finance theory, it is well known that there is equivalence between no arbitrage and the existence of a risk-neutral probability measure, Q . Following the convention in the literature, we assume the following risk-neutral (Q) process:

$$\frac{dS_t}{S_t} = (r_t - \delta_t - \lambda_t^Q \mu_S^Q)dt + \sqrt{V_t}dW_t^S(Q) + Z_t^S(Q)dN_t(Q), \quad (4.3)$$

$$\begin{aligned} dV_t &= (\kappa(\theta - V_t) + \eta_v V_t)dt + \rho\sigma_v\sqrt{V_t}dW_t^S(Q) \\ &\quad + \sqrt{1 - \rho^2}\sigma_v\sqrt{V_t}dW_t^v(Q) + Z_t^v(Q)dN_t(Q), \end{aligned} \quad (4.4)$$

where η_v is a diffusive volatility risk premium, $W_t^S(Q)$ and $dW_t^v(Q)$ are uncorrelated Brownian motions under risk-neutral measure, $N_t(Q) \sim Poi(\lambda_t^Q)$, $\ln(1 + Z_t^S(Q)) \sim N(\mu_S^Q, (\sigma_S^Q)^2)$, and $Z_t^v(Q) \sim \exp(\mu_v^Q)$. The Girsanov theorem imposes the restriction that the parameters, σ_v , ρ , κ , and θ are identical across both measures, P and Q . For remaining parameters, the differences in magnitude between both measures have an interpretation as risk premia. For example, Broadie et al. (2007) define the mean price jump risk premium as $\mu_S - \mu_S^Q$, the volatility of price jumps risk premium as $\sigma_S^Q - \sigma_S$, and the volatility jump risk premium as $\mu_v^Q - \mu_v$. For constant jump risk premia models (e.g., SVJ and SVCJ), we

arrive independently. However, Eraker et al. (2003) find that the SVIJ improves only marginally upon the SVCJ in capturing time-series dynamics.

assume that λ_t^Q is same as λ . In addition to the constant jump arrival intensity models, we consider additional risk-neutral dynamics for the SVJ and SVCJ, by allowing for a state-dependent risk-neutral jump arrival intensity, which is assumed to be linear in spot volatility, $\lambda_t^Q = \lambda_1^Q V_t$ (e.g., Bates, 2000; Pan, 2002). By the state-dependent jump arrival intensity, we can incorporate time-varying jump risk premia. We denote these specifications by the SVJtv and SVCJtv, respectively. The risk-neutral time-varying specification is possible since no arbitrage assumption imposes a much weaker restriction for change of measure on a jump process than on a diffusion process. Therefore, for the jump process, jump arrival intensities under both measures may differ in their current level, and also have different degrees of persistence and time-varying volatility (e.g., Singleton, 2006).

Our AJD specifications under the objective and risk-neutral measures implicitly assume the following pricing-kernel (or stochastic discount factor) process (denoted by M_t) in Equation (4.5).²⁶ Under the assumption that the pricing kernel, M_t , depends only on the current state, the pricing kernel can have an economic interpretation as an intertemporal marginal rate of substitution for consumptions at time t (e.g., Cochrane, 2001; Bates, 2000). Note that, without the drift term, " $-r_t dt$," in the following stochastic differential equation, the process amounts to the Radon-Nykodym derivative (dQ/dP) process.

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda_t^S dW_t^S - \Lambda_t^v dW_t^v - (\Gamma_t dN_t - (\lambda - \lambda_t^Q) dt), \quad (4.5)$$

where $\Lambda_t = (\Lambda_t^S, \Lambda_t^v)' = (\eta_S \sqrt{V_t}, -(1 - \rho^2)^{-1/2} (\rho \eta_S + \eta_v / \sigma_v) \sqrt{V_t})'$ represents market prices of diffusive risk²⁷, and $\Gamma_t = 1 - \lambda_t^Q f^Q(Z_t^S, Z_t^v) / (\lambda f^P(Z_t^S, Z_t^v))$ is a market price of jump risk with $f^P(\cdot, \cdot)$ and $f^Q(\cdot, \cdot)$ being the objective and risk-neutral

²⁶For the pricing-kernel process, we largely follow the notations and descriptions in Singleton (2006).

²⁷Therefore, we assume "completely affine" market prices of diffusive risk (e.g., Singleton, 2006).

joint densities for jump increments in S_t and V_t upon jump occurring, respectively.

By using Equation (4.1), (4.2) and (4.5), one can derive instantaneous covariances between each state variable and pricing kernel as given by Equation (4.6), (4.7), (4.8), and (4.9). Those covariances imply that various diffusive and jump risk premia reflect the compensation for bearing systematic risks.

$$Cov\left(\frac{dS_t}{S_t}, \frac{dM_t}{M_t}\right) = -\eta_S V_t dt, \quad (4.6)$$

$$\begin{aligned} Cov\left(\frac{\Delta S_t}{S_t}, \frac{\Delta M_t}{M_t}\right) &= \frac{\lambda_t^Q}{\lambda} (E^Q[Z_t^S] - E[Z_t^S]) \\ &= \frac{\lambda_t^Q}{\lambda} \left(\exp(\mu_S^Q + \frac{1}{2}\sigma_S^2) - \exp(\mu_S + \frac{1}{2}\sigma_S^2) \right), \end{aligned} \quad (4.7)$$

$$Cov\left(dV_t, \frac{dM_t}{M_t}\right) = \eta_V V_t dt, \quad (4.8)$$

$$\begin{aligned} Cov\left(\Delta V_t, \frac{\Delta M_t}{M_t}\right) &= \frac{\lambda_t^Q}{\lambda} (E^Q[Z_t^V] - E[Z_t^V]) \\ &= \frac{\lambda_t^Q}{\lambda} (\mu_V^Q - \mu_V), \end{aligned} \quad (4.9)$$

where Δ represents a jump increment upon jump occurring.

If we interpret the pricing kernel as an intertemporal marginal rate of substitution, one would naturally expect negative signs for the covariances between price and pricing kernel, (4.6) and (4.7), both of which are for diffusive and jump increments, respectively, since overall stock index tends to covary positively with wealth level. As suggested by the traditional consumption-CAPM, economic agents would then require an excessive return for bearing such systematic risks (i.e., $\eta_S > 0$ and $\mu_S - \mu_S^Q > 0$). In contrast, there seems to be little well-established economic theory for the covariances between volatility and pricing kernel, (4.8), and (4.9). As Singleton (2006) notes, the previous empirical option studies (e.g., Pan, 2002; Eraker, 2004; and Broadie et al., 2007) have provided a positive diffusive volatility risk premium ($\eta_S > 0$), though often statistically insignificant for jump-incorporated

models. This implies that volatility and wealth tend to move in the opposite direction. The leverage effect can be one potential reason for this negative relation. In this case, when volatility is high, option prices (particularly, for long-date options) are higher than what would be implied by historical volatility. Likewise, the result from Broadie et al. (2007) (i.e., $\mu_v^Q > \mu_v$) suggests positive covariance between the jump increments of volatility and pricing kernel.

Given estimated risk-neutral parameters, a European call option at time t (denoted by C_t) with maturity date T and strike price X is priced by the following formula:

$$C_t = E^Q \left[e^{-\int_t^T r_u du} (S_T - X)^+ \mid \mathcal{F}_t \right]. \quad (4.10)$$

One of the advantages in the AJD class is that there exists a closed-form formula for option pricing up to the Fourier inversion. Bates (1996) and Duffie et al. (2000) provide an option pricing method for various AJD specifications. We will briefly describe the method below. This advantage comes from the fact that many AJD models have a closed-form conditional characteristic function.

Hull and White (1987) shows that stochastic volatility can capture some anomalies of the Black-Scholes model, particularly the volatility-smile (or volatility-smirk) effect in option prices. Furthermore, the leverage effect can be explicitly taken into account by the parameter, ρ , or by negative mean jump size (i.e., negative μ_S). The jump component in the SVJ or SVCJ can provide additional flexibility in capturing some important features of asset return dynamics such as conditional skewness and leptokurtosis. In particular, the (risk-neutral) jump component is useful in capturing a highly steep volatility smirk observed in short-dated OTM (out of the money) put options.

4.2.2 Nonlinear Asymmetric GARCH

Contrary to the AJD class, the GARCH class of models assume a discrete-time process. Among a variety of GARCH models, we consider the nonlinear asymmetric GARCH (NGARCH) model. The NGARCH model is essentially the linear GARCH model that explicitly considers the leverage effect (negative correlation between returns and volatility innovations). For this specification, Duan (1995) proposes a theoretical framework for option pricing under a local risk-neutral probability measure.

Following Duan (1995), suppose the following stock price and volatility processes under the objective measure P :

$$\begin{aligned}\ln \frac{S_{t+1}}{S_t} &= r - \delta + \lambda \sigma_{t+1} - \frac{1}{2} \sigma_{t+1}^2 + \sigma_{t+1} \varepsilon_{t+1}, \\ \sigma_{t+1}^2 &= \beta_0 + \beta_1 \sigma_t^2 + \beta_2 \sigma_t^2 (\varepsilon_t^2 - \theta)^2\end{aligned}\tag{4.11}$$

where ε_{t+1} is a standard normal random variable, r is a risk-free interest rate, δ is a dividend yield, λ is a unit risk premium for the stock, and θ is the parameter capturing the leverage effect. What characterizes the above specification from other GARCH models is an incorporation of the leverage effect coefficient, θ , into the volatility process. Moreover, the drift term in the price process is more complicated than usual GARCH specifications. However, like other GARCH models, its model-implied volatility path up to the current date is observable given parameters and a past history of returns, which enables us to price option contracts without option market information.

Duan (1995) shows that the local risk neutrality assumption provides the fol-

lowing risk-neutral process:

$$\begin{aligned}\ln \frac{S_{t+1}}{S_t} &= r - \delta - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}\xi_{t+1} \\ \sigma_{t+1}^2 &= \beta_0 + \beta_1\sigma_t^2 + \beta_2\sigma_t^2(\xi_{t+1} - \theta - \lambda)^2\end{aligned}\quad (4.12)$$

where $\xi_{t+1} = \varepsilon_{t+1} + \lambda$ is a standard normal random variable under the risk neutral measure, Q .

This model, however, provides no closed-form formula for option pricing, so option value should be calculated via the Monte Carlo simulation.²⁸ For the efficient Monte Carlo simulation, Empirical Martingale Simulation (EMS) method is proposed by Duan and Simonato (1998). The EMS method ensures that the price estimated by simulation satisfies the rational option pricing bounds.²⁹ We adopt their EMS method for the Monte Carlo simulation, and set the number of simulations to be 10,000.

We consider two specifications: the NGARCH0 with the restriction that $\theta = 0$ and the NGARCH without this restriction. By considering these two models, we can address a contribution from modeling the leverage effect in the NGARCH specification. Another reason is that the restricted NGARCH0 model is similar in its specification to Bollerslev's (1986) GARCH(1,1) and JP Morgan's (1996) Risk-Metrics, which have been popularly used for the VaR implementation in finance industry. Therefore, our empirical results for the NGARCH0 are expected to pro-

²⁸Heston and Nandi (2000) propose a variant of GARCH model which provide a closed-form option pricing formula up to the Fourier inversion. Hence, the Monte Carlo simulation is not necessary for pricing option contracts. In its simplest specification, Heston and Nandi assumes the following volatility process under P ,

$$\sigma_{t+1}^2 = \beta_0 + \beta_1\sigma_t^2 + \beta_2(\varepsilon_{t+1}^2 - \theta\sigma_t)^2.$$

²⁹The associated GAUSS codes are available from Duan's webpage (www.rotman.utoronto.ca/~jcduan/). We use his code to evaluate model-implied option prices.

vide useful information on how the two popular models would have performed in our empirical study.

4.3 Methodology

This section describes the methodology used throughout this study to evaluate a density forecast performance and to price option contracts for each model.

As noted above, we employ Hong and Li's (2005) nonparametric specification testing method for density forecast evaluations. The dynamic probability integral transform, which is a key factor in this method, requires a closed-form conditional density function. For the NGARCH models, which satisfy this requirement, we can simply apply Hong and Li's (2005) original method. Unfortunately, the AJD models have no closed-form conditional density function. Even when we can back out the spot volatilities from option market data (e.g. in the case of SV-OPT, SVJ-OPT, SVCJ-OPT, SVJtv, and SVCJtv), we still don't have a conditional density as a closed form. Furthermore, with return data alone (without spot volatility) their conditional densities require all the past history of returns as conditioning variables (e.g., in the case of SV-PF, SVJ-PF, and SVCJ-PF). In this section, we propose dynamic probability integral transform schemes that can handle these two cases. Then, we discuss option pricing method for the AJD class of models. The option pricing method for NGARCH models has been outlined in Section 4.2.2.

4.3.1 Dynamic Probability Integral Transform via Inverse Fourier Transform

Suppose that the options market is fully integrated with the corresponding spot market. Then, in order to forecast density for an AJD model, we can use model-implied spot volatilities extracted from option prices. One difficulty is that, even with the help of extracted volatility, we still cannot obtain a closed-form condi-

tional density function for the AJD class of models. There are some alternative methods that handle this problem. One can use Ait-Sahalia's (2002a, 2002b) Hermite polynomial method or the simulation methods of Pedersen (1995), Elerian et al. (2001) and Brandt and Santa-Clara (2002) to approximate the transition density. These methods are computationally expensive, particularly when the sample size is large. Furthermore, most methods are exposed to a discretization bias to some extent.

Fortunately, our specification for the AJD has a closed-form conditional moment generating function (henceforth CMGF). The CMGF, $\phi_{t,T}(u)$, is defined as

$$\phi_{t,T}(u) = E[e^{u \ln S_T} | \mathcal{F}_t], \quad (4.13)$$

where $u \in \mathbb{C}$. Note that the filtration, \mathcal{F}_t , can be replaced by S_t and V_t from the Markov property of the AJD process. The CMGF of the AJD process then has the following exponential affine form:

$$\phi_{t,T}(u) = \exp(\phi_{0t} + \phi_{1t} \ln S_t + \phi_{2t} V_t) \quad (4.14)$$

where ϕ_{0t} , ϕ_{1t} and ϕ_{2t} are complex functions of $\tau (= T - t)$ and u . This explains why the AJD process has the term "affine" in its name.

Using the exponential affine form of the CMGF for the AJD models, we can implement a dynamic probability integral transform by an inverse Fourier transform. In fact, this method is widely known as the Levy inversion method (under the risk-neutral Q -measure) in the option pricing literature (Heston, 1993; Bates, 1996; Duffie et al., 2000). Our method extends this method to an objective measure version.

Suppose that we are given the CMGF of the AJD process. A generalized residual at time T conditioning on the information set at time t can be computed

by the following Levy inversion formula:

$$z_{t,T} = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im} [\phi_{t,T}(iu)e^{-iu \ln s_T}]}{u} du. \quad (4.15)$$

The proof is provided in Appendix 4.1. Since this method does not involve any discretization bias, though it does have some integration error, we can compute generalized residuals at any forecast horizon ($\tau = T - t$) without worrying about the discretization bias. The CMGF of our AJD specification is given by (4.16), which is based on Bates (1996) and Duffie et al. (2000). Suppose that $\tau = T - t$, $\bar{k} = \exp(\mu_S + \frac{1}{2}\sigma_S^2) - 1$, and $\gamma = \sqrt{(\rho\sigma_v u - \kappa_v)^2 + \sigma_v^2(u - u^2)}$. The CMGF of the AJD process in (4.1) and (4.2) is then

$$\phi_{t,T}(u) = \exp(u \ln S_t + A(u, \tau) + B(u, \tau)V_t + \lambda C(u, \tau)(1 + \bar{k})^u e^{-\frac{1}{2}\sigma_S^2(u-u^2)} - \lambda\tau), \quad (4.16)$$

where

$$\begin{aligned} A(u, \tau) &= \mu u \tau - \frac{\kappa \theta \tau}{\sigma_v^2} (\rho \sigma_v u - \kappa - \gamma) \\ &\quad - \frac{2\kappa \theta}{\sigma_v^2} \ln \left[1 + \frac{1}{2} (\rho \sigma_v u - \kappa - \gamma) \frac{1 - e^{\gamma \tau}}{\gamma} \right], \\ B(u, \tau) &= \frac{u - u^2}{\rho \sigma_v u - \kappa + \gamma \frac{1 + e^{\gamma \tau}}{1 - e^{\gamma \tau}}}, \\ C(u, \tau) &= \frac{\gamma - \rho \sigma_v u + \kappa}{\gamma - \rho \sigma_v u + \kappa + \mu_v (u - u^2)} \cdot \tau + \frac{-2(u - u^2)\mu_v}{\gamma^2 - (\rho \sigma_v u - \kappa - \mu_v (u - u^2))^2} \\ &\quad \ln \left(1 - \frac{1}{2} \frac{\gamma + \rho \sigma_v u - \kappa - \mu_v (u - u^2)}{\gamma} (1 - \exp(-\gamma \tau)) \right). \end{aligned}$$

The integration in (4.15) can be evaluated quickly by the Gauss-Legendre quadrature for suitably chosen intervals. For example, we use three intervals for integration: $[0, 10]$, $[10, 100]$, and $[100, 1000]$. Alternatively, one may be able to use the numerical integration method introduced by Pan (2002). In her method, letting $I(u) = \text{Im} [\phi(iu)e^{-iu \ln s_T}]$ denote the integrand, one can approximate the

integration by

$$\tilde{z} = \frac{1}{2} - \frac{1}{\pi} \sum_{n=0}^{\lfloor U/\Delta u \rfloor} \frac{I((n+1/2)\Delta u)}{n+1/2}, \quad (4.17)$$

where $\lfloor x \rfloor$ is an integer such that $\lfloor x \rfloor - 1 < x \leq \lfloor x \rfloor$, and truncation and discretization errors can be controlled by suitably chosen values for U and Δu . We have found that both numerical integration methods give almost identical results. However, considering the computing speed, we adopt the former Gauss-Legendre integration method.

4.3.2 Dynamic Probability Integral Transform via Particle Filtering

If the options market is separated from the corresponding spot market, probably due to some market-microstructural reasons, we should forecast densities by relying only on time series of stock returns. Without observable spot volatilities, we cannot exploit the Markov property of the AJD process, which makes it hard to directly apply Hong and Li's (2005) method. Now the dynamic probability integral transform is infeasible with the usual methods. To handle this problem, we employ the "particle filtering" method. We have already introduced the particle filtering method in Chapter 3. Here, we illustrate the APF algorithm applicable to the simple SV model. The algorithm for the SVCJ has been introduced by Appendix 3.1 in Chapter 3.

Suppose the following SV model from the Euler discretization scheme.

$$y_t = \alpha + \sqrt{V_t} \varepsilon_t, \quad (4.18)$$

$$V_t = V_{t-1} + \kappa(\theta - V_{t-1}) + \rho\sigma_v(y_t - \alpha) + \sqrt{1 - \rho^2}\sigma_v\sqrt{V_{t-1}}\eta_t, \quad (4.19)$$

where $y_t = \ln(S_t/S_{t-1})$, $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$, and stochastic volatility V_t is an unobservable latent state variable. We take the following steps recursively.

- **Step 1:** Suppose that we are given N particles for volatility, $\{V_{t-1}^{(i)} \mid y^{t-1}\}_{i=1}^N$, where $V_{t-1}^{(i)}$'s are simulated based on the past history of returns, $y^{t-1} = \{y_1, \dots, y_{t-1}\}$. By sampling $\{\eta_t^{(i)}\}_{i=1}^N$ from i.i.d. $N(0, 1)$, we can obtain $\{V_t^{(i)} \mid y^{t-1}\}_{i=1}^N$ through (4.19). Then, with $\{V_t^{(i)} \mid y^{t-1}\}_{i=1}^N$ and y_t (a realized stock return at time t), we can approximate the generalized residual at time t by

$$\begin{aligned}
z_t &= \int_{-\infty}^{y_t} p(y \mid y^{t-1}) dy \\
&= \int_{-\infty}^{y_t} \int_0^{\infty} \int_0^{\infty} p(y \mid V_t) p(V_t \mid V_{t-1}, y_{t-1}) p(V_{t-1} \mid y^{t-1}) dV_t dV_{t-1} dy \\
&\simeq \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} p(y \mid V_t^{(i)}) dy \\
&\simeq \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{y_t} \phi(y \mid \alpha, V_t^{(i)}) dy
\end{aligned} \tag{4.20}$$

where $\phi(\cdot \mid \mu, \sigma^2)$ is a normal density function with mean μ and variance σ^2 . After computing the generalized residual at time t , we discard the simulated volatilities, $\{V_t^{(i)}\}_{i=1}^N$.³⁰

- **Step 2:** This step begins to update the particles from $\{V_{t-1}^{(i)} \mid y^{t-1}\}_{i=1}^N$ to $\{V_t^{(i)} \mid y^t\}_{i=1}^N$. First, we compute an auxiliary variable, $\widehat{V}_t^{(i)}$ for each i , which is a conditional expectation of the volatility at time t , given $V_{t-1}^{(i)}$. In other words,

$$\widehat{V}_t^{(i)} = V_{t-1}^{(i)} + \kappa(\theta - V_{t-1}^{(i)}) + \rho\sigma_v(y_{t-1} - \mu). \tag{4.21}$$

Using $\widehat{V}_t^{(i)}$, we evaluate the first stage weight for each i :

$$w_t^{(i)} \propto \phi(y_t \mid \mu, \widehat{V}_t^{(i)}). \tag{4.22}$$

With the weight for each i , we resample the existing particles, $\{V_{t-1}^{(i)}\}_{i=1}^N$, which provides resampled particles, $\{V_{t-1}^{k(i)}\}_{i=1}^N$.

³⁰If we used the SIR algorithm rather than the APF, the simulated volatilities could be used for updating the particles.

- **Step 3:** For each $k(i)$, we generate volatility at time t by drawing $\eta_t^{k(i)}$ from i.i.d. $N(0, 1)$:

$$V_t^{k(i)} = V_{t-1}^{k(i)} + \kappa(\theta - V_{t-1}^{k(i)}) + \sigma_v \rho (y_{t-1} - \alpha) + \sigma_v \sqrt{1 - \rho^2} \sqrt{V_{t-1}^{k(i)}} \eta_t^{k(i)}, \quad (4.23)$$

which provides $\{V_t^{k(i)}\}_{i=1}^N$.

- **Step 4:** As a final step, resample the particles, $\{V_t^{k(i)}\}_{i=1}^N$ with the second stage weights defined by

$$\pi_t^{k(i)} \propto \frac{p(y_t | V_t^{k(i)})}{p(y_t | \widehat{V}_t^{k(i)})} = \frac{\phi(y_t | \alpha, V_t^{k(i)})}{\phi(y_t | \alpha, \widehat{V}_t^{k(i)})}. \quad (4.24)$$

From the second stage resampling, we can obtain $\{V_t^{(i)} | y^t\}_{i=1}^N$.

- **Step 5:** Go to Step 1, and compute the generalized residual at time $t + 1$.

The remaining step is to compute Hong and Li's (2005) statistics as in Chapter 2. In our analysis, we set the number of particles to be 2,500. Our separate experiment shows that a smaller number of particles like 1,000 provide very close results.

4.3.3 Option Pricing for the AJD Class of Models

For the AJD class of models, we apply the option pricing method proposed by Bates (1996) and Duffie et al. (2000). Heston (1993) first developed a closed-form solution up to the Fourier inversion for a simple SV specification (without jump). Later, Duffie et al. (2000) generalize Heston's method, including double jump specifications (jumps in both return and volatility).

For a reason which will be explained in Section 4.4.1, we consider a forward price at time t with expiration date T (hereafter $F_{t,T}$), which can be computed

from call and put options with the same maturity date (T) and strike price (X) via put-call parity. The forward and spot prices are related to each other as follows.

$$F_{t,T} = E^Q[S_T | \mathcal{F}_t] = S_t e^{(r-\delta)\tau}, \quad (4.25)$$

where r and δ are assumed to be constant for notational convenience. Note that the forward and spot prices are identical at maturity date T , that is $F_{t,T} = S_T$.

Given a spot price process under the risk neutral measure Q as in (4.4), its forward counterpart solves the following stochastic differential equations:

$$\begin{aligned} \frac{dF_{t,T}}{F_{t,T}} &= -\lambda^Q \mu_S^Q dt + \sqrt{V_t} dW_t^S(Q) + Z_t^S(Q) dN_t(Q), \\ dV_t &= (\kappa(\theta - V_t) + \eta_v V_t) dt + \rho \sigma_v \sqrt{V_t} dW_t^S(Q) \\ &\quad + \sqrt{1 - \rho^2} \sigma_v \sqrt{V_t} dW_t^v(Q) + Z_t^v(Q) dN_t(Q), \end{aligned} \quad (4.26)$$

where all the notations are the same as in (4.4).

Note that $r_t - \delta_t$ is dropped from the stock price process in (4.4). Like the spot price process, the Girsanov theorem constrains the parameters σ_v , ρ , κ , and θ to be identical across both measures P and Q .

Under this setting, the call option contract with maturity date T and strike price X is priced at time t by

$$C_t = E^Q[e^{-r\tau}(F_{t,T} - X)^+ | \mathcal{F}_t] = e^{-r\tau}(F_{t,T}P_1 - XP_2). \quad (4.27)$$

The above P_1 and P_2 can be evaluated by the Levy inversion formula such that

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Im}[\phi_j(iu)e^{-iu \ln X}]}{u} du, \quad (4.28)$$

where $\phi_{t,T}(u)$ is the CMGF for the process in (4.26) ($\phi_{t,T}(u) = E^Q[e^{u \ln F_{t,T}} | \mathcal{F}_t]$), $j = 1, 2$, $\phi_1(iu) = \phi_{t,T}(iu)$, and $\phi_2(iu) = \phi_{t,F}(iu + 1)$. Note that Equation (4.28) is very similar to Equation (4.15). Explicit formula for P_j is provided in Appendix 4.2.

Similar to the case for the AJD, when a call option is priced through a forward price process in the NGARCH setting, " $r - \delta$ " term has to be dropped out from (4.12).

4.4 Data and Model Parameter Estimation

4.4.1 Data

To estimate each model's parameter under the objective measure (P), we use S&P500 stock index returns from January 1987 to December 2000. Excluding weekends and holidays, we have 3,538 daily observations. We use 3-month T-bill rates for risk-free short-rate,³¹ which is required to estimate the NGARCH models and to evaluate option prices.

We save the data from January 2001 to December 2007 with 1,757 observations to evaluate out-of-sample density forecast performances. To examine each model's performances under different market conditions, we divide the out-of-sample into two subsamples: Subsample 1 (2001-2003) and Subsample 2 (2004-2007). Both subsamples are characterized by different volatility levels. We separately evaluate density forecast and option pricing performances for each model across both subsamples. In terms of the VIX index, the average volatility levels are 25.0 for Subsample 1, and 14.7 for Subsample 2. During Subsample 1, there had been the bursting of the Dot-Com bubble, the attacks on 9/11, and the commencement of the Iraqi war, which led to a highly volatile stock market. Figure 4.1 displays the history of the VIX index throughout both in- (from 1990) and out-of-samples.

Our option contract data set (source: DeltaNeutral) consists of daily closing prices for the S&P 500 index option contracts traded in the CBOE (Chicago Board of Options Exchange). The S&P 500 index options are among the most actively

³¹Source: FRB H.15 historical data (www.federalreserve.gov)

traded options in the world. Since the contracts are European style, an early exercise premium does not matter. We take the averages of bid and ask quotes to compute option prices.

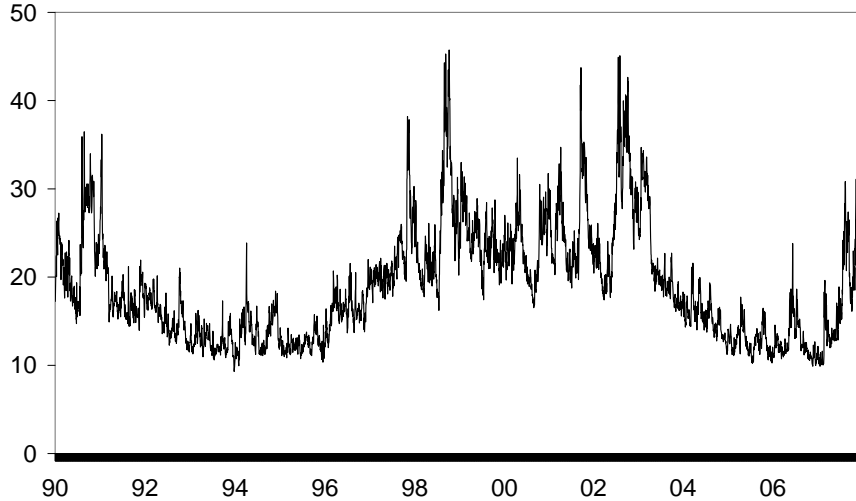


Figure 4.1: Historical Volatilities (VIX index)³²

However, as Ait-Sahalia and Lo (1998) and Eraker (2004) argue, it is hard to observe the underlying index at the exact times when the option prices are recorded,³³ which may induce non-synchronicity bias. Moreover, we cannot observe expected future dividend yields. To avoid these problems, as Ait-Sahalia and Lo (1998) suggest, we back out the forward value of the underlying index through a put-call parity. As indicated by (4.26) and (4.27), option prices implied by a model do not involve a dividend rate any more when the forward index is used. The put-call parity at time t is given by

$$C(S_t, X, T - t, r, \delta) + Xe^{-r(T-t)} = P(S_t, X, T - t, r, \delta) + F_{t,T}e^{-r(T-t)}. \quad (4.29)$$

³²Formerly, the VIX was calculated from S&P 100 options (current ticker VXO). As of September 22, 2003, the VIX began to use options on the S&P 500. The above figure connects both indices.

³³There is a 15-minute difference between the close of major stock markets such as the AMEX, NASDAQ, and NYSE (where a large portion of stocks included in the index are traded), and the Chicago options market.

where $C(\cdot)$ and $P(\cdot)$ are call and put prices at time t respectively, with underlying price S_t , strike X and expiration date T . Since we know that $F_{t,T} = S_T$, the option prices for forward and spot indices are the same when strike prices and maturity structures coincide for two options.

To evaluate both density forecast and option pricing performances of the option-based models, we need to extract filtered spot volatilities for each model. For this purpose, we choose a representative option for each day. Similar methods are used in Pan (2002) and Broadie et al. (2007). We roughly follow Pan (2002) for the representative option selection scheme: among all available options for each day, we select an option contract with a time-to-maturity as close as possible to 30 calendar days and a moneyness (the ratio of strike to underlying price) as close as possible to 1. This scheme guarantees that the most liquid options are selected to represent a model-implied spot volatility for each day.

Now we describe the selection of option contract sample used for evaluating option pricing performances. To reduce a computational burden, we select every Wednesdays' option contracts from January 2001 to December 2007 and then apply the following selection criteria to avoid a liquidity-related bias: We exclude (1) option contracts with maturity less than six days and more than one year, (2) option contracts with quotes lower than $\$3/8$, and (3) option contracts with moneyness (strike/underlying index) higher than 1.1 and less than 0.9. To establish these criteria, we referred to Bakshi et al. (1997), Chernov and Ghysels (2000), and Pan (2002). Consequently, we end up with 31,087 call option contracts (the same number of put option contracts are used to back out the underlying forward prices).³⁴

³⁴The number of call option contracts for each maturity and moneyness category is shown in Table 4.5.

4.4.2 Objective Parameter Estimation

We estimate the NGARCH models via the maximum likelihood by using their closed-form conditional density. The AJD models, however, have no closed-form conditional density, and they involve unobservable latent state variables such as stochastic volatility and jump, making their estimation more challenging. Recently, a variety of estimation methods have been developed for this class of models.³⁵ Among the alternatives, we select the Bayesian MCMC (Monte Carlo Markov Chain).³⁶

Although the AJD models are continuous-time models, they are usually estimated from discretely sampled data. Following Eraker et al. (2003), we adopt the Euler-Maruyama discretization scheme for a diffusion part and the Bernoulli approximation for a jump component. Although the Euler-Maruyama scheme may induce a discretization bias, it has been documented in the literature that the induced bias is relatively small for daily frequency data. Eraker et al. (2003) provide a simulation study in support of the discretization scheme for daily frequency. We consider the following discretized specification:

$$\Delta \ln S_t = \alpha + \sqrt{V_t} \varepsilon_t + J_t Z_t^S, \quad (4.30)$$

$$\begin{aligned} V_t = & V_{t-1} + \kappa(\theta - V_{t-1}) + \rho\sigma_v(\Delta \ln S_{t-1} - \alpha - J_{t-1}Z_{t-1}^S) \\ & + \sqrt{1 - \rho^2}\sigma_v\sqrt{V_{t-1}}\eta_t + J_{t-1}Z_t^v, \end{aligned} \quad (4.31)$$

³⁵For example, there are the GMM by Melino and Turnbull (1990) and Andersen and Sorensen (1996); the QMLE by Harvey et al. (1994) and Harvey and Shephard (1996); the EMM by Andersen et al. (1999) and Andersen et al. (2002); the simulation-based maximum likelihood by Danielsson (1994) and Danielsson and Richard (1993); and the Markov Chain Monte Carlo (MCMC) method by Jacquier et al. (1994), Kim et al. (1998), Eraker (2001), Eraker et al. (2003) and Yu (2005).

³⁶The MCMC method is essentially a likelihood-based estimation method. It is known that, in terms of efficiency, the likelihood-based inference methods are superior to the moment-based counterparts such as the GMM and EMM methods. Notably, Andersen et al. (1999) provide their Monte Carlo simulation results showing that the MCMC outperforms the EMM method in estimating a log stochastic volatility model.

where $(\varepsilon_t \eta_t)' \sim \text{i.i.d. } N(0, I_2)$, $J_t \sim \text{i.i.d. } \text{Ber}(\lambda)$, $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$, and $Z_t^v \sim \text{i.i.d. } \exp(\mu_v)$.³⁷

For implementing the MCMC, we use a recently-developed Bayesian statistical software called "WinBUGS." Meyer and Yu (2000) and Yu (2005) show that WinBUGS performs well in estimating log stochastic volatility models. Based on the WinBUGS codes from Yu (2005),³⁸ we have developed our own WinBUGS code applicable to the AJD class of models. Following Eraker *et al.* (2003) and Yu (2005), we ran the MCMC algorithm for 110,000 iterations, discarding the first 10,000 as a burn-in-period to achieve the convergence of the chain. For each parameter to be estimated, we use the same priors as in Eraker *et al.* (2003).³⁹

Table 4.1: Parameter Estimates for the NGARCH Models

	λ	θ	β_0	β_1	β_2	log-likelihood
NGARCH0	0.0704 (0.0162)	- (-)	1.8e-006 (4.6e-007)	0.8884 (0.0141)	0.1017 (0.0120)	11633.7
NGARCH	0.0302 (0.0161)	0.7562 (0.0751)	2.3e-006 (5.1e-007)	0.8455 (0.0181)	0.0883 (0.0112)	11682.7

Note: The models are nested by the specification:

$$\ln \frac{S_{t+1}}{S_t} = r_t + \lambda \sigma_{t+1} - \frac{1}{2} \sigma_{t+1}^2 + \sigma_{t+1} \varepsilon_{t+1}, \text{ and}$$

$$\sigma_{t+1}^2 = \beta_0 + \beta_1 \sigma_t^2 + \beta_2 \sigma_t^2 (\varepsilon_t^2 - \theta)^2. \text{ Standard errors are in parentheses.}$$

Table 4.1 and 4.2 report the estimation results. For the NGARCH models, we obtained the estimates for β_1 and β_2 whose magnitudes are close to typical estimates in the GARCH literature. The sums of β_1 and β_2 are close to one, suggesting that volatility dynamics are highly persistent. Notably, we obtained

³⁷Note that we let the parameter α approximate the drift term in Equation (4.1) as in Eraker *et al.* (2003). They argue that a more complicate specification in the drift does not improve upon this approximation.

³⁸They are available from Jun Yu's webpage (www.mysmu.edu/faculty/yujun).

³⁹The priors are $\mu \sim N(1, 25)$, $\kappa\theta \sim N(0, 1)$, $\kappa \sim N(0, 1)$, $\sigma_v^2 \sim IG(2.5, 0.1)$, $\rho \sim U(-1, 1)$, $\lambda \sim \text{Beta}(2, 40)$, $\mu_S \sim N(0, 100)$, $\sigma_S^2 \sim IG(5.0, 20)$, $\mu_v \sim G(20, 10)$, where G refers to a Gamma distribution, IG to the Inverse Gamma distribution, and U to the standard uniform distribution.

a significant θ in the NGARCH, indicating a substantial leverage effect in the GARCH process. For the AJD class of models, our estimates are close to those from Eraker et al. (2003), partly due to the same priors for parameter estimation. However, we obtained larger magnitudes for ρ and σ_v for each AJD model. In fact, our magnitudes are closer to those from the previous option pricing studies such as Bakshi et al. (1996), Pan (2002), and Eraker (2004). Specifically, our estimates for ρ and σ_v are -0.61 and 0.13, respectively, for the SVJ model, while Eraker et al. (2003), using time-series data, obtains -0.47 and 0.10, respectively, in the same SVJ specification. From options data alone, Bakshi et al. (1996) obtains -0.57 and 0.15, respectively. As noted, Bakshi et al. (1997) and Bates (2000) argue that σ_v and ρ estimated from options data are too large in magnitude to be consistent with time-series dynamics. Contrary to their argument, our estimates from time-series data are closer to those estimated from option data, indicating that the large magnitudes for ρ and σ_v from the previous option studies may have been simply due to different sample periods.

Table 4.2: Objective Parameter Estimates for the AJD Class of Models

	SV	SVJ	SVCJ
α	0.0401 (0.0128)	0.0432 (0.0125)	0.0475 (0.0125)
κ	0.0288 (0.0052)	0.0204 (0.0048)	0.0362 (0.0060)
θ	1.0210 (0.1133)	0.9856 (0.1340)	0.7026 (0.0711)
σ_v	0.1813 (0.0128)	0.1450 (0.0132)	0.1459 (0.0131)
ρ	-0.4960 (0.0528)	-0.5912 (0.0511)	-0.5712 (0.0500)
λ		0.0066 (0.0025)	0.0049 (0.0017)
μ_S		-3.3490 (1.4190)	-4.6460 (1.1250)
σ_S		3.7040 (0.7329)	2.7100 (0.8013)
μ_v			2.1040 (0.4025)

Note: The models are nested by the specification: $\Delta \ln S_t = \alpha + \sqrt{V_t} \varepsilon_t + J_t Z_t^S$, and $V_t = V_{t-1} + \kappa(\theta - V_{t-1}) + \rho \sigma_v (y_t - \alpha) + \sqrt{1 - \rho^2} \sigma_v \sqrt{V_{t-1}} \eta_t + J_{t-1} Z_t^v$, where $(\varepsilon_t, \eta_t)' \sim \text{i.i.d. } N(0, I_2)$, $J_t \sim \text{i.i.d. } \text{Ber}(\lambda)$, $Z_t^S \sim \text{i.i.d. } N(\mu_S, \sigma_S^2)$, and $Z_t^v \sim \text{i.i.d. } \exp(\mu_v)$. The estimates correspond to percentage changes in the index value. Standard deviations of posteriors are reported in parentheses.

4.4.3 Risk-neutral Parameter Estimation

In order to evaluate option prices or to extract model-implied spot volatilities for each AJD model, we need to estimate risk-neutral parameters as in (4.26). There are a few difficulties that make the estimation challenging. The estimation procedure usually involves a huge amount of panel data of option contracts. Generally, objective functions for the associated optimization procedure are highly nonlinear, so convergence in the algorithm may occur very slowly. Furthermore, there is no consensus on the distribution of option pricing errors, leading to a likelihood-based inference infeasible (e.g., Bates, 2000).

Several estimation strategies have been pursued in the literature to estimate the parameters using option data. For instance, Bakshi et al. (1997) employ a simple nonlinear least square method to minimize the sum of squared pricing errors using cross-sectional option data. They recalibrate parameters on a daily basis then evaluate one-day-ahead option contracts with the calibrated parameters. In contrast, adopting some error component structure, Bates (2000) estimates the parameters, which are assumed to be constant throughout the sample, by using quasi-maximum likelihood based on Kalman filtering method. The method-of-moment type methods have been used in some studies using joint return and option data. For example, Chernov and Ghysels (2000) use the EMM (Gallant and Tauchen, 1996) method to estimate the Heston model (i.e., SV from our notations), and Pan (2002) proposes the IS-GMM (implied-state GMM) exploiting the closed-form conditional characteristic function of the AJD process. The Bayesian MCMC method has been introduced into the empirical option literature by Eraker (2004).

Those past studies have tried to reduce computational burdens in several dimensions. For instance, many previous studies have used weekly data (e.g., mainly

Wednesday’s contracts). Some studies use only one to three option contracts per day (e.g., Chernov and Ghysels, 2000; Pan, 2002; Eraker, 2004). In contrast, Bates (2000) and Broadie et al. (2007), though computationally expensive, include as many option contracts as possible and estimate the parameters, aiming at better capturing the cross-sectional feature of option prices.

Following the previous studies, we sample Wednesday’s closing option prices for in-sample period (1990-2000) but include as many observations per day as possible. As before, in order to avoid a liquidity-related bias, we apply the same selection criteria as introduced in Section 4.4.1 and use all put options to back out the underlying forward prices. Our options data set, unfortunately, exhibits a lot of variations in the numbers of available option contracts across different dates. We select the dates in which the number of contracts is greater than 20 after the selection criteria are applied. Consequently, we end up with 433 sample dates for estimation, and the total number of call option contracts is 26,413 (the daily average of 61), so, even after our selection procedure, we still have a sufficiently large number of option contracts compared to other previous studies.

For the number of risk-neutral parameters to be estimated for each model, we aim to be as parsimonious as possible, based on the existing theories and the past empirical studies. As noted, we constrain the parameters σ_v , ρ , κ , and θ to be identical across both objective and risk-neutral measures, based on Girsanov theorem. Like the past option studies such as Pan (2002) and Eraker (2004), we also impose the restrictions for constant jump risk premia models such that the jump parameters λ^Q and σ_S^Q are equal to their objective-measure counterparts (i.e., λ and σ_S respectively).⁴⁰ For the AJD models with jumps, we constrain the

⁴⁰Broadie et al. (2007) argue that, generally, it is only possible to estimate the jump compensator, $\lambda^Q \mu_S^Q$, and not the individual components separately. Thus, λ^Q is often constrained to be equal to λ in the literature. Also, as in Broadie et al. (2007), the restriction that $\sigma_S^Q = \sigma_S$

diffusive volatility risk premium, η_v , to be zero. It has been documented in the past studies that, for the jump models, it is difficult to precisely estimate η_v and that the estimate is statistically insignificant.⁴¹ For all parameters listed so far, we take their values from the time-series estimates in Table 4.2. A similar risk-neutral parameter estimation scheme is used by Broadie et al. (2007). As a result of our restrictions, we have only to estimate one or two risk-neutral parameters for each model (e.g., η_v for SV; μ_S^Q for both SVJ and SVJtv; μ_S^Q and μ_v^Q for both SVCJ and SVCJtv) along with spot volatility for each date.

For the time-varying jump risk premia models, such as SVJtv and SVCJtv, we employ a linear jump intensity specification, that is, $\lambda_t^Q = \lambda_1^Q V_t$, which has been studied by Bates (2000) and Pan (2002). However, instead of separately estimating λ_1^Q , we calibrate the value by dividing λ (constant jump arrival intensity) by θ (long-run average volatility), where the values of λ and θ are taken from the time-series estimates (i.e., from the estimates for SVJ and SVCJ, respectively). Consequently, our time-varying jump intensities for the SVJtv and SVCJtv are artificially designed to be the same as constant ones for SVJ and SVCJ, respectively, when volatility is at its long-run average level. Similarly, as before, this calibration scheme comes from the argument in Broadie et al. (2007) that the individual components from the jump compensator, $\lambda_t^Q \mu_S^Q$ ($= \lambda_1^Q V_t \mu_S^Q$), are not separately identified.

We estimate both risk-neutral parameters and spot volatilities jointly by non-linear least squares (NLS), that is, we minimize the sum of squared differences

comes from the Lucas economy equilibrium models in Bates (1988), which assume power utility over consumption or wealth.

⁴¹According to Pan's (2002) Lagrange-multiplier test for the SVJ model with time-varying jump intensity (i.e. SVJtv in our model notations), the null hypothesis that $\eta_v = 0$ is not rejected against the alternative, $\eta_v \neq 0$. In a similar vein, Broadie et al.(2007) show that there is no significant difference in option pricing performances between the SVJ (or SVCJ) models with and without diffusive volatility risk premium.

between market and model-implied Black-Scholes implied volatilities (BSIV) as in Equation (4.32).

$$(\widehat{\theta}^Q, \widehat{V}) = \arg \min \sum_{t=1}^T \sum_{n=1}^{O_t} [BSIV_{n,t}^{Market} - BSIV_{n,t}^{Model}(V_t, \theta^Q | \theta^P)]^2, \quad (4.32)$$

where V_t is a spot volatility at date t , V is a spot-volatility vector such that $V = (V_1, \dots, V_T)'$, T is the number of sample dates ($T = 433$), O_t is the number of option prices on date t , θ^Q is a vector of free risk-neutral parameters, θ^P is vector of both objective-measure and constrained parameters, $BSIV_{n,t}^{Market}$ is a market-observed BSIV, and $BSIV_{n,t}^{Model}(V_t, \theta^Q | \theta^P)$ is a model-implied BSIV as a function of both V_t and θ^Q given θ^P . For notational simplicity, we suppress strike price $K_{t,n}$, maturity $\tau_{t,n}$, underlying (implied) forward price $F_{t,n}$ and short-rate r_t for each $BSIV_{n,t}$. This implied-volatility metric enables us to avoid undue weights on relatively illiquid deep in-the-money or long-dated options (e.g., Pan, 2002) since the BSIVs have the same order of magnitude across different strikes and maturities. Broadie et al. (2007) also use this BSIV metric to estimate their risk-neutral parameters.

However, the above minimization problem involves a highly nonlinear objective function and too many parameters (i.e., too many spot volatilities) with respect to which the objective function is minimized. We circumvent this computational difficulty by concentrating the parameter $\widehat{\theta}^Q$ out of the objective function. That is, we transform each V_t into a function of θ^Q by using the first order condition with respect to each V_t , and reduce the dimension of the above minimization problem. Now the objective function in Equation (4.32) has only to be minimized with respect to θ^Q . In some models (e.g., SV, SVJ, and SVCJ), however, we encountered some difficulty in obtaining the convergence of the concentrated minimization algorithm. The convergence seems to be very sensitive to the choice of initial value

for θ^Q . For those models where the convergence doesn't occur, we adopt the grid search method, though computationally expensive.

Our risk-neutral parameter estimates are reported in Table 4.3, where the RMSEs in the last row represent root mean squared errors (%) in terms of the BSIV metric. It turns out that the time-varying jump risk premia models, such as SVJtv and SVCJtv, offer the best in-sample fit for option pricing in terms of the RMSE. Among the two, the SVCJtv improves modestly upon the SVJtv. In contrast, the constant jump risk premia models (SVJ and SVCJ) exhibit a poor in-sample fit. Notably, unlike the previous studies, the SVCJ performs the worst among all models in terms of the RMSE, even worse than the SV.

Table 4.3: Risk-neutral Parameter Estimates for the AJD Models

	SV	SVJ	SVCJ	SVJtv	SVCJtv
η_v	0.003 (0.0001)	-	-	-	-
μ_S^Q	-	-3.71 (0.06)	-5.38 (0.13)	-9.20 (0.06)	-9.59 (0.10)
μ_v^Q	-	-	1.34 (0.08)		1.25 (0.06)
RMSE	1.950	1.754	1.957	1.340	1.195

Note: The numbers in parentheses are heteroskedasticity-robust standard errors for nonlinear least squares. RMSEs are root mean squared error in terms of the BSIV (multiplied by 100). The other risk-neutral jump parameters are restricted by the objective parameters such that $\lambda^Q = \lambda$, $\lambda_1^Q = \lambda/\theta$, and $\sigma_S^Q = \sigma_S$ (see Table 4.2).

For the mean jump-in-return size, μ_S^Q , the time-varying jump risk premia models provide larger magnitudes than the constant jump risk premia models: e.g., we obtain -3.7 (SVJ) vs. -9.2 (SVJtv) and -5.4 (SVCJ) vs. -9.6 (SVCJtv). Thus, we see that both constant jump risk premia models exhibit only a mild price jump risk premium (i.e., $\mu_S - \mu_S^Q$), that is, 0.36 and 0.74 for SVJ and SVCJ, respectively. Rather surprisingly, the mean jump-in-volatility sizes, μ_v^Q , for both SVCJ

and SVCJtv are estimated to be smaller in magnitude than the objective-measure counterpart μ_v (equal to 2.1 as in Table 4.2), implying a counter-intuitive negative volatility jump risk premium. In our case, it is likely that a large proportion of jump risk premia is absorbed by μ_S^Q rather than μ_v^Q .

To uncover a source for the poor performances of the constant jump risk premia models, we conduct some experiment as follows. For each model, we estimate risk-neutral parameters and spot volatilities on a daily basis, which provides 433 pairs of spot volatility and risk-neutral parameter vector. This looks similar to the daily recalibration scheme used by Bakshi et al. (1997). For each model, we compute a correlation between the estimated spot volatility (\widehat{V}_t) and mean jump-in-return size ($\widehat{\mu}_{S,t}^Q$). For the SVJ and SVCJ, the correlations are computed to be -0.77 and -0.78, respectively. Thus, a return jump risk premium ($\mu_S - \mu_S^Q$) tends to be large (small) when spot volatility is high (low). This indicates the presence of time-varying jump risk premia correlated with volatility level. In contrast, for both SVJtv and SVCJtv, we obtain only mild correlations between \widehat{V}_t and $\widehat{\mu}_{S,t}^Q$, which are -0.07 and 0.19 for the SVJtv and SVCJtv, respectively. We conduct a similar experiment for the SV model where we compute a correlation between \widehat{V}_t and $\widehat{\eta}_{v,t}$. Likewise, we obtain a relatively large magnitude of correlation (equal to -0.64), which also seems to indicate a possible time-varying feature of volatility risk premium. It seems most likely that the outcomes of our experiment support the importance of time-varying risk premia.

4.5 Empirical Results

4.5.1 Option Pricing Performance Evaluation

Now we evaluate out-of-sample option pricing performances for alternative stock return model specifications. For each AJD models, a model-implied spot volatility,

which is required to price all cross-sectional option contracts for a given date, is extracted from a representative short-dated ATM option, as described in Section 4.4.1. In the case of NGARCH models, conditional variances are available from the past history of returns by their natures, that is, no option information is used for pricing options. As before, we measure option pricing errors by the root-mean squared error (RMSE) between the model-implied and market-observed Black-Scholes implied volatilities (BSIV).

Table 4.4 reports overall pricing errors of each model for the overall out-of-sample (2001-2007), Subsample 1 (2001-2003), and Subsample 2 (2004-2007). For further investigation, Table 4.5 also reports the option pricing errors separated into nine categories according to moneyness and time-to-maturity. The numbers in the parentheses for both tables indicate average pricing errors, which would reveal, if any, a systematic pricing bias. For a pictorial exposition, Figure 4.2 and 4.3 plot model-implied BSIV curves along with market-observed BSIVs across different maturities for two representative days: November 28, 2001 from Subsample 1, and March 9, 2005 from Subsample 2. Those representative days are chosen so that the volatility level on each date (e.g., 25.9 and 12.7, respectively, in terms of the VIX index) is close to the average over the corresponding subsample (25.0 and 14.7, respectively).

Similar to the in-sample results, our out-of-sample analysis shows that modeling time-varying jump risk premia provide a substantial improvement in pricing cross-sectional option prices over time, which is consistent with the result from Pan (2002). Among the models, the SVCJtv model exhibits the smallest RMSE (1.2%), whose magnitudes are close to those from in-sample fit (See Table 4.3). It is worth noting that the SVCJtv improves upon the SVJtv since the SVCJtv model has

rarely been studied in the empirical option studies except for Eraker (2004). His empirical results, however, do not evidence the role of time-varying jump risk premia or jump-in-volatility in pricing options. Contrary to his result, our result supports the roles of both components in describing cross-sectional variations of option prices over time.

It is also clear from Table 4.4 that both time-varying jump risk premia models, the SVJtv and SVCJtv, perform consistently well across both high- and low-volatility subsamples. Furthermore, Table 4.5 indicates that, for the overall sample, both models uniformly dominate the other models for every pair of maturity and moneyness. Particularly, both time-varying jump risk premia models capture the volatility smirk emerging in the ITM categories quite well. It has been documented in the literature that a volatility smirk tends to be less pronounced on low-volatility days. By help of flexible jump-arrival intensity, both models can be easily adapted to any level of smirk. The upper-right panels in Figure 4.2 and 4.3 highlight this adaptability. We, however, find that their performances become worse in the long-maturity OTM category (with the RMSE of 2.1% and 1.9%, respectively). We will revisit this issue shortly.

As reported in Table 4.4, among the constant jump risk premia AJD models (e.g., SV, SVJ, and SVCJ), the SVJ model (with the overall RMSE 1.9%) outperforms the others across both subsamples. The incorporation of jump-in-volatility in SVCJ does not seem to improve upon the SVJ. Although the SVCJ nests the SVJ, the smaller magnitude of σ_S^2 in SVCJ, constrained by its time-series estimate, might perhaps make the SVCJ perform poorer. Our result is different from Broadie et al. (2007), who conclude that the jump-in-volatility provides a modest improvement.

We should mention that, for both SVJ and SVCJ models, there is some difficulty in extracting model-implied spot volatilities from the representative options. For some extremely low-volatility days, we could not even obtain positive model-implied spot volatilities because, on those tranquil days, model-implied (constant) jump risk premia are too large to be reconciled with the market-observed representative option prices. This clearly evidences the time-series inconsistency of constant jump risk premia specifications. To handle this difficulty, we impose a lower limit of 0.001 on model-implied spot volatilities for both SVJ and SVCJ. This limit value is equivalent to as low as 3.2% standard deviation on an annual basis. During our sample period, this value is never reached by model-implied spot volatilities for the SV, SVJtv, and SVCJtv models. Our separate experiment (not reported) has shown that the magnitude of the lower limit has only a minor effect on the pricing performance.

Table 4.4: Option Pricing Errors

	Overall Sample (2001-2007)	Subsample 1 (2001-2003)	Subsample 2 (2004-2007)
BS	3.20 (-1.04)	3.04 (0.03)	3.27 (-1.52)
NGARCH0	3.41 (-0.19)	4.14 (-1.81)	3.03 (0.52)
NGARCH	2.98 (-0.06)	3.85 (-2.40)	2.50 (0.97)
SV	2.20 (-0.55)	2.56 (-1.49)	2.03 (-0.14)
SVJ	1.86 (-0.36)	2.22 (-1.27)	1.68 (0.04)
SVCJ	2.16 (-0.46)	2.89 (-1.78)	1.74 (0.12)
SVJtv	1.33 (0.29)	1.34 (0.00)	1.32 (0.41)
SVCJtv	1.22 (0.19)	1.40 (0.05)	1.14 (0.26)

Note: All pricing errors are measured as the root-mean squared error (RMSE) between the model-implied and market observed Black-Scholes implied volatilities (multiplied by 100). The numbers in parentheses are the average errors.

Table 4.5: Option Pricing Errors by Moneyness-Maturity Categories

		Overall Sample		
		k<0.97	[0.97,1.03]	k>1.03
t<60	BS	5.1 (-4.7)	1.3 (-0.1)	2.9 (2.4)
	NGARCH0	4.7 (-3.6)	2.8 (-0.1)	3.8 (1.9)
	NGARCH	3.5 (-1.8)	2.7 (0.1)	3.2 (0.3)
	SV	3.1 (-2.6)	1.0 (0.0)	2.0 (1.4)
	SVJ	2.2 (-1.6)	0.9 (0.0)	1.8 (0.9)
	SVCJ	2.0 (-1.1)	1.1 (0.1)	1.9 (0.9)
	SVJtv	1.1 (-0.3)	0.8 (0.0)	1.6 (0.1)
	SVCJtv	1.0 (-0.1)	0.7 (0.0)	1.5 (0.0)
	(# of obs.)	1214	1711	1320
60<t<180	BS	4.0 (-3.5)	1.7 (-0.6)	2.6 (1.7)
	NGARCH0	3.3 (-2.0)	2.6 (0.1)	3.3 (1.9)
	NGARCH	2.8 (-0.4)	2.8 (0.3)	2.9 (0.3)
	SV	2.4 (-1.7)	1.6 (-0.2)	1.7 (0.7)
	SVJ	2.1 (-1.4)	1.4 (0.0)	1.5 (0.6)
	SVCJ	2.6 (-1.4)	1.9 (0.2)	1.8 (0.1)
	SVJtv	1.2 (-0.4)	1.2 (0.7)	1.6 (1.1)
	SVCJtv	1.1 (-0.4)	1.1 (0.5)	1.5 (0.8)
	(# of obs.)	788	979	1124
t>180	BS	3.9 (-3.1)	2.6 (-0.9)	2.7 (0.7)
	NGARCH0	2.6 (-0.2)	2.8 (1.0)	3.6 (2.4)
	NGARCH	2.8 (0.6)	3.0 (0.7)	3.1 (0.9)
	SV	2.8 (-1.5)	2.6 (-0.7)	2.5 (0.1)
	SVJ	2.6 (-1.2)	2.3 (-0.5)	2.3 (0.2)
	SVCJ	3.2 (-1.7)	2.9 (-1.0)	2.7 (-0.4)
	SVJtv	1.3 (0.1)	1.6 (1.0)	2.1 (1.7)
	SVCJtv	1.3 (-0.3)	1.4 (0.7)	1.9 (1.4)
	(# of obs.)	767	763	821

Note: All pricing errors are measured as the root-mean squared error (RMSE) between the model-implied and market observed Black-Scholes implied volatilities (multiplied by 100). The numbers in parentheses are the average errors. k denotes moneyness, which is defined as the ratio of strike to underlying forward price, and τ denotes the remaining days to expiration in terms of calendar days.

Table 4.5 (Continued)

		Subsample1		
		k<0.97	[0.97,1.03]	k>1.03
t<60	BS	4.8 (-4.3)	1.3 (-0.1)	2.6 (2.2)
	NGARCH0	6.6 (-5.3)	4.2 (-2.1)	4.0 (0.2)
	NGARCH	5.3 (-3.9)	4.0 (-2.3)	3.9 (-1.8)
	SV	3.2 (-2.8)	1.2 (0.0)	1.7 (1.0)
	SVJ	2.9 (-2.4)	1.0 (0.0)	1.6 (0.8)
	SVCJ	2.8 (-2.2)	1.4 (0.1)	1.9 (0.8)
	SVJtv	1.4 (-0.4)	0.8 (-0.1)	1.6 (-0.6)
	SVCJtv	1.3 (-0.0)	0.7 (-0.1)	1.6 (-0.9)
	(# of obs.)	1214	1711	1320
60<t<180	BS	3.0 (-2.2)	1.9 (0.3)	3.2 (2.4)
	NGARCH0	4.9 (-3.8)	3.6 (-2.0)	2.9 (0.1)
	NGARCH	4.0 (-2.7)	3.7 (-2.5)	3.4 (-2.2)
	SV	3.3 (-3.1)	2.1 (-1.7)	1.3 (-0.7)
	SVJ	2.8 (-2.6)	1.7 (-1.3)	1.2 (-0.5)
	SVCJ	3.8 (-3.5)	2.6 (-2.2)	1.9 (-1.4)
	SVJtv	1.2 (-0.6)	1.1 (0.3)	1.5 (0.7)
	SVCJtv	1.1 (-0.4)	1.1 (0.3)	1.5 (0.5)
	(# of obs.)	788	979	1124
t>180	BS	3.1 (-1.0)	2.9 (0.5)	3.8 (2.2)
	NGARCH0	3.3 (-2.3)	2.7 (-1.2)	2.4 (0.1)
	NGARCH	2.9 (-2.0)	3.0 (-2.0)	2.9 (-1.9)
	SV	4.1 (-3.8)	3.4 (-3.1)	3.8 (-2.3)
	SVJ	3.6 (-3.4)	2.9 (-2.6)	2.8 (-2.0)
	SVCJ	4.8 (-4.5)	4.0 (-3.7)	2.4 (-3.0)
	SVJtv	1.3 (-0.5)	1.3 (0.3)	1.8 (1.1)
	SVCJtv	1.5 (-0.2)	1.6 (0.5)	2.1 (1.2)
	(# of obs.)	767	763	821

Note: All pricing errors are measured as the root-mean squared error (RMSE) between the model-implied and market observed Black-Scholes implied volatilities (multiplied by 100). The numbers in parentheses are the average errors. k denotes moneyness, which is defined as the ratio of strike to underlying forward price, and τ denotes the remaining days to expiration in terms of calendar days.

Table 4.5 (Continued)

		Subsample2		
		k<0.97	[0.97,1.03]	k>1.03
t<60	BS	5.2 (-4.9)	1.3 (-0.1)	3.0 (2.5)
	NGARCH0	3.9 (-3.1)	2.1 (0.6)	3.6 (3.0)
	NGARCH	2.6 (-1.1)	2.1 (1.0)	2.6 (1.6)
	SV	3.0 (-2.5)	1.0 (0.0)	2.2 (1.6)
	SVJ	2.0 (-1.3)	0.9 (0.0)	1.9 (1.0)
	SVCJ	1.7 (-0.7)	1.0 (0.1)	1.9 (1.0)
	SVJtv	1.0 (-0.3)	0.8 (0.1)	1.5 (0.6)
	SVCJtv	0.9 (-0.1)	0.8 (0.1)	1.4 (0.5)
	(# of obs.)	3864	4812	1947
60<t<180	BS	4.3 (-4.0)	1.6 (-0.9)	2.1 (1.3)
	NGARCH0	2.5 (-1.4)	2.1 (0.1)	3.5 (3.0)
	NGARCH	2.2 (0.4)	2.3 (1.4)	2.6 (1.8)
	SV	2.0 (-1.2)	1.3 (0.4)	1.9 (1.5)
	SVJ	1.8 (-0.9)	1.3 (0.5)	1.7 (1.2)
	SVCJ	2.0 (-0.7)	1.5 (0.6)	1.7 (1.1)
	SVJtv	1.2 (-0.3)	1.2 (0.8)	1.7 (1.3)
	SVCJtv	1.1 (-0.5)	1.0 (0.6)	1.5 (1.1)
	(# of obs.)	788	979	1124
t>180	BS	4.2 (-3.9)	2.4 (-1.8)	1.6 (-0.2)
	NGARCH0	2.2 (0.7)	2.9 (2.2)	4.1 (3.7)
	NGARCH	2.8 (1.8)	3.0 (2.2)	3.3 (2.6)
	SV	2.1 (-0.5)	2.0 (0.6)	2.3 (1.5)
	SVJ	2.0 (-0.3)	1.9 (0.7)	2.2 (1.5)
	SVCJ	2.3 (-0.6)	2.1 (0.4)	2.2 (1.2)
	SVJtv	1.3 (0.3)	1.8 (1.3)	2.3 (2.1)
	SVCJtv	1.1 (-0.4)	1.3 (0.7)	1.8 (1.5)
	(# of obs.)	1829	1399	1328

Note: All pricing errors are measured as the root-mean squared error (RMSE) between the model-implied and market observed Black-Scholes implied volatilities (multiplied by 100). The numbers in parentheses are the average errors. k denotes moneyness, which is defined as the ratio of strike to underlying forward price, and τ denotes the remaining days to expiration in terms of calendar days.

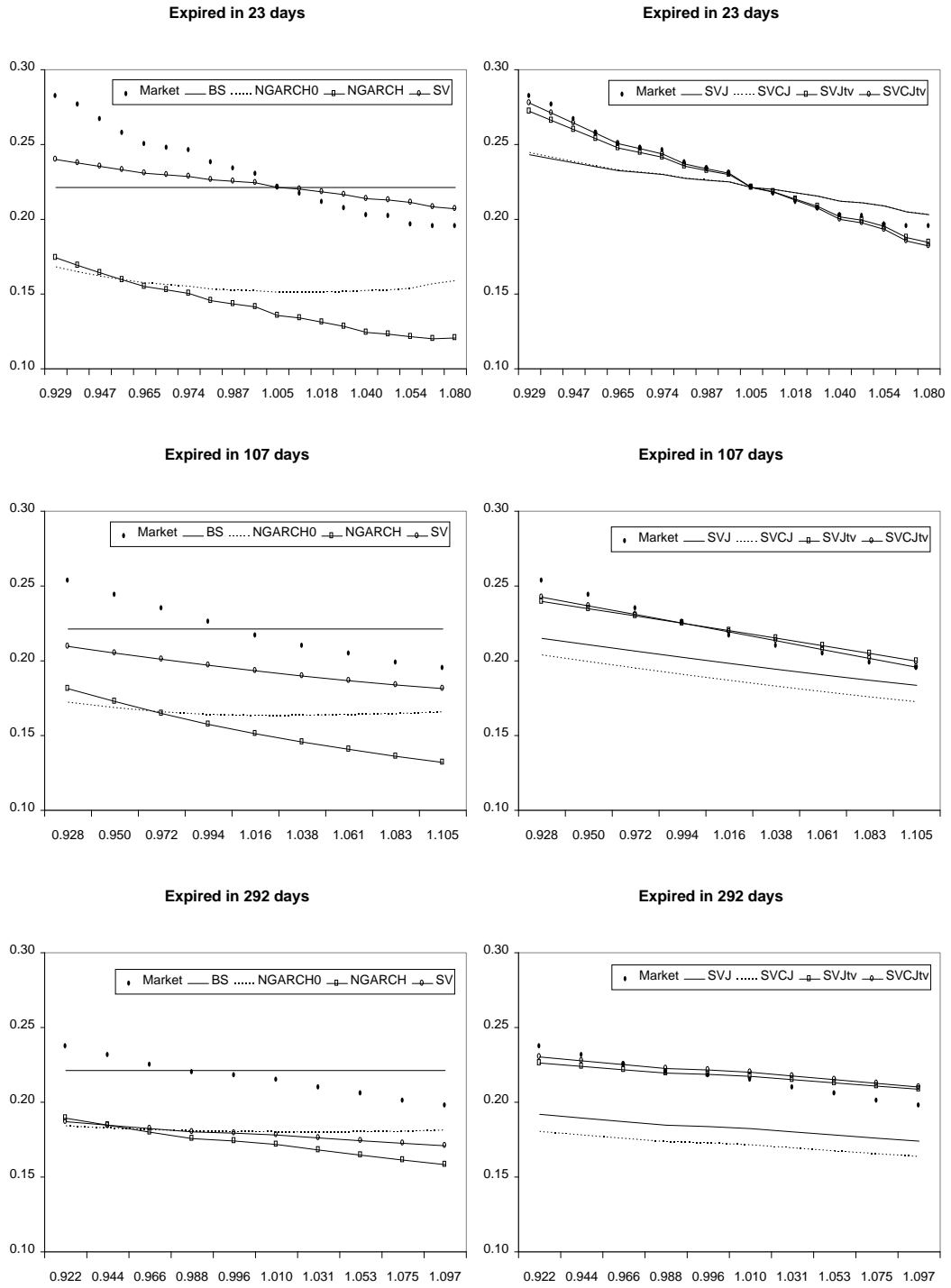


Figure 4.2: Model-implied BSIV's for a High-volatility Day (11/28/2001)

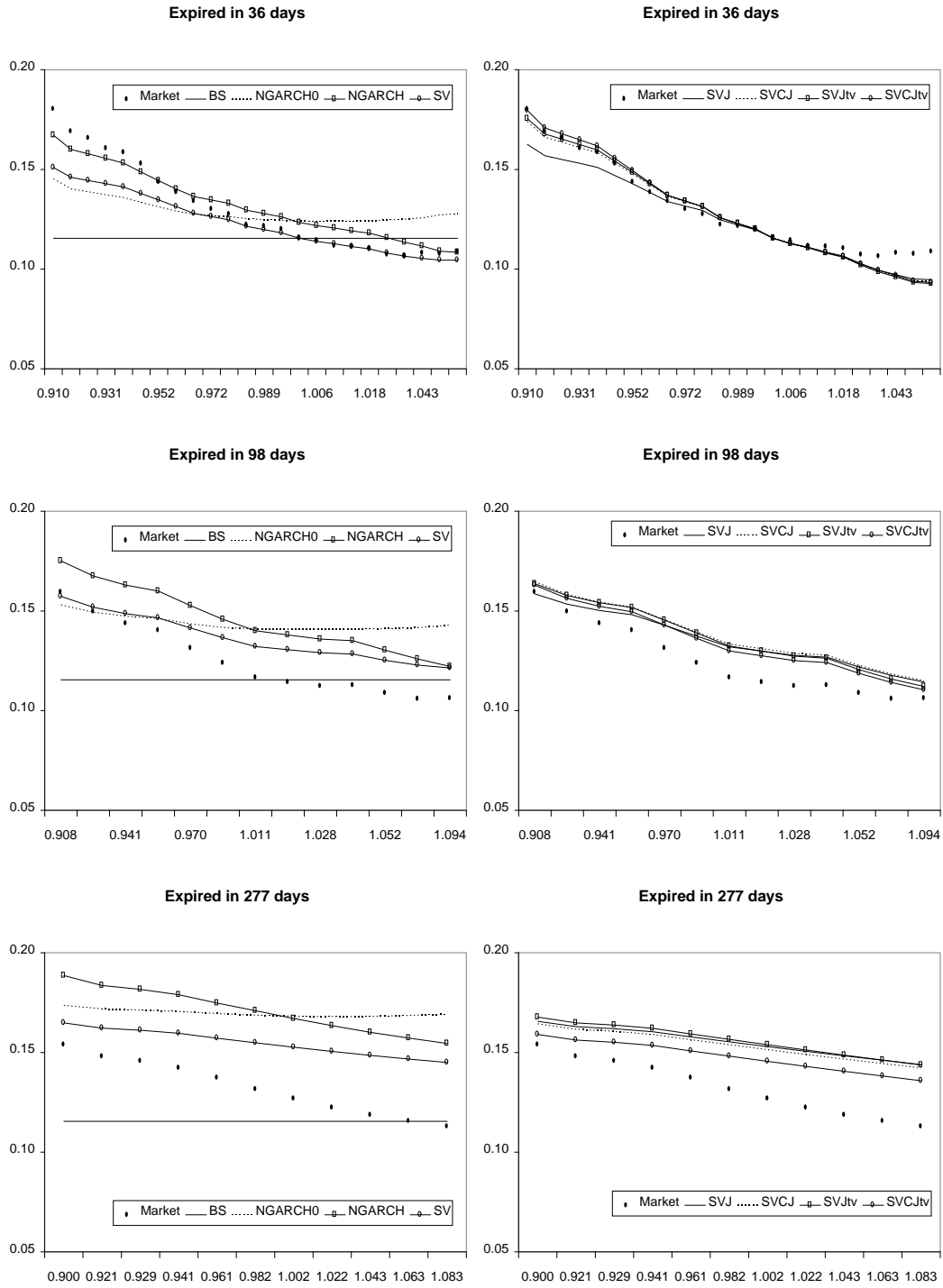


Figure 4.3: Model-implied BSIV's for a Low-volatility Day (3/9/2005)

All models other than the time-varying jump risk premia models provide poor pricing performances in the short-dated ITM category whose volatility-smirk effect is known to be the most substantial. Among those constant jump risk premia models, the SVJ and SVCJ captures the volatility smirk in this category relatively well in the tranquil subsample, whereas they generate too gentle smirk to correctly evaluate the options in the same category in the turbulent subsample. Their inability to consistently capture the smirk at different volatility levels is clearly due to the constant jump risk premia structure. Also, this inability is highlighted by the upper-right panels in Figure 4.2 and 4.3. We see that the failure to capture the smirk arises more severely for the models without jump component (e.g., BS, NGARCH, NGARCH0, and SV). They nearly all tend to underprice the short- and medium-maturity ITM options across both subsamples, and more severely over the high-volatility period.

As noted above, our empirical result is in contrast to Eraker (2004). He argues that the benefit from incorporating price jumps to the AJD specification is only modest. In his empirical study, the pricing performance of the SV is similar even to that of the SVCJtv. In a similar vein, Bakshi et al. (1997) argue that, once stochastic volatility is modeled, adding other features will usually lead to second-order pricing improvements. Our contrasting result is partly due to a different sample period. Our sample period is long enough to cover both high- and low-volatility periods. However, Eraker's (2004) sample period (i.e., 1991-1996) exhibits relatively low-volatility level with the average VIX index of 14.6 (See Figure 4.1). When volatility is low, the role of jump risk premia becomes less pronounced, which might have made the performance of the SV model in Eraker (2004) less distinguishable from that of the SVJ or SVCJ. In fact, our result also shows that

the SV performs better during the low-volatility period (Subsample 2). Similarly, the other models without jump such as the NGARCH0 and NGARCH also show much better performance in Subsample 2.

The lower panels in Figure 4.3 illustrate that all models other than the BS seem to systematically overprice long-maturity options on the low-volatility day. Moreover, most models, except for BS, SVJtv, and SVCJtv, underestimate long-dated option on the high-volatility day as can be seen in Figure 4.2. As pointed out by Chernov and Ghysels (2000), the BSIV, unlike spot volatility, implies the average volatility level over the remaining time to maturity of an option contract. It is likely that our model specifications impose a stronger mean-reversion tendency in the risk-neutral volatility dynamics than what is implied by the market data. One possible explanation for this systematic bias is suggested by Chernov and Ghysels (2000). They argue that long-maturity contracts may reflect long memory, which is not captured by our estimated parameters. Thus, modeling long memory might be a possible consideration to improve pricing performance in the long-maturity categories. Supporting this view, it appears that the BS model provides little systematic bias for long-maturity options. Note that the BS model is an extreme case of long-memory volatility as it assumes a constant volatility. Alternatively, Bates (2000) suggests two factor models from the consideration that one-factor models can do a poor job in capturing the term structures of implicit volatilities over time, based on the evidences from his currency option study (e.g., Bates, 1996). However, one should be careful about choosing a more sophisticated model as it might be subject to an overfitting problem.

Finally, it should be mentioned that both NGARCH models are treated rather unfairly compared to the AJD models. In our study, the NGARCH models eval-

uate options based solely on time-series information without using any option market data. Chernov and Ghysel (2000) provide their option pricing empirical results⁴² that the NGARCH0 (with $\theta = 0$) performs poorly in pricing option contracts and that it is outperformed even by the BS model. They also document that the NGARCH (with nonzero θ) performs slightly worse than the NGARCH0. Unlike their empirical result, our analysis shows that the NGARCH outperforms the NGARCH0 in pricing options. Furthermore, after introducing the leverage effect, the NGARCH prices options better than the BS model, particularly during the low-volatility period. As seen from Figure 4.2, the NGARCH can generate a sufficiently steep "asymmetric" smirk for the BSIV curve in the ITM area during high-volatility days, whereas the NGARCH0 generate only a "symmetric" smile. However, a spot volatility level implied by the NGARCH models is systematically inconsistent with market option prices. During the high (low) volatility period, this model tends to underestimate (overestimate) a volatility level implied by options market. If option market information had been exploited by the NGARCH model like the AJD models, though obviously inconsistent with the model assumption, we might possibly have obtained a better result for this specification.

4.5.2 Density Forecast Performance Evaluation

In what follows, we examine density forecast performances for the stock return models. The forecasting time horizon is set to be one day, which is typically used in practice to compute value-at-risk measure. As mentioned earlier, we employ Hong and Li's (2005) statistics to evaluate each model's one-day-ahead density forecast performance. Table 4.6 reports $\widehat{W}(p)$ statistics at three different lags ($p = 5, 10,$ and 20) for overall sample (2001-2007), Subsample 1 (2001-2003), and Subsample 2

⁴²Chernov and Ghysel (2000) conduct their out-of-sample option pricing analysis for the sample period, November 1993 through October 1994, when the average VIX index is as low as 13.4.

(2004-2007). Also, Figure 4.4 illustrates $\widehat{Q}(j)$ statistics across different lags as well as kernel marginal densities of generalized residuals. To simplify our discussion, we will interpret our empirical result in terms of $\widehat{W}(p)$ at p equal to 5.

For each AJD specification (i.e., SV, SVJ, and SVCJ), we adopt two different density forecasting strategies. First, we use only historical return data to forecast densities. For this purpose, we employ the particle filters, and compute the corresponding Hong and Li (2005) statistics as in Section 4.3.2. Those return-based AJD models are denoted by the SV-PF, SVJ-PF, and SVCJ-PF. Both NGARCH and NGARCH0 models also use only time-series information. Second, we use option market information to forecast densities. That is, given an AJD specification, we extract a model-implied spot volatility from a representative option for each date,⁴³ and then forecast density with the extracted model-implied spot volatility. Those option-based models are denoted by the SV-OPT, SVJ-OPT, and SVCJ-OPT. In addition, the SVJtv and SVCJtv are considered as option-based models equipped with time-varying jump risk premia. We also address the BS-OPT model as a benchmark. For those option-based models, Hong and Li's (2005) statistics are computed by the Fourier-inversion method as in Section 4.3.1.

The success of an option-based model depends on whether its model-implied spot volatilities (filtered from options) are consistent with true volatility level implied by time-series dynamics (e.g., Eraker, 2004). Thus, jump risk premia structure will play an important role in the density forecasting ability by affecting the magnitudes of the spot volatilities. The larger jump risk premia will induce the smaller magnitude of filtered spot volatility.

As reported in Table 4.6, our density forecast analysis indicates that all models

⁴³As before, the representative option at each date is chosen so that its time-to-maturity and moneyness are as close as possible to 30 calendar days and one, respectively.

are, unfortunately, rejected under a conventional significance level (e.g., critical value of 2.33 for 1% significance level). Some time-series studies (not using options data), such as Andersen et al. (2001) and Eraker et al. (2002), could not find evidence of misspecification for jump-equipped AJD models (e.g., SVCJ in Eraker et al. and SVJ in Andersen et al.). Our rejection of all models might be due to a higher power of Hong and Li's (2005) test than existing testing methods.

Table 4.6 shows that the performances of some models differ substantially across two subsamples. In particular, the option-based models seem to show larger differences. Most models perform better in Subsample 1 (high-volatility period). Exceptionally, the SV-OPT model performs better in Subsample 2 (low-volatility period). Interestingly, the SV-OPT outperforms all other models in Subsample 2 ($\widehat{W}(5)$ equal to 9.7), whereas it performs the worst in Subsample 1 ($\widehat{W}(5)$ equal to 15.8). As noted above, the role of jump risk premia is less important when volatility is low. Not distorted by jump risk premia, the SV-OPT model seems to be able to capture a true volatility level during those tranquil times.

Table 4.6: $\widehat{W}(p)$ Statistics

Lags (p)	Overall			Subsample 1			Subsample 2		
	5	10	20	5	10	20	5	10	20
BS-OPT	32.9	45.3	63.5	7.7	10.3	15.9	28.6	39.7	57.1
NGARCH0	24.4	31.0	43.0	5.7	6.6	9.6	25.3	33.8	48.4
NGARCH	18.9	24.3	33.6	2.0	2.1	2.6	25.7	34.2	49.5
SV-PF	10.9	12.4	16.2	5.8	7.6	11.1	13.5	16.3	22.7
SVJ-PF	9.8	11.7	15.8	3.8	4.7	7.3	12.1	14.9	21.0
SVCJ-PF	11.5	13.4	18.1	9.6	13.1	19.3	12.2	14.5	20.5
SV-OPT	17.9	24.0	33.9	15.8	21.7	32.2	9.7	13.1	19.0
SVJ-OPT	9.8	12.8	18.4	8.9	12.3	19.0	21.1	30.2	42.9
SVCJ-OPT	12.2	16.4	23.7	15.1	21.0	31.1	28.3	40.3	57.6
SVJtv	20.6	27.2	38.1	6.8	9.3	14.9	16.7	22.0	30.5
SVCJtv	21.8	29.0	40.7	6.8	9.3	14.9	18.3	24.4	33.9

Note: The statistics are computed by $\widehat{W}(p) = \frac{1}{\sqrt{p}} \sum_{j=1}^p \widehat{Q}(j)$.

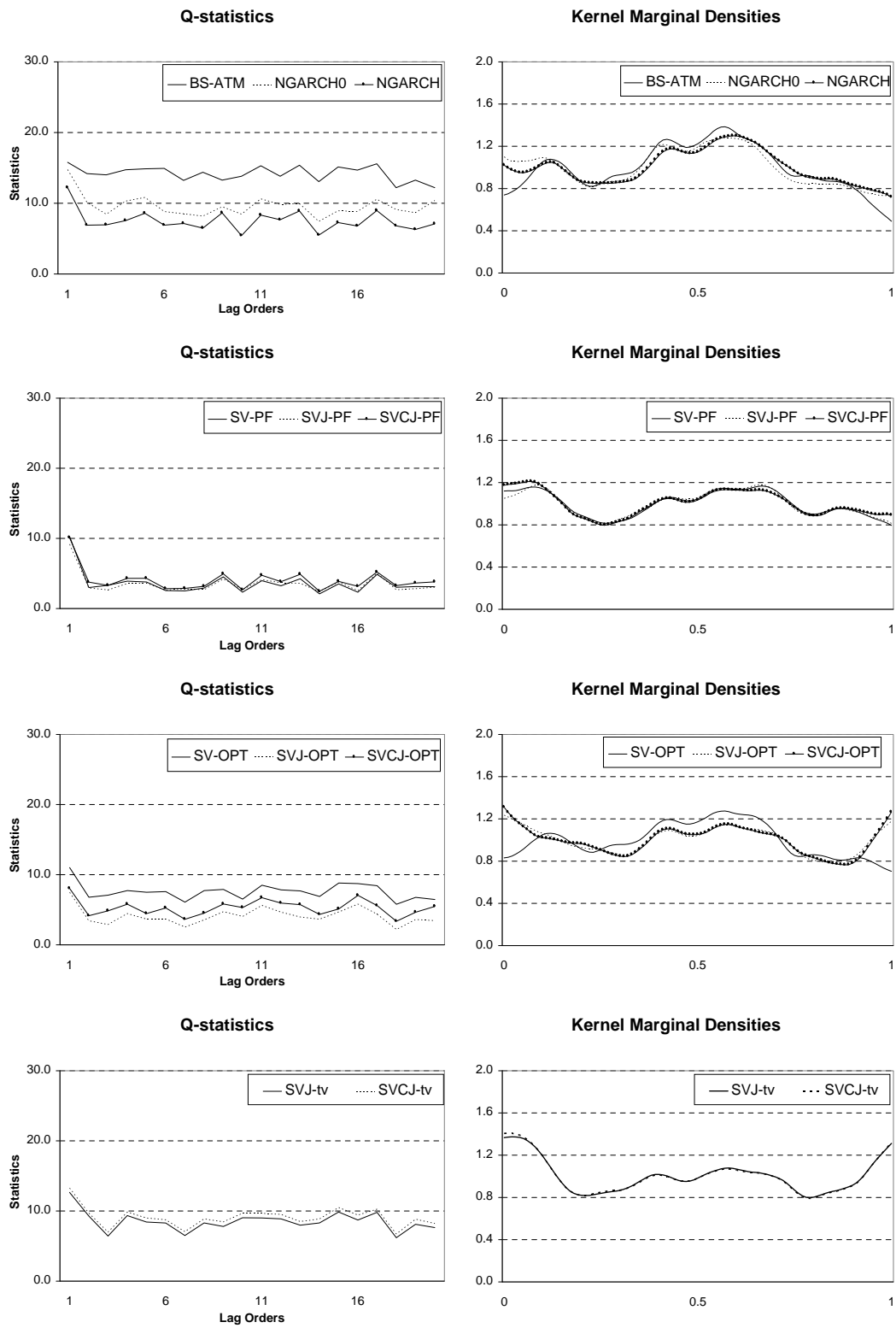


Figure 4.4: $\hat{Q}(j)$ Statistics and Kernel Marginal Densities

As a benchmark, we examine the density forecast performance for the BS-OPT model. It performs the worst among the models for overall sample (with $\widehat{W}(5)$ equal to 32.9). The upper-right panel in Figure 4.4 indicates that its kernel marginal density of generalized residuals are lower than unity at both tails, suggesting that there are fewer observations around the tails than are predicted by the model. As has long been documented in the literature, this result implies that the Black-Scholes implied volatility tends to be too high to be reconciled with true historical volatility level. Thus, the BS model is clearly inconsistent with time-series dynamics. To reconcile both time-series and risk-neutral dynamics, one possible solution might be to reduce the magnitude of volatility by introducing additional risk premia or stochastic volatility into the model. The remaining option-based models will be examined in this context.

For the overall sample, both return- and option-based SVJ models (i.e., SVJ-PF and SVJ-OPT) turn out to be the least misspecified models with their $\widehat{W}(5)$ statistics equal to 9.8 (Table 4.6). They are followed by two other return-based models (i.e., SV-PF and SVCJ-PF). Contrary to our option pricing results, the constant jump risk premia model (i.e., SVJ-OPT) turns out to forecast densities better than both time-varying jump risk premia models (i.e., SVJtv and SVCJtv). However, a further investigation reveals that the SVJ-OPT does very poorly during Subsample 2 (8.9 for Subsample 1 \rightarrow 21.2 for Subsample 2 in terms of $\widehat{W}(5)$). This worse performance also occurs to the other constant jump risk premia model, that is, SVCJ-OPT. In fact, their kernel marginal densities of generalized residuals for Subsample 2 (not reported here) exhibit highly pronounced peaks in both ends of the densities. That is, those models are underestimating tail risk during the tranquil times. In order to be reconciled with time-series data, larger magnitudes

of spot volatilities should have been extracted from option prices during the low-volatility period. However, it is obviously impossible for those models having a constant nature of jump risk premia. Also, recall that we could not even extract positive spot volatilities from options for some extremely low-volatility dates.⁴⁴ In contrast, the return-based AJD models (i.e., SV-PF, SVJ-PF, and SVCJ-PF) seem to exhibit relatively stable performances across both subsamples.

In our analysis, the time-varying jump risk premia models (i.e., SVJtv and SVCJtv) perform poorly in forecasting densities. Their $\widehat{W}(5)$ statistics are as high as 20.6 and 21.8, respectively, for overall sample. This result is rather surprising as the time-varying jump risk premia have caused a remarkable success in describing cross-sectional option prices. It seems that their time-varying specifications are not consistent with time-series dynamics, though consistent with risk-neutral dynamics implied by cross-sectional options data. As shown in Figure 4.4, their kernel marginal densities tend to underestimate tail risks. A further investigation reveals that, contrary to the BS-OPT, they consistently underestimate tail risks across both subsamples (not reported here), that is, their model-implied spot volatilities extracted from options data are too low to be consistent with true volatility level implied by time-series dynamics. Perhaps, smaller jump risk premia for the SVJtv and SVCJtv (e.g., smaller magnitude of λ_1 or μ_S^Q) would have provided a better result.⁴⁵

Our mixed results (between option pricing and density forecast evaluations) for the role of time-varying jump risk premia might be because the options market is

⁴⁴Similar to our option pricing study, we impose an artificial lower limit of spot volatility (e.g., 0.001) for some extremely low volatility dates.

⁴⁵As a separate experiment, we evaluate the density forecast performances for the SVJtv and SVCJtv with different risk-neutral mean jump-in-return size. We artificially reduce the magnitude of μ_S^Q by 3%p for both models. As a result, we obtain the $\widehat{W}(5)$ statistics of 10.3 for both SVJtv and SVCJtv, accompanied by mildly worse option pricing performances (the overall RMSE of 1.4% for both SVJtv and SVCJtv).

somehow segmented from the spot market. Pan (2002) argues that this segmentation can occur due to some option-specific factors such as liquidity. Otherwise, it might be that the AJD models are misspecified. However, it is hard to tell which one dominates. We should conclude that, among our AJD models, we could not find a specification that reconciles both time-series and risk-neutral dynamics.

Finally, it is notable that the return-based AJD models perform better than both NGARCH models which are also return-based. Furthermore, the performances of the return-based AJD models are more stable across two subsamples. In practice, the GARCH-type models, including RiskMetrics, have been widely used in value-at-risk implementation. Our results indicate that the AJD models can be considered as an alternative tool to implement value-at-risk.

4.6 Conclusion

This chapter has addressed out-of-sample option pricing and density forecast performances for the affine jump diffusion (AJD) models by using the S&P 500 stock index and associated option contracts. For comparison purposes, we have also considered nonlinear asymmetric GARCH (NGARCH) models, which provide a convenient option pricing method among the GARCH-type models.

We have evaluated one-day-ahead density forecast performance for each model by using Hong and Li's (2005) testing method. Our study has made methodological contributions by extending the existing Hong and Li's (2005) testing method to be applicable to the famous AJD class of models, whether or not model-implied spot volatilities are available. For either case, we propose (i) the Fourier inversion of the closed-form conditional characteristic function and (ii) the Monte Carlo integration based on the particle filters proposed by Johannes et al. (2008).

From our option pricing analysis, we have found strong evidence in favor of

time-varying jump risk premia. In particular, the SVCJtv, which has rarely been studied in the literature, are the most successful in fitting cross-sectional option prices over time. Both time-varying jump risk premia models (i.e., SVJtv and SVCJtv) dominate the other models uniformly across both high- and low-volatility periods and for every maturity-moneyness category. However, we find that all AJD models somehow tend to systematically underprice (overprice) long-dated options on high-volatility (low-volatility) days, indicating that our AJD models impose a stronger mean-reversion tendency in the risk-neutral volatility dynamics than what is implied by long-dated market option prices. As Chernov and Ghysels (2000) argue, modeling long memory into risk-neutral volatility dynamics might be able to correct the pricing bias for long-maturity options. This could be an interesting issue that will be addressed in future research.

However, in terms of density forecast, we could not find an AJD specification that reconciles the dynamics implied by both time series and options data across both subsamples. Our density forecast analysis shows that all models are somehow misspecified in terms of Hong and Li's (2005) statistics. However, we find that the return-based AJD models (particularly, SVJ-PF), by using the particle filters, exhibit relatively good and stable density forecasting performances across both subsamples compared to the option-based counterparts. The time-varying jump risk premia in both SVJtv and SVCJtv, which have drastically improved option pricing performance, seem to be inconsistent with time-series dynamics. Their model-implied spot volatilities extracted from options data turns out to be systematically lower than what is implied by time-series data.

Finally, we should mention that our density forecasting analysis focuses only on one-day-horizon density forecast. Such a short time-horizon forecasting is a major

interest of financial institutions who manage a liquid market portfolio. However, there may be some cases where a longer horizon density forecast is more useful. For instance, central banks are interested in forecasting long-term densities for interest rates, exchange rates, and stock market indices in order to obtain information on business cycles. Because of its forward-looking nature, option market information might help to assess future economic conditions. Also, a density forecasting ability of a given model might vary across different time horizons. In this sense, a longer time-horizon density forecasting evaluation for a stock return model would be an interesting topic for future research.

APPENDICES

Appendix 4.1: Dynamic Probability Integral Transform via Fourier Inversion for the SVCJ Model

Suppose that a conditional moment generating function of some AJD process in (4.2) is available such as $\phi_{t,T}(u) = E[e^{u \ln S_T} | \mathcal{F}_t]$. I_t denotes information set at time t , S_T denotes a random variable for asset price at time T , and s_T denotes a realized asset price at time T . We follow a usual proof method used in the option pricing literature.

$$z_{t,T} = \int_{-\infty}^{s_T} p(s | I_t) ds = 1 - E[1_{[S_T \geq s_T]} | \mathcal{F}_t].$$

Then, we have

$$\begin{aligned} E[1_{[S_T \geq s_T]} | \mathcal{F}_t] &= E[1_{[\ln S_T \geq \ln s_T]} | \mathcal{F}_t] \\ &= E\left[\frac{1}{2} + \frac{1}{2} \text{sign}\left(\ln \frac{S_T}{s_T}\right) \mid \mathcal{F}_t\right] \\ &= E\left[\frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin\left(u \ln \frac{S_T}{s_T}\right)}{u} du \mid \mathcal{F}_t\right] \\ &= E\left[\frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\sin\left(u \ln \frac{S_T}{s_T}\right)}{u} du \mid \mathcal{F}_t\right] \\ &= E\left[\frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{e^{iu \ln \frac{S_T}{s_T}} - e^{-iu \ln \frac{S_T}{s_T}}}{2iu} du \mid \mathcal{F}_t\right] \\ &= E\left[\frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}\left[e^{iu \ln \frac{S_T}{s_T}}\right]}{u} du \mid \mathcal{F}_t\right] \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}\left[E\left(e^{iu \ln \frac{S_T}{s_T}} \mid \mathcal{F}_t\right)\right]}{u} du \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}\left[E\left(e^{iu \ln S_T} \mid \mathcal{F}_t\right) e^{-iu \ln s_T}\right]}{u} du, \end{aligned}$$

where $\text{sign}(\theta) = 1$ for $\theta > 0$, $\text{sign}(\theta) = -1$ for $\theta < 0$, and $\text{sign}(\theta) = 0$ for $\theta = 0$, and, in the third equality, we use

$$\lim_{T \rightarrow \infty} \int_{-T}^T \frac{\sin(\theta t)}{t} dt = \pi \cdot \text{sign}(\theta).$$

As a result, we have that $z_{t,T} = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}[\phi_{t,T}(iu)e^{-iu \ln s_T}]}{u} du$.

Appendix 4.2: Option Pricing Formula for the AJD Class of Models

This appendix further details the option pricing formula that shows up in Section 4.4. The underlying of the option contract is a forward-price. We consider the SVCJ model. The SV and SVJ models are just a simpler case of the SVCJ model. The content of this appendix is based on Bates (1996), and Duffie et al. (2000).

In Equation (4.28) from Section 4.4, the evaluation of P_1 and P_2 ⁴⁶ involves the integration of the function involving the CMGF, $\phi_j(u)$ for $j = 1, 2$. For the process in (4.26), the CMGF has a closed-form as follows.

$$\begin{aligned} \phi_j(u) = & \exp[u \ln S_t + A_j + B_j V_t + \lambda^Q C_j (1 + \bar{k})^{\mu_j + 1/2} (1 + \bar{k})^u e^{\sigma_S^2 (\mu_j u + u^2/2)} - \\ & \lambda^Q \tau (1 + \bar{k})^{\mu_j + 1/2}], \end{aligned}$$

⁴⁶ P_2 is a probability that S_T is greater than X under the risk-neutral measure, that is, $P_2 = E^Q[1_{[S_T > X]} | \mathcal{F}_t]$. On the other hand, P_1 is the probability of the same event, but under some forward measure, that is, $P_1 = E^* [1_{[S_T > X]} | \mathcal{F}_t] = E^Q \left[\frac{S_T}{E^Q[S_T]} 1_{[S_T > X]} | \mathcal{F}_t \right]$, where $*$ denotes the forward measure.

where

$$\begin{aligned}
A_j(u, \tau) &= -\lambda^Q \bar{k} u \tau - \frac{\kappa^Q \theta^Q \tau}{\sigma_v^2} (\rho \sigma_v u - \beta_j - \gamma_j) \\
&\quad - \frac{2\kappa^Q \theta^Q}{\sigma_v^2} \ln \left[1 + \frac{1}{2} (\rho \sigma_v u - \beta_j - \gamma_j) \frac{1 - e^{\gamma_j \tau}}{\gamma_j} \right], \\
B_j(u, \tau) &= -2 \frac{\mu_j u + \frac{1}{2} u^2}{\rho \sigma_v u - \beta_j + \gamma_j \frac{1 + e^{\gamma_j \tau}}{1 - e^{\gamma_j \tau}}}, \\
C_j(u, \tau) &= \frac{\gamma_j - \rho \sigma_v u + \beta_j}{\gamma_j - \rho \sigma_v u + \beta_j - 2\mu_v (\mu_j u + u^2/2)} \cdot \tau \\
&\quad + \frac{4(\mu_j u + u^2/2)\mu_v}{\gamma_j^2 - (\rho \sigma_v u - \beta_j + 2\mu_v (\mu_j u + u^2/2))^2} \\
&\quad \ln \left(1 - \frac{1}{2} \frac{\gamma_j + \rho \sigma_v u - \beta_j + 2\mu_v (\mu_j u + u^2/2)}{\gamma_j} (1 - \exp(-\gamma_j \tau)) \right)
\end{aligned}$$

and

$$\begin{aligned}
\beta_1 &= \kappa_v - \rho \sigma_v, \quad \beta_2 = \kappa_v, \quad \mu_1 = \frac{1}{2}, \quad \mu_2 = -\frac{1}{2}, \\
\gamma_j &= \sqrt{(\rho \sigma_v u - \beta_j)^2 - 2\sigma_v^2 (\mu_j u + \frac{1}{2} u^2)}, \\
\bar{k} &= \exp(\mu_S + \frac{1}{2} \sigma_S^2) - 1.
\end{aligned}$$

The remaining procedure is to compute an option price through (4.27) and (4.28). For numerical integration, Gauss-Legendre quadrature can be used as explained in Section 4.2.

REFERENCES

- [1] Ahn, D., and B. Gao, 1999, "A Parametric Nonlinear Model of Term Structure Dynamics," *Review of Financial Studies*, 12, 721-762.
- [2] Ait-Sahalia, Y., 1996, "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385-426.
- [3] Ait-Sahalia, Y., 2002a, "Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach," *Econometrica*, 70, 223-262.
- [4] Ait-Sahalia, Y., 2002b, "Closed-Form Expansion for Multivariate Diffusions," working paper, Princeton University.
- [5] Akgiray, V., and G. Booth, 1986, "Stock Price Processes with Discontinuous Time Paths: An Empirical Examination," *The Financial Review*, 21, 163-184.
- [6] Andersen, T., L. Benzoni, and J. Lund, 2002, "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239-1284.
- [7] Andersen, T., H. Chung, and B. Sorensen, 1999, "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study," *Journal of Econometrics*, 91, 61-87.
- [8] Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003-2049.
- [9] Ball, C.A., and W.T. Torous, 1985, "On Jumps in Common Stock Prices and their Impact on Call Option Pricing," *Journal of Finance*, 40, 155-173.

- [10] Bates, D., 1996, "Jumps and Stochastic Volatility: Exchange Rate Process Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69-107.
- [11] Bates, D., 2000, "Post-'87 Crash Fears in the S&P500 Futures Option Market," *Journal of Econometrics*, 94, 181-238.
- [12] Bera, A.K., and A. Ghosh, 2002, "Density Forecast Evaluation with Applications in Finance," unpublished manuscript, University of Illinois at Urbana-Champaign, Dept. of Economics.
- [13] Berg, A., R. Meyer, and J. Yu, 2004, "Deviance information Criterion for Comparing Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 22, 107-120.
- [14] Bhardwaj, G., V. Corradi, V., and N.R. Swanson, 2007, "A Simulation Based Specification Test for Diffusion Processes," *Journal of Business and Economic Statistics*, forthcoming.
- [15] Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- [16] Bollerslev, T., and H.O. Mikkelsen, 1996, "Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 73, 151-184.
- [17] Brandt, M., and P. Santa-Clara, 2002, "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets," *Journal of Financial Economics*, 63, 161-210.

- [18] Broadie, M., M. Chernov, and M. Johannes, 2007, "Model Specification and Risk Premia: The Evidence from Futures Options," *Journal of Finance*, 62, 1453-1490.
- [19] Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209-1227.
- [20] Chen, S.X. and J. Gao, 2005, "On the Use of the Kernel Method for Specification Tests of Diffusion Models," Working Paper, University of Western Australia.
- [21] Chen, S.X. and J. Gao, 2007, "An Adaptive Empirical Likelihood Test for Time Series Models," *Journal of Econometrics*, forthcoming.
- [22] Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen, 2003, "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225-257.
- [23] Chernov, M. and E. Ghysels, 2000, "Estimating of Stochastic Volatility Models for the Purpose of Option Pricing," *Computational Finance* 1999.
- [24] Chernov, M. and E. Ghysels, 2000, "A Study towards a Unified Approach to the Joint Estimation of the Objective and Risk Neutral Measures for the Purpose of Option Valuation," *Journal of Financial Economics*, 56, 407-458.
- [25] Chib, S., 1995, "Marginal Likelihood from a Gibbs Output," *Journal of the American Statistical Association*, 90, 1313-1321.
- [26] Christoffersen, P., and F.X. Diebold, 1996, "Further Results on Forecasting and Model Selection Under Asymmetric Loss," *Journal of Applied Econometrics*, 11, 561-72.

- [27] Clark, P.K., 1973, "A subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135-155.
- [28] Clements, M.P., and J. Smith, 2000, "Evaluating the Forecast Densities of Linear and Non-Linear Models: Applications to Output Growth and Unemployment," *Journal of Forecasting*, 19, 255-276.
- [29] Cochrane, J., 2005, *Asset Pricing*, Princeton University Press, Princeton, NJ.
- [30] Corradi, V., 2000, "Reconsidering the Continuous-Time Limit of the GARCH(1,1) Process," *Journal of Econometrics*, 96, 145-153.
- [31] Corradi, V. and N.R. Swanson, 2005, "A Bootstrap Specification Test for Diffusion Processes," *Journal of Econometrics*, 124, 117-148.
- [32] Corradi, V. and N.R. Swanson, 2006, "Simulation Based Predictive Density Estimation and Testing for Diffusion Processes," Working Paper, Rutgers University and Warwick University.
- [33] Cox, J.C., J.E. Ingersoll, and S.A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.
- [34] Das, S.R., 2002, "The Surprise Element: Jump in Interest Rate," *Journal of Econometrics*, 106, 27-65.
- [35] Das, S.R., and R. Uppal, 1998, "International Portfolio Choice with Systemic Risk," Working Paper, Harvard Business School and MIT.
- [36] Diebold, F.X., T. Gunther, and A. Tay, 1998, "Evaluating Density Forecasts with Application to Financial Risk Management," *International Economic Review*, 39, 863-883.

- [37] Diebold, F.X., J. Hahn, and A. Tay, 1999, "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns of Foreign Exchange," *Review of Economics and Statistics*, 81, 661-673.
- [38] Duan, J., 1995, "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13-32.
- [39] Duffie, D., J. Pan, and K. Singleton, 2000, "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343-1376.
- [40] Elerian, O., S. Chib, and N. Shephard, 2001, "Likelihood Inference for Discretely Observed Nonlinear Diffusions," *Econometrica*, 69, 959-993.
- [41] Engle, R., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.
- [42] Eraker, B., 2004, "Do Stock Prices and Volatility Jump? Reconciling Evidences from Spot and Option Prices," *Journal of Finance*, 59, 1367-1403.
- [43] Eraker, B., M. Johannes, and N. Polson, 2003, "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 53, 1269-1300.
- [44] Gallant, A.R., and G. Tauchen, 1996, "Which Moments to Match?," *Econometric Theory*, 12, 657-681.
- [45] Gallant, A.R., and G. Tauchen, 1998, "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions," *Journal of American Statistical Association*, 93, 10-24.
- [46] Gao, J., and M.L.King, 2004, "Adaptive Testing in Continuous-Time Diffusion Models," *Econometric Theory*, 20, 844-882.

- [47] Granger, C., 1999, "The Evaluation of Econometric Models and of Forecasts," invited lecture at the Far Eastern Meeting of the Econometric Society, Singapore.
- [48] Granger, C., and M.H. Pesaran, 2000, "A Decision Theoretic Approach to Forecasting Evaluation," in *Statistics and Finance: An Interface*, eds. Chan W.S., W.K.Li and H.Tong, London: Imperial College, pp.261-278
- [49] Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility, with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-344.
- [50] Hong, Y., and H. Li, 2005, "Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37-84.
- [51] Hong, Y., H. Li and F. Zhao, 2004, "Out-of-Sample Performance of Discrete-Time Spot Interest Rate Models," *Journal of Business & Economic Statistics*, 22, 457-473.
- [52] Hong, Y., H. Li and F. Zhao, 2007, "Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates," *Journal of Econometrics*, 141, 736-776.
- [53] Hull, J., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 52, 281-300.
- [54] Jacquier, E., N. Polson, and P. Rossi, 1994, "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economics Statistics*, 12, 70-87.
- [55] Jarrow, R., and E. Rosenfield, 1984, "Jump Risks and the Intertemporal Capital Asset Pricing Model," *Journal of Business*, 57, 337-351.

- [56] Johannes, M., and N. Polson, 2005, "MCMC for Financial Econometrics," in Y. Ait-Sahalia and L. Hansen (Eds.), *Handbook of Financial Econometrics*, Amsterdam: Elsevier-North-Holland.
- [57] Johannes, M., N. Polson, and J. Stroud, 2008, "Learning about Jumps and Stochastic Volatility: Filtering Stochastic Differential Equations with Jumps," *Review of Financial Studies*, forthcoming.
- [58] Kallsen, J., and M.S. Taqqu, 1998, "Option Pricing in ARCH-type models," *Mathematical Finance*, 8, 13-26.
- [59] Kim, S., N. Shephard, and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economics Studies*, 65, 361-394.
- [60] Li, F., and G. Tkacz, 2006, "A Consistent Bootstrap Test for Conditional Density Functions with Time-Series Data," *Journal of Econometrics*, 92, 101-147.
- [61] Li, F., 2007, "Testing the Parametric Specification of the Diffusion Function in a Diffusion Process," *Econometric Theory*, 23, 221-250.
- [62] Melino, A., and Turnbull, 1990, "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics*, 45, 239-265.
- [63] Merton, R.C., 1976, "Option Pricing when Underlying Stock Returns are discontinuous," *Journal of Financial Economics*, 3, 125-144.
- [64] Meyer, R., and J. Yu, 2000, "BUGS for a Bayesian Analysis of Stochastic Volatility Models," *Econometrics Journal*, 3, 198-215.

- [65] Nelson, D., 1990, "ARCH Models as Diffusion Approximation," *Journal of Econometrics*, 45, 7-39.
- [66] Pan, J., 2002, "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3-50.
- [67] Pedersen, A.R., 1994, "Uniform Residuals for Discretely Observed Diffusion Processes," *Technical Report 292*, Department of Statistics, Institute of Mathematics, University of Aarhus.
- [68] Pedersen, A.R., 1995, "A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete observations," *Scandinavian Journal of Statistics*, 22, 55-71.
- [69] Pitt, M., and Shephard, N., 1999, "Filtering via simulation: Auxiliary Particle Filter," *Journal of American Statistical Association*, 94, 590-599.
- [70] Press, S.J., 1967, "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317-335.
- [71] Rosenberg, B., 1972, "The Behavior of Random Variables with Nonstationary Variance and the Distribution of Security Prices," University of California, Berkeley.
- [72] Rosenblatt, M., 1952, "Remarks on a multivariate transformation," *Annals of Mathematical Statistics*, 23, 470-472.
- [73] Scott, D.W., 1992, *Multivariate Density Estimation, Theory, Practice, and Visualization*, John Wiley and Sons, New York.
- [74] Singleton, K., 2006, *Empirical Dynamic Asset Pricing*, Princeton University Press, Princeton, NJ.

- [75] Taylor, S., 1982, "Financial Returns Modelled by the Product of Two Stochastic Processes - a Study of Daily Sugar Prices 1961-79," In Anderson, O.D., (ed.), *Time Series Analysis: Theory and Practice. 1* 203-26, Amsterdam: North-Holland.
- [76] Taylor, S., 1986, *Modelling Financial Time Series*, John Wiley and Sons, London.
- [77] Thompson, S.B., 2002, "Specification Tests for Continuous Time Models," Working Paper, Harvard University.
- [78] Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177-188.
- [79] Yu, J., 2005, "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics*, 127, 165-178.

Chapter 5

Conclusion

In this dissertation, by extending the existing Hong and Li's (2005) testing method, we have developed several specification testing methods applicable to popular asset pricing models. We focus on how to implement a dynamic probability integral transform under different situations. Chapter 2 has introduced simulation method to implement Hong and Li's (2005) transition density-based test for the case where there is no closed-form transition density. Our Monte Carlo study shows that the proposed simulation test has very similar sizes and powers to Hong and Li's (2005) test using the closed form of the transition density. Moreover, the performance of the test is robust to the choice of the number of simulation iterations and the number of discretization steps between adjacent observations. In Chapter 3 and 4, we have proposed the testing method applicable to various stochastic volatility models including the affine jump diffusion models, whether or not model-implied spot volatilities are available. When they are not available, we have proposed the Monte Carlo integration based on the particle filtering approach (e.g., Johannes et al., 2008). On the other hand, when model implied spot volatilities are available from options data for the famous affine jump diffusion models, one can use the Fourier inversion of their closed-form conditional characteristic function proposed in Chapter 4.

Based on the proposed testing methods, we conduct a comprehensive empirical studies on some popular stock return models by using the S&P 500 index and the associated option contracts (Chapter 3 and 4). Our time-series study in Chapter 3 shows that all models are misspecified in terms of Hong and Li (2005) statistics. However, among the models considered, the stochastic volatility models perform

relatively well in both in- and out-of-samples. We also find that modeling leverage effect provides a substantial improvement in the log stochastic volatility models. Our value-at-risk performance analysis also supports stochastic volatility models rather than GARCH models. Our results provide a practical implication that stochastic volatility models can be a possible alternative to the widely used GARCH models in the VaR implementation.

In Chapter 4, we extend our study to the risk-neutral dynamics implied by the AJD models in the out-of-sample context. Especially, we have focused on the role of time-varying jump risk premia. Our empirical option pricing analysis shows strong evidence in favor of time-varying jump risk premia. We, however, find that, given our AJD specifications, option market information is inconsistent with time-series dynamics in terms of density forecast. Among our AJD models, we could not find a specification that successfully reconcile the dynamics implied by both time-series and options data across both subsamples. Our mixed results (between option pricing and density forecast evaluations) for the role of time-varying jump risk premia might be because the options market is somehow segmented from the spot market due to some option-specific factors such as liquidity. Otherwise, it might be that the AJD models are misspecified, although it is hard to tell which one dominates the other.

REFERENCES

- [1] Johannes, M., N. Polson, and J. Stroud, 2008, “Learning about Jumps and Stochastic Volatility: Filtering Stochastic Differential Equations with Jumps,” *Review of Financial Studies*, forthcoming.
- [2] Hong, Y., and H. Li, 2005, “Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates,” *Review of Financial Studies*, 18, 37-84.