Integrated Time-Series Analysis of Spot and Option Prices

Jun Pan*

Graduate School of Business
Stanford University

First draft : October 15, 1998
This draft : November 28, 1999

Abstract: This paper examines the joint time series of the S&P 500 index and near-the-money short-dated option prices with an arbitrage-free model, capturing both stochastic volatility and jumps. Jump-risk premia uncovered from the joint data respond quickly to market volatility, becoming more prominent during volatile markets. This form of jump-risk premia is important not only in reconciling the dynamics implied by the joint data, but also in explaining the volatility "smirks" of cross-sectional options data. Further diagnostic tests suggest a stochastic-volatility model with two factors — one strongly persistent, the other quickly mean-reverting and highly volatile.

*Graduate School of Business, Stanford University, Stanford, CA 94305. Email: junpan@stanford.edu.
I benefited from discussions with professors and doctoral students in the finance program at Stanford. I am especially grateful to my advisers Darrell Duffie and Ken Singleton, who inspired me to take on this project, and helped push it all the way through with their insightful comments and warm encouragement. I also benefited from extensive discussions with Jun Liu, and valuable comments from Andrew Ang, Joe Chen, Mark Ferguson, Peter Glynn, and Monika Piazzesi. This paper can be downloaded at `http://www.stanford.edu/~junpan/`.

1 Introduction

This paper offers an integrated time-series study of the joint transition behavior of the S&P 500 index $\{S_t\}$ and near-the-money short-dated option prices $\{C_t\}$. Examining these joint spot and option data through an arbitrage-free jump-diffusion model, we find compelling evidence of a jump-risk premium that responds quickly to market volatility — becoming more prominent during volatile markets. This jump-risk premium is important, not only in reconciling the dynamics implied by the time series $\{S_t, C_t\}$, but also in explaining changes over time of the "smiles" and "smirks" found in cross-sectional options data.

The joint time-series data on spot and option prices strongly reject pure-diffusion models (such as the Heston [1993] model), but do not disagree with the jump-diffusion models considered in this paper. The stochastic-volatility model of Heston [1993] is examined both with and without volatility-risk premia. Excluding volatility-risk premia from the Heston [1993] model results in significant inconsistency between the level of volatility observed in the spot market and that implied, through the model, by the options market. Specifically, given the level of spot-market volatility, this Heston [1993] model *without* volatility-risk premia severely under-prices options. Including volatility-risk premia in the Heston [1993] model resolves this "under-pricing" problem, but the volatility-risk premia thus estimated from $\{S_t, C_t\}$ imply an explosive volatility process under the "risk-neutral" measure, and leads to severely over-priced long-dated options.

The jump-diffusion model of Bates [1997] extends the Heston [1993] model by incorporating jumps in returns whose arrival intensity depends on the level of volatility. We consider two variants of the Bates [1997] model, one incorporating premia for jump risk ("SVJ0"), the other incorporating premia for both volatility and jump risks ("SVJ"). The SVJ0 and SVJ models both fit the joint time-series data well. Key to this goodness of fit is the presence of non-trivial jump-risk premia that are highly correlated with market volatility. The jump-risk premia uncovered by the SVJ model, however, are almost twice those uncovered by the SVJ0 model. The estimated SVJ model compensates for its "overstated" jump-risk premia (relative to the SVJ0 model) with negative volatility-risk premia. The "overstated" jump-risk premia in the SVJ model result in exaggerated volatility "smirks" for short- and medium-dated options, and (through negative volatility-risk premia) under-priced long-dated options.

An empirical analysis of spot and option prices is challenged by the richness of the data. On any given day, one observes a cross section of options with different times to expiration and degrees of moneyness. Such a rich structure of options data is also highly informative, as each option reflects, from a different perspective, expectations regarding both the underlying price dynamics and investors' risk attitudes. For example, the importance of jump-risk premia is only uncovered when we examine the spot and option data jointly.

In order to capture some of this rich information regarding the underlying return dynamics and risk attitudes, this paper adopts an "implied-state" generalized method of moments (IS-GMM) estimation strategy. For a given set ϑ of model parameters, we proxy for the unobserved volatility V_t with an option-implied volatility V_t^ϑ, inverted from the time-t spot price S_t and a near-the-money option price C_t, using the model-implied option-pricing formula. (The true stochastic volatility V_t is recovered only at the true model parameter ϑ_0.)

Access to the option-implied stochastic volatility V^ϑ allows us to explore the joint distribution of spot and option prices by focusing directly on the dynamic structure of the state variables (S, V). This approach[1] is especially attractive in our parametric setting, as the conditional moment-generating function of (S, V) is explicitly known. By taking advantage of the analytically-derived conditional moments of (S, V), we construct "optimal" moment conditions in the spirit of Hansen [1985]. These analytical conditional moments also also used as a rich set of diagnostic tools for investigating model mis-specification.

Our econometric approach differs conceptually from those of Bates [1997] and Bakshi, Cao, and Chen [1997], who focus mainly on cross-sectional options data. Both of these studies estimate the "risk-neutral" model parameters and the current level V_t of stochastic volatility by minimizing the sum of squared differences between the model-implied and market-observed cross-sectional option prices.[2] Their estimation strategies do not take advantage of a time-series model, and data, for the joint spot and option price process. That strategy leaves open whether or not the price dynamics inverted from cross-sectional options data are indeed consistent with the time-series properties of spot and option data. (Further consistency tests in both Bates [1997] and Bakshi, Cao, and Chen [1997] show that the answer is "no.") The integrated time-series approach of this paper is also better adapted to the task of uncovering the nature of risk premia embedded in option prices.

Our findings are also consistent with, and partially explain, some previously reported empirical findings. Using a joint time series of S&P 500 index and option prices, Chernov and Ghysels [1999] find that the stochastic-volatility model of Heston [1993] is strongly rejected by the joint data. Assuming that the S&P 500 index is a one-factor diffusion, Aït-Sahalia, Wang, and Yared [1998] compare the state-price density estimated from cross-sectional S&P 500 option prices to that inferred from time-series data on the S&P 500 index, and find the difference to be significant. Aït-Sahalia, Wang, and Yared [1998] also report that the difference can be partially reconciled by adding a jump component to the index dynamics. Evidence of volatility-risk premia is reported by Guo [1998], Benzoni [1998], Chernov and Ghysels [1999], and Poteshman [1998]. These papers, however, do not consider jump-risk premia, leaving open whether such volatility-risk premia are proxying for missing jump-risk premia (as is shown here to be the case).

This paper also addresses some new econometric issues raised by the use of options with time-varying contract variables, such as times to expiration and moneyness. Because the option-implied volatility V^ϑ depends on such contract variables, time variation in contract variables could introduce spurious effects in V^ϑ (unless, of course, $\vartheta = \vartheta_0$), which in turn introduces a form of "nuisance parameter" in the moment conditions. This problem is not specific to the IS-GMM approach. It applies to other estimation approaches (such as simulated method of moments or maximum-likelihood estimation) that use the same time series of option prices. This issue of time-varying contract variables, inherent in exchange-

[1]An alternative approach can be found in Chernov and Ghysels [1999], who build moment conditions directly on S&P 500 returns and option prices (measured in Black-Scholes implied-volatility), using the SNP/EMM empirical strategy of Gallant and Tauchen [1998].

[2]Bakshi, Cao, and Chen [1997] minimize the sum of squared pricing differences on a daily basis, leaving the model parameters vary from day to day. Bates [1997], on the other hand, minimizes the sum of squared pricing errors over the entire time-series, keeping the model parameters fixed.

traded options data, has been largely ignored in the previous literature. In this paper, we prove consistency and asymptotic normality of the IS-GMM estimators, under mild technical conditions on the time-varying contract variables.

Even for the jump-diffusion models considered in this paper, there remains significant room for improvement. In particular, our diagnostic tests indicate (*i*) mis-specification for the term structure of volatility, and (*ii*) poor fits for the third and fourth conditional moments of stochastic volatility. Both of these findings may call for a stochastic-volatility model with two factors — one strongly persistent, the other quickly mean-reverting and highly volatile. The second finding also suggests the possibility of jumps in volatility.

The form of jump-risk premia considered in this paper can also be extended. In particular, this paper allows only for risk premia for the size of jumps, assigning no premia for uncertainty regarding the timing of jumps. While aversion to both types of jump uncertainty is realistic and potentially important for option valuation, concerns over our ability to separately identify these two types of jump-risk premia motivate us to "lump" them together for purpose of estimating the model.

The rest of the paper is organized as follows. Section 2 sets up an arbitrage-free jump-diffusion pricing model and provides the associated option-pricing formula. Section 3 introduces IS-GMM estimation, establishes conditions for strong consistency and asymptotic normality of IS-GMM estimators with time-varying contract variables, and provides details on the construction of "optimal" moment conditions. Section 4 summarizes the empirical findings, and Section 5 concludes the paper. Technical details are provided in appendices.

2 The Model

This section provides the dynamic model of spot prices, and a specification of risk premia that determine options prices.

2.1 The Data-Generating Process

We fix a probability space $(\Omega, \mathcal{F}, P)$ and an information filtration $(\mathcal{F}_t)$ satisfying the usual conditions,[3] and let S be the ex-dividend price process of a security that pays dividends at a stochastic proportional rate q, whose specification will follow shortly. The stochastic-volatility model with state-dependent jumps (SVJ model) is parameterized as

$$\begin{aligned} dS_t &= \left[r_t - q_t + \eta^s V_t - (\lambda_0 + \lambda_1 V_t)\,\mu^*\right] S_t\,dt + \sqrt{V_t} S_t\,dW_t^{(1)} + dZ_t \\ dV_t &= \kappa_v(\bar{v} - V_t)\,dt + \sigma_v\sqrt{V_t}\,\left(\rho\,dW_t^{(1)} + \sqrt{1-\rho^2}\,dW_t^{(2)}\right)\,, \end{aligned} \tag{2.1}$$

where $W = \left[W^{(1)}, W^{(2)}\right]^\top$ is an adapted standard Brownian motion in $\mathbb{R}^2$, $\rho \in (0,1)$ is a constant coefficient controlling correlation between the "Brownian" shocks to S and V, r is a short interest-rate process defined below, and Z is a pure-jump process to be defined shortly. The stochastic-volatility process V is an autonomous one-factor "square-root" process, characterized by Feller [1951], with constant long-run mean $\bar{v}$, mean-reversion rate κ_v,

[3] See, for example, Protter [1990].

and volatility coefficient σ_v. The jump-event times $\{\mathcal{T}_i : i \geq 1\}$ of the pure-jump process Z arrive with a state-dependent stochastic intensity process $\{\lambda_0 + \lambda_1 V_t : t \geq 0\}$, for some non-negative constants λ_0 and λ_1. (The conditional probability at time t of another jump before $t + \Delta t$ is, for some small Δt, approximately $(\lambda_0 + \lambda_1 V_t)\Delta t$.) At the i-th jump time $\mathcal{T}_i$, the price jumps from $S(\mathcal{T}_i-)$ to $S(\mathcal{T}_i-)\exp(U_i^s)$, where U_i^s is normally distributed with mean μ_J and variance σ_J^2, independent of W, of inter-jump times, and of U_j^s for $j \neq i$. The *mean relative jump size* is $\mu = E(\exp(U^s) - 1) = \exp(\mu_J + \sigma_J^2/2) - 1$. Finally, η^s and μ^* are constant coefficients associated with premia for "Brownian" return risk and jump risk. respectively. The jump model is of the Cox-process type: Conditional on the path of V, jump arrivals are Poisson with time-varying intensity $\{\lambda_0 + \lambda_1 V_t : t \geq 0\}$. (See, for example, Brémaud [1981].)

The short interest-rate process r is of the type modeled by Cox, Ingersoll, and Ross [1985]. Specifically, r and the dividend-rate process q are defined by

$$\begin{aligned} dr_t &= \kappa_r(\bar{r} - r_t)\,dt + \sigma_r\sqrt{r_t}\,dW_t^{(r)} \\ dq_t &= \kappa_q(\bar{q} - q_t)\,dt + \sigma_q\sqrt{q_t}\,dW_t^{(q)}\,, \end{aligned} \tag{2.2}$$

where $W^{(r)}$ and $W^{(q)}$ are independent adapted standard Brownian motions in $\mathbb{R}$, independent also of W and Z. Similar to the stochastic-volatility process V, both r and q are autonomous one-factor square-root processes with constant long-run means ($\bar{r}$ and $\bar{q}$), mean-reversion rates (κ_r and κ_q), and volatility coefficients (σ_r and σ_q). Our formulation of r and q precludes possible correlation between r and q, as well as more plausible and richer dynamics for the short-rate process. For the short-dated options used to fit our model, however, the particular stochastic nature of interest rates r and dividend yields q plays a relatively minor role.[4]

Except for the stochastic short-rate process r and dividend-rate process q, the SVJ model specified in (2.1) is a special case of that of Bates [1997], which in turn extends the stochastic-volatility model (SV) of Heston [1993]. Compared with the Merton [1976] jump extension of the Black and Scholes [1973] model, two important characteristics introduced by the SVJ model are: (1) volatility is itself stochastic, driven by a series of random shocks that could be either positively or negatively correlated with the random shocks driving the price process; (2) the jump-arrival intensity is state-dependent.

Appendix A provides a more precise technical specification of the model by specifying the infinitesimal generator of the state process $(\ln S, V, r, q)$.

2.2 The State-Price Density

The concept of a state-price density (or pricing kernel) is central to the dynamic asset-pricing literature. In essence, a state-price density process π is such that the time-t price of a claim paying Y_s at some future time s is given (under technical conditions) by $E_t(\pi_s Y_s)/\pi_t$, where E_t denotes $\mathcal{F}_t$-conditional expectation. Under technical conditions, the existence of a state-price density ensures the absence of arbitrage, and conversely. (See, for example, Duffie

[4]In this paper, we choose to treat r and q as stochastic processes, as opposed to time-varying constants, in order to accommodate stochastic interest rates and dividend yields, which vary in the data, and whose levels indeed affect even short-dated option prices. This approach could also be potentially useful for studies of very long-dated options such as leaps.

[1996] and references therein.) In our setting of stochastic volatility and jumps, markets are incomplete, and the state-price density is not unique. Our approach[5] is to focus on a candidate state-price density that prices the three important sources of risks: diffusion price shocks, jump risks, and volatility shocks. Without much loss for our pricing problem (because of their mild effects on short-dated option prices), we assume that the interest-rate and dividend-rate risks are not priced.

Consider a candidate state-price density π of the form

$$\pi_t = \exp\left(-\int_0^t r_\tau \, d\tau\right) \mathcal{E}\left(-\int_0^t \zeta_\tau \, dW_\tau\right) \exp\left(\sum_{i,\mathcal{T}_i \leq t} U_i^\pi\right) , \tag{2.3}$$

where $\mathcal{E}(\,\cdot\,)$ denotes the stochastic exponential,[6] and where ζ and U_i^π are defined as follows. Associated with the Brownian motions $W^{(1)}$ and $W^{(2)}$ are their respective market prices of risks $\zeta^{(1)}$ and $\zeta^{(2)}$, defined by

$$\zeta_t^{(1)} = \eta^s \sqrt{V_t} , \qquad \zeta_t^{(2)} = -\frac{1}{\sqrt{1-\rho^2}} \left(\rho\eta^s + \frac{\eta^v}{\sigma_v}\right) \sqrt{V_t} , \tag{2.4}$$

where η^s and η^v are constant coefficients. The jump risks are priced with the *i.i.d.* random variables $U_1^\pi, U_2^\pi, \ldots$, normally distributed with mean μ_π and variance σ_π^2, and independent of W, W^r, W^q, and inter-jump times. We suppose that U_i^π and U_j^s are independent for $i \neq j$, and that U_i^π and U_i^s have correlation ρ_π. The most general form of jump-risk premia is obtained by treating μ_π, σ_π, and ρ_π as free parameters. In this paper, however, we constrain the mean relative jump size in the state-price density to be zero. That is, $\mu_\pi + \sigma_\pi^2/2 = 0$. We postpone a motivation for this constraint to the next subsection when we introduce a "risk-neutral" measure associated with this state-price density π. We let $\mu^* = \exp(\mu_J + \sigma_J\sigma_\pi\rho_\pi + \sigma_J^2/2) - 1$, which can be interpreted as the mean relative jump size under the "risk-neutral" measure associated with π. (A formal definition given in the following section.)

We now show that (2.3) indeed defines a state-price density. Let

$$\mathcal{S} = \left\{ S_t \exp\left(\int_0^t q_\tau \, d\tau\right) : 0 \leq t \leq T \right\} , \quad \mathcal{B} = \left\{ \exp\left(\int_0^t r_\tau \, d\tau\right) : 0 \leq t \leq T \right\} ,$$

be the total gain processes generated by holding one unit of the underlying security and one dollar in the bank account, respectively. For π to be a state-price density, the deflated processes $\mathcal{S}^\pi = \pi\mathcal{S}$ and $\mathcal{B}^\pi = \pi\mathcal{B}$ are required, by definition, to be local martingales. Appendix B shows that this indeed rules out arbitrage opportunities involving $\mathcal{S}$ and $\mathcal{B}$, under natural conditions on dynamic trading strategies.

[5] An alternative approach is preference-based equilibrium pricing, for which the state-price density arises from marginal rates of substitution evaluated at equilibrium consumption streams. See Lucas [1978]. Also, see Naik and Lee [1990] for an extension to jumps, Pham and Touzi [1996] for an extension to stochastic volatility, and Detemple and Selden [1991] for an analysis on the interactions between options and stock markets.

[6] The stochastic exponential of a continuous semi-martingale X, with $X_0 = 0$, is defined by $\mathcal{E}(X)_t = \exp\left(X_t - [X,X]_t/2\right)$, where $[X,X]$ is the total quadratic-variation process.

Finally, we show that $\mathcal{S}^\pi$ and $\mathcal{B}^\pi$ are indeed local martingales. By Ito's Formula,

$$d\mathcal{S}_t^\pi = \left(\sqrt{V_t} - \zeta_t^s\right) \mathcal{S}_t^\pi \, dW_t^s - \zeta_t^v \, \mathcal{S}_t^\pi \, dW_t^v + \left[\exp\left(U_{\mathcal{N}_t}^\pi + U_{\mathcal{N}_t}^s\right) - 1\right] \mathcal{S}_{t-}^\pi \, d\mathcal{N}_t - (\lambda_0 + \lambda_1 V_t)\mu^* \mathcal{S}_t^\pi \, dt$$
$$d\mathcal{B}_t^\pi = -\zeta_t^s \, \mathcal{B}_t^\pi \, dW_t^s - \zeta_t^v \, \mathcal{B}_t^\pi \, dW_t^v + \left[\exp\left(U_{\mathcal{N}_t}^\pi\right) - 1\right] \mathcal{B}_{t-}^\pi \, d\mathcal{N}_t \, ,$$

where $\mathcal{N}_t$ is the number of price jumps by time t. We see that $\mathcal{S}^\pi$ and $\mathcal{B}^\pi$ are in fact local martingales, by using the fact that, for any $i \geq 1$, U_i^π and U_i^s are independent of $\{V_t\}$ and that

$$E\left[\exp\left(U_i^\pi + U_i^s\right) - 1\right] = \exp\left(\mu_\pi + \frac{\sigma_\pi^2}{2} + \mu_J + \sigma_\pi \, \sigma_s \, \rho + \frac{\sigma_J^2}{2}\right) - 1 = \mu^*$$
$$E\left[\exp\left(U_i^\pi\right) - 1\right] = \exp\left(\mu_\pi + \frac{\sigma_\pi^2}{2}\right) - 1 = 0 \, .$$

2.3 The Risk-Neutral Dynamics

In order to rule out arbitrage involving not only the underlying spot and interest rate markets, but also the options markets, it is enough to assign a price $E_t\left(\pi_T \, (S_T - K)^+\right)$ to any call option expiring at time t with strike price K, and likewise for put options. For the purpose of arbitrage-free derivative pricing, however, it is generally convenient to transform the pricing calculation to those under the associated "risk-neutral measure." (Harrison and Kreps [1979]) For this, we define a density process ξ by

$$\xi_t = \pi_t \exp\left(\int_0^t r_s \, ds\right) = \mathcal{E}\left(-\int_0^t \zeta_\tau \, dW_\tau\right) \exp\left(\sum_{\{i:\, \mathcal{T}_i \leq t\}} U_i^\pi\right) \, , \quad 0 \leq t \leq T \, . \tag{2.5}$$

Applying Ito's Formula, one can show that ξ is a local martingale. If ξ is actually a martingale,[7] then ξ uniquely defines an equivalent martingale measure Q. Both r and q have the same joint distribution under Q as under the data-generating measure P. The dynamics of (S, V) under Q are of the form[8]

$$\begin{aligned} dS_t &= \left[r_t - q_t - (\lambda_0 + \lambda_1 V_t)\mu^*\right] S_t \, dt + \sqrt{V_t} \, S_t \, dW_t^{(1)}(Q) + dZ_t^Q \, , \\ dV_t &= \left[\kappa_v \left(\bar{v} - V_t\right) + \eta^v V_t\right] dt + \sigma_v \sqrt{V_t} \left(\rho \, dW_t^{(1)}(Q) + \sqrt{1 - \rho^2} \, dW_t^{(2)}(Q)\right) \, , \end{aligned} \tag{2.6}$$

where $W(Q) = \left[W^{(1)}(Q) \, , \, W^{(2)}(Q)\right]$ is a standard Brownian motion under Q defined by

$$W_t(Q) = W_t + \int_0^t \zeta_s \, ds \, , \quad 0 \leq t \leq T \, . \tag{2.7}$$

This can be shown as an application of Levy's Characterization Theorem. See, for example, Karatzas and Shreve [1991]. The pure-jump process Z^Q has an distribution under Q that is

[7] Appendix C gives a sufficient Novikov-like condition on the model parameters, for ξ to be a martingale.

[8] Appendix A provides a more precise technical specification of the model by specifying the infinitesimal generator of the state process $(\ln S, V, r, q)$ under the risk-neutral measure Q.

identical to the distribution of Z under P defined in (2.1), except that, for any $i \geq 1$, U_i^s is normally distributed with Q-mean μ_J^* and Q-variance σ_J. In particular, one can show that $\mu_J^* = \mu_J + \sigma_J \sigma_\pi \rho_\pi$. The mean relative jump size of S under Q is $\mu^* = E^Q\left(\exp(U^s) - 1\right) = \exp\left(\mu_J^* + \sigma_J^2/2\right) - 1$.

We now focus on the types of jump-risk premia. By allowing the risk-neutral mean relative jump size μ^* to be different from its data-generating counterpart μ, we accommodate a premium for jump-size risk. Similarly, a premium for jump-timing risk can be incorporated, if we allow the coefficients λ_0^* and λ_1^* for the risk-neutral jump-arrival intensity to be different from their respective data-generating counterparts λ_0 and λ_1. In this paper, however, we focus only on the risk premium for jump-size and ignore the risk premium for jump-timing by supposing[9] that $\lambda_0^* = \lambda_0$ and $\lambda_1^* = \lambda_1$. With this assumption, all jump risk premia will be artificially absorbed by the jump-size risk premium coefficient $\mu - \mu^*$, resulting in an time-t expected excess rate of return for jump risk of $(\lambda_0 + \lambda_1 V_t)(\mu - \mu^*)$. We adopt this approach mainly out of empirical concern over our ability to separately identify the risk premia for jump timing and jump size. For example, the arrival intensity of price jumps, as well as the mean relative jump size μ, could be difficult to pin down using the S&P 500 index data under a GMM estimation approach.

Premia for "conventional" return risks ("Brownian" shocks) are parameterized by $\eta^s V_t$, for a constant coefficient η^v. This is similar to the risk-return trade-off in a CAPM framework. Premia for "volatility" risks, on the other hand, are not as transparent, since volatility is not directly traded as an asset. Because volatility is itself volatile, options may reflect an additional volatility risk premium. Volatility risk is priced via the extra term $\eta^v V_t$ in the risk-neutral dynamics of V in (2.6). For a positive coefficient η^v, the time-t instantaneous mean growth rate of the volatility process V is therefore $\eta^v V_t$ higher under the risk-neutral measure Q than under the data-generating measure P. Since option prices respond positively to the volatility of the underlying price in this model, option prices are increasing in η^v.

The linear form of the volatility-risk premia $\eta^v V_t$ could be relaxed by introducing the polynomial form $\eta_0 + \eta_1 V_t + \eta_2 V_t^2 + \cdots + \eta_l V_t^l$, for some constant coefficients $\eta_0, \eta_1, \eta_2, \ldots, \eta_l$. In our specification, however, we rule out the possibility that $\eta_0 \neq 0$. The quadratic term $\eta_2 V_t^2$ seems an interesting case, which is not, however, studied in this paper.

2.4 Option Pricing

The model parameters are $\theta_r = [\kappa_r\,, \bar{r}\,, \sigma_r]^\top$, $\theta_q = [\kappa_q\,, \bar{q}\,, \sigma_q]^\top$, and

$$\vartheta = [\kappa_v\,, \bar{v}\,, \sigma_v\,, \rho\,, \eta^s\,, \eta^v\,, \lambda_0\,, \lambda_1\,, \mu\,, \sigma_J\,, \mu^*]^\top\,, \tag{2.8}$$

where the vectors θ_r and θ_q include the model parameters for the interest-rate process r and the dividend-rate process q, respectively. We will focus on the parameter vector ϑ, an element of a parameter space $\Theta \subset \mathbb{R}^{n_\vartheta}$ with $n_\vartheta = 11$.

[9]This is a direct consequence of our specification of the state-price density π in Section 2.2. Specifically, we impose the constraint $\mu_\pi + \sigma_\pi^2/2 = 0$. One can show that $\lambda_0^* = \lambda_0 \exp\left(\mu_\pi + \sigma_\pi^2/2\right)$ and $\lambda_1^* = \lambda_1 \exp\left(\mu_\pi + \sigma_\pi^2/2\right)$. The constraint $\mu_\pi + \sigma_\pi^2/2 = 0$ therefore corresponds to $\lambda_0^* = \lambda_0$ and $\lambda_1^* = \lambda_1$. Similarly, the risk-neutral standard deviation σ_J^* of the jump amplitude can also be different from its data-generating counterpart σ_J.

Let C_t denote the time-t price of a European-style call option on S, struck at K_t and expiring at $T = t + \tau_t$. Assuming that $E(\pi_T S_T) < \infty$, we have

$$C_t = \frac{1}{\pi_t} E_t \left[\pi_T \left(S_T - K \right)^+ \right] = E_t^Q \left[\exp\left(- \int_t^T r_u \, du \right) \left(S_T - K \right)^+ \right] . \tag{2.9}$$

In order to calculate the expectation in (2.9), we adopt a transform-based approach. (See, for example, Stein and Stein [1991], Heston [1993], Bates [1997], Bakshi, Cao, and Chen [1997], Bakshi and Madan [1999], and Duffie, Pan, and Singleton [1999].) Specifically, for any $c \in \mathbb{C}$, the time-t conditional transform of $\ln S_T$, when well defined, is given by

$$\psi^\vartheta(c, V_t\,, r_t\,, q_t\,, T-t) = E_t^Q \left[\exp\left(- \int_t^T r_u \, du \right) e^{c \, \ln S_T} \right] .$$

Under certain integrability conditions (Duffie, Pan, and Singleton [1999]),

$$\psi^\vartheta(c, v, r, q, \tau) = \exp\Big(\alpha(c, \tau, \vartheta, \theta_r, \theta_q) + \beta_v(c, \tau, \vartheta)\, v + \beta_r(c, \tau, \theta_r)\, r + \beta_q(c, \tau, \theta_q)\, q \Big) , \tag{2.10}$$

where α, β_v, β_r, and β_q are shown explicitly in Appendix D. For notational simplicity, the dependence of ψ on θ_r and θ_q is not shown.

Letting $k_t = K_t / S_t$ be the time-t "strike-to-spot" ratio, the time-t price of a European-style call option with time-to-expiration τ_t can be calculated as

$$C_t = S_t \, f(V_t, \vartheta, r_t, q_t, \tau_t, k_t) ,$$

where $f : \mathbb{R}_+ \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to [0,1]$ is defined by

$$f(v, \vartheta, r, q, \tau, k) = \mathcal{P}_1 - k \, \mathcal{P}_2 , \tag{2.11}$$

with

$$\begin{aligned} \mathcal{P}_1 &= \frac{\psi(1, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}\left(\psi(1 - iu, v, r, q, \tau) e^{i\,u(\ln k)} \right)}{u} \, du \\ \mathcal{P}_2 &= \frac{\psi(0, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}\left(\psi(-iu, v, r, q, \tau) e^{i\,u(\ln k)} \right)}{u} \, du , \end{aligned} \tag{2.12}$$

where $\operatorname{Im}(\cdot)$ denotes the imaginary component of a complex number.

The integrations in (2.12) are typically carried out by a numerical scheme, a potential source of computational burden and numerical errors. In Appendix E, we introduce a new numerical inversion scheme that offers both computational efficiency and error control by taking advantage of the fact that the transform ψ is explicitly known.

3 Estimation

This section focuses on a strategy for using market-observed data to estimate the parameters θ_r, θ_q, and ϑ of the state process (S, V, r, q) and state-price density π. We adopt a two-stage approach. The first stage obtains the maximum-likelihood (ML) estimates of θ_r and θ_q, using time series of interest rates and dividend yields, respectively. The second stage treats the ML estimates of θ_r and θ_q as true parameters, and adopts an "implied-state" GMM estimation strategy for ϑ, using a joint time series of spot and option prices. A direct one-stage estimation of $(\vartheta, \theta_r, \theta_q)$ is feasible and does not involve any new conceptual difficulties. We choose this two-stage approach for simplicity. Any loss of efficiency from the two-stage approach is expected to be small, as the particular stochastic natures of r and q play a relatively minor role in pricing the short-dated options that we use to estimate ϑ. This two-stage approach also allows us to focus more easily, in the second stage, on the dynamic implications of the joint spot price and option data.

Treating the ML estimates θ_r and θ_q as true parameters, we show in this section that "implied-state" GMM estimators are consistent. This result can be easily extended to our two-stage setting, given the consistency of the ML estimators $\hat{\theta}_r$ and $\hat{\theta}_q$.

3.1 "Implied-State" GMM Estimators

Fixing some time interval Δ, we sample the continuous-time state process $\{S_t, V_t, r_t, q_t\}$ at discrete times $\{0, \Delta, 2\Delta, \ldots, N\Delta\}$, and denote the sampled process $\{S_{n\Delta}, V_{n\Delta}, r_{n\Delta}, q_{n\Delta}\}$ by $\{S_n^\Delta, V_n^\Delta, r_n^\Delta, q_n^\Delta\}$. Letting

$$y_n = \ln S_n^\Delta - \ln S_{n-1}^\Delta - \int_{(n-1)\Delta}^{n\Delta} (r_u - q_u)\, du \tag{3.1}$$

denote the date-n "excess" return, it is easy to see that transition distribution of $\{y_n, V_n^\Delta\}$ depends only on parameter vector ϑ, and not on θ_r or θ_q. For the purpose of estimating ϑ, we construct $n_h \geq n_\vartheta$ moment conditions of the form

$$E_{n-1}^{\vartheta_0}\left[h\left(y_{(n,n_y)}, V_{(n,n_v)}, \vartheta_0\right)\right] = 0\,, \tag{3.2}$$

where ϑ_0 is the true model parameter, $h : \mathbb{R}^{n_y} \times \mathbb{R}_+^{n_v} \times \Theta \to \mathbb{R}^{n_h}$ is some test function to be chosen,[10] E_{n-1}^{ϑ} denotes $\mathcal{F}_{(n-1)\Delta}$-conditional expectation under the transition distribution of (y, V) associated with parameter ϑ, and, for some positive integers n_y and n_v,

$$y_{(n,n_y)} = \left[y_n, y_{n-1}, \ldots, y_{n-n_y+1}\right]^\top \quad \text{and} \quad V_{(n,n_v)} = \left[V_n^\Delta, V_{n-1}^\Delta, \ldots, V_{n-n_v+1}^\Delta\right]^\top$$

denote the "n_y-history" of y and the "n_v-history" of V, respectively.

As with the generalized method of moments (GMM) approach of Hansen [1982], the set (3.2) of moment conditions allows for an exploration of the dynamic structure of the state variables S and V. Setting our situation apart from that of a typical GMM, we do not directly observe, at each date n, the stochastic volatility V_n^Δ. Our approach is to take advantage of the date-n market-observed spot price S_n and option price C_n, and explore the option-pricing relation $c_n = C_n/S_n = f(V_n^\Delta, \vartheta_0)$. In fact, we will formally introduce, in the paragraph below, a ϑ-proxy of V_n^Δ, obtained by inverting this option-pricing relation at

[10] We assume that h is continuously differentiable and integrable in the sense of (3.2).

some candidate parameter $\vartheta \in \Theta$. Given this ϑ-proxy of V_n^Δ, denoted V_n^ϑ, we can construct the sample analogue of the moment condition (3.2) by

$$G_N(\vartheta) = \frac{1}{N} \sum_{n \le N} h\left(y_{(n,n_y)}, V^\vartheta_{(n,n_v)}, \vartheta\right), \tag{3.3}$$

and define the "implied-state" GMM (IS-GMM) estimator by $\hat{\vartheta}_N$ by

$$\hat{\vartheta}_N = \arg\min_{\vartheta \in \Theta} \; G_N(\vartheta)^\top \mathcal{W}_N \, G_N(\vartheta), \tag{3.4}$$

where $\{\mathcal{W}_n\}$ is an $(\mathcal{F}_{n\Delta})$-adapted sequence of $n_h \times n_h$ positive semi-definite distance matrices.

We now formally introduce the concept of option-implied volatility V^ϑ. Let $c_n = C_n/S_n$ be the price-to-spot ratio of the option observed on date n, with time τ_n to expiration and strike-to-spot ratio k_n. We have, using the option-pricing function f defined by (2.11),

$$c_n = f(V_n^\Delta, \vartheta_0, r_n^\Delta, q_n^\Delta, \tau_n, k_n), \tag{3.5}$$

where $\vartheta_0 \in \Theta$ is the true model parameter. Let $\Xi \subset [0,1] \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ denote the domain of invertibility (with respect to volatility) of the option-pricing function f of (2.11), in that Ξ is the maximal set for which a mapping $g : \Xi \to \mathbb{R}_+$ is uniquely defined by

$$f(g(c, \vartheta, r, q, \tau, k), \vartheta, r, q, \tau, k) = c, \tag{3.6}$$

for all $(c, \vartheta, r, q, \tau, k) \in \Xi$. We suppose that the parameter space Θ is defined so that, for any observation date n and all $\vartheta \in \Theta$, we have $(c_n, \vartheta, r_n, q_n, \tau_n, k_n) \in \Xi$. In effect, this is a joint property of the data and Θ, akin to an assumption that the model is not shown to be mis-specified. Indeed, in the empirical results to follow, inversion was possible at all data points. For any $\vartheta \in \Theta$, we can therefore define the date-n option-implied volatility by

$$V_n^\vartheta = g(c_n, \vartheta, r_n^\Delta, q_n^\Delta, \tau_n, k_n). \tag{3.7}$$

One important property of V_n^ϑ is that that the true date-n stochastic volatility V_n^Δ is retrieved when V_n^ϑ is evaluated at the true model parameter ϑ_0.

We also note that the sample analogue (3.3) of the moment condition (3.2) requires observations of the excess return y. In order to construct the excess-return process y defined by (3.1), we need to observe, at any time t, the continuous-time processes r and q. In practice, however, we observe r and q at a fixed time interval Δ. In our estimation, we use $\tilde{y}_n = \ln S_n^\Delta - \ln S_{n-1}^\Delta - \left(r_{n-1}^\Delta - q_{n-1}^\Delta\right)\Delta$ as a proxy for y_n. For a relatively short time interval Δ (our data are weekly), the effect of this approximation error on our results is assumed to be small.[11]

Our IS-GMM approach falls into a group of estimation strategies for state variables that can only be observed up to unknown model parameters.[12] This econometric setting arises in

[11] Alternative proxies for $\int_{(n-1)\Delta}^{n\Delta} (r_t - q_t)\, dt$, such as $\left(r_n^\Delta - q_n^\Delta\right)\Delta$ and $\left[\left(r_n^\Delta + r_{n-1}^\Delta\right)/2 - \left(q_n^\Delta + q_{n-1}^\Delta\right)/2\right]\Delta$, are also considered. The empirical results reported in this paper are robust with respect to all three proxies.

[12] In the stochastic-volatility framework of Hull and White [1987], Renault and Touzi [1996] develop an MLE-based two-step iterative procedure. Applications of simulated method of moments to option and underlying spot markets can be found in Pastrorello, Renault, and Touzi [1996] and Chernov and Ghysels [1999]. More recently, Singleton [1999a] develops a conditional-characteristic-function-based estimation method for the general class of affine jump-diffusions. In particular, the implied-state approach using GMM is also discussed in Singleton [1999a].

many other empirical applications. For example, zero- and coupon-bond yields, exchange-traded interest-rate option prices, over-the-counter interest-rate cap and floor data, and swaptions can all in principle be used to invert for an otherwise-unobserved multi-factor state variable that governs the dynamics of the short interest rate process. Dai and Singleton [1997] provide an example in a swap-curve setting. More recently, Piazzesi [1999] adds Federal Reserve target rates and macro-economic variables to the swap-curve setting. As another example, an increasingly popular approach in the literature (on defaultable bonds, in particular) is to model the uncertain mean arrival rate of economic events through some stochastic intensity process.[13] If there exist market-traded instruments whose values are linked to such events, then the otherwise-unobserved intensity processes can be "backed out."

It should be noted that, on each date n, our IS-GMM approach uses only one option price C_n from the entire date-n cross section of option prices, ignoring the additional information potentially contained in the unused cross-sectional option data. Under the null that the model is true, the "neglected" information is in fact redundant. This approach benefits from its avoidance of a relatively *ad hoc* assumption of "pricing errors" for cross-sectional options data. Such a strategy also allows us to focus first only on the tension between spot and option prices, and then extrapolate the IS-GMM estimation result to cross-sectional option data.

Finally, this approach of inverting for a proxy of the otherwise-unobserved state variable V can be extended to cases in which V is multi-dimensional. For example, in the two-factor stochastic-volatility model of Bates [1997], the price process S is driven by two unobserved stochastic volatility factors, $V^{(1)}$ and $V^{(2)}$, which are not directly observed. For this, we could collect, on each date n, the prices of two options with distinct contract variables $(\tau_n^1, k_n^1) \neq (\tau_n^2, k_n^2)$. Under mild technical conditions, we can use the option pairs to obtain proxies for the implied volatility state variables.

3.2 Large-Sample Properties of IS-GMM Estimators

An inherent feature of exchange-traded options is that certain contract variables, such as time τ_n to expiration and strike-to-spot ratio k_n, vary from observation to observation. As the option-implied stochastic volatility V_n^ϑ depends on τ_n and k_n, this variation in contract variables introduces a form of nuisance-dependency to the moment conditions that may affect the large-sample properties of the IS-GMM estimators. In this section, we establish the strong consistency and asymptotic normality of IS-GMM estimators under assumptions of weak time-stationarity of $\{\tau_n\}$ and geometric ergodicity of $\{y_n, V_n, r_n, q_n, k_n\}$. The results established in this section could be useful in other applications using exchange-traded derivative securities.[14]

[13] See, for example, Duffie and Singleton [1999] and Duffee [1999], and references therein.

[14] For exchange-traded derivatives, this situation of time-varying contract variables almost always arises. In over-the-counter markets, however, contract variables on regularly quoted derivative prices are usually constant over time. See Brandt and Santa-Clara [1999] for an application to over-the-counter derivatives.

3.2.1 Stationarity Assumption for Contract Variables

Figure 1 plots $\{\tau_n, k_n\}$ for a time series of S&P 500 options, where τ_n is chosen closest to 30 days to expiration (with a lower bound of 15 days), and where $k_n = K_n/S_n$, with K_n selected nearest to S_n from a grid of available strike prices.[15] Qualitatively, we see that $\{\tau_n\}$ is "repetitive," in an almost deterministic fashion according to the business calendar, while $\{k_n\}$ evolves in a random fashion that can be thought of as a sample path drawn from a stationary process.

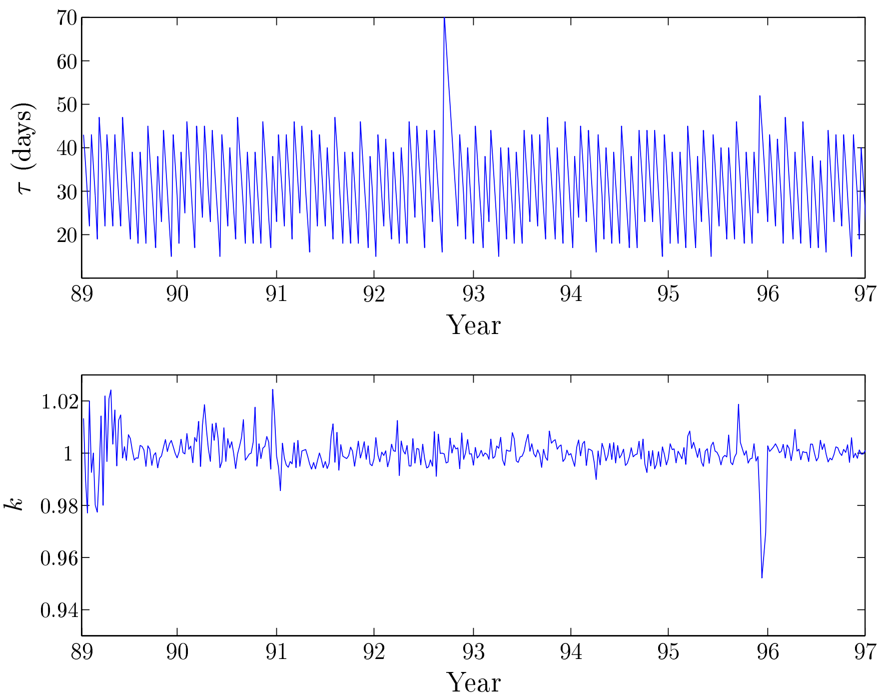

Figure 1: Time series of contract variables: time-to-expiration τ and strike-to-spot ratios k.

Given the nearly periodic feature of $\{\tau_n\}$, the usual mixing conditions used for consistency are difficult to justify. For example, suppose that $\{\tau_n\}$ is of the form (40, 33, 26, 19, 40, 33, 26, 19, ...). Then on date n, depending on where we start initially, τ_n can be 40, 33, 26, or 19. Effectively, this chain has an infinitely long "memory," contrary to the usual mixing property.[16] In this paper, we take an alternative approach, and assume that $\{\tau_n\}$ takes only finitely many outcomes, and satisfies a time-stationarity property (Assumption 3.1 below) that is weaker than typical mixing conditions. In the above example, for instance, $\{\tau_n\}$ is time stationary because the fraction of observations for which $\tau_n = 40$ converges to 0.25, and likewise for each of the other outcomes of τ_n. Such an assumption of finitely many outcomes is characteristic of many derivative contract variables, such as the indicator

[15] To be more precise, we select K_n to be closest to the daily average of spot prices on the n-th day. See also Section 4.1.

[16] The "mixing" property of a Markov chain can be intuitively explained by a physical analogue: the location of a particle or gaseous mixture becomes less and less dependent on its initial position as time progress. See Gallant and White [1988] and references therein.

for "put" versus "call," the exchange identity (for example, CBOE, CME, or PHLX) from which the derivative securities are observed, the maturity of the underlying instruments (in the case of interest-rate derivatives), or multiple selections of an underlying.

An appropriate stationarity assumption for the dynamic behavior of the strike-to-spot ratio $\{k_n\}$, on the other hand, is not as clear. In particular, the evolution of $\{k_n\}$ could be quite complicated, depending on the evolution over time of the strike-price grid, which is driven by detailed institutional features of the equity index option market. In this paper, our consistency result can be based on the assumption that $\{k_n\}$ is, joint with $\{y_n, V_n, q_n, r_n\}$, geometrically ergodic, as stated more precisely below and in Appendix I.

3.2.2 Consistency

We first establish a link between the ϑ-proxy V_n^ϑ and the true volatility state variable V_n^Δ by letting $V_n^\vartheta = \nu(V_n^\Delta, \vartheta, r_n^\Delta, q_n^\Delta, \tau_n, k_n)$, where $\nu : \mathbb{R}_+ \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is defined by

$$\nu(v, \vartheta, r, q, \tau, k) = g(f(v, \vartheta_0, r, q, \tau, k), \vartheta, r, q, \tau, k)\,, \tag{3.8}$$

where g is defined by (3.6), using the fact that $c_n = f(V_n^\Delta, \vartheta_0, r_n^\Delta, q_n^\Delta, \tau_n, k_n)$. We note that $\nu(v, \vartheta_0, r, q, \tau, k) = v$.

Next, letting $X_n = \left[y_{(n,n_y)}, V_{(n,n_v)}, r_{(n,n_v)}, q_{(n,n_v)}, k_{(n,n_v)}\right]$ denote the "n_y-history" of y and the "n_v-histories" of r, q, k, and τ, and letting $Y_n = \tau_{(n,n_v)}$ denote the "n_v-history" of τ, we write

$$H(X_n, \vartheta, Y_n) = h\left(y_{(n,n_y)}, \nu\left(V_{(n,n_v)}, \vartheta, r_{(n,n_v)}, q_{(n,n_v)}, \tau_{(n,n_v)}, k_{(n,n_v)}\right), \vartheta\right)\,, \tag{3.9}$$

where $r_{(n,n_v)} = [r_n, r_{n-1}, \ldots, r_{n-n_v+1}]$, and, analogously, $q_{(n,n_v)}$, $k_{(n,n_v)}$, and $\tau_{(n,n_v)}$ are the n_v-dimensional vectors consisting of q_n, k_n, τ_n, and their respective lags. As outlined in the previous subsection, reasonable stationarity assumptions for X and Y are rather different, and are treated separately.

Assumption 3.1 (Time Stationarity of Y) *$\{Y_n\}$ has finitely many outcomes, denoted $\{1, 2, \ldots, I\}$. For each outcome i and each positive integer N, let $A_N^{(i)} = \{n \le N : Y_n = i\}$ be the dates, up to N, on which Y has outcome i. For each i, there is some $w_i \in [0, 1]$, such that*

$$\lim_N \frac{\# A_N^{(i)}}{N} = w_i \quad a.s.\,, \tag{3.10}$$

where $\#(\,\cdot\,)$ denotes cardinality.

Appendix I shows the state vector $\{y_n, V_n, r_n, q_n\}$ is geometrically ergodic, under easy-to-check parameter restrictions on ϑ. Assuming further that $\{k_n\}$ and $\{y_n, V_n, r_n, q_n\}$ are jointly geometrically ergodic, we know that $X_n = \left[y_{(n,n_y)}, V_{(n,n_v)}, r_{(n,n_v)}, q_{(n,n_v)}, k_{(n,n_v)}\right]$ is geometrically ergodic, since it includes only finitely many lags of the joint process. The pointwise strong law of large numbers (SLLN) part of Assumption 3.2 below then follows from Glynn [1999], under Assumption 3.1 and the additional assumption of independence between X

and Y. This independence assumption is trivially satisfied in our setting because Y is deterministic.[17]

Assumption 3.2 (USLLN of $A^{(i)}$-Sampling) *For each outcome i of Y, letting*

$$G_N^{(i)}(\vartheta) = \frac{1}{\# A_N^{(i)}} \sum_{n \in A_N^{(i)}} H(X_n, \vartheta, i)\,,$$

$G_\infty^{(i)}(\vartheta) = \lim_N G_N^{(i)}(\vartheta)$ *exists (pointwise SLLN), and*

$$\sup_{\vartheta \in \Theta} |G_N^{(i)}(\vartheta) - G_\infty^{(i)}(\vartheta)| \to 0 \quad a.s.\,. \tag{3.11}$$

Given the pointwise-SLLN portion of Assumption 3.2, in order to establish the uniform SLLN of Assumption 3.2, it is typical to assume some form of Lipschitz condition on $H(x, \vartheta, i)$ as a function of ϑ. Examples of such conditions include the Lipschitz and derivative conditions of Andrews [1987] and the first-moment-continuity condition of Hansen [1982].

We now establish the uniform strong law of large numbers (USLLN) of $\{H(X_n, \vartheta, Y_n)\}$, key step step to establishing the strong consistency of $\{\hat{\vartheta}_N\}$. A proof is given in Appendix H.

Proposition 3.1 (USLLN of $H(X, \vartheta, Y)$) *Under Assumptions 3.1 and 3.2, for each ϑ, $G_\infty(\vartheta) = \lim_N G_N(\vartheta)$ exists, and*

$$\sup_{\vartheta \in \Theta} |G_N(\vartheta) - G_\infty(\vartheta)| \to 0 \quad a.s.\,,$$

where $G_N(\vartheta)$, defined by (3.3), is the sample moment of the observation function.

Finally, to show strong consistency of the IS-GMM estimator $\{\hat{\vartheta}_N\}$, we adopt the following two standard assumptions.

Assumption 3.3 (Convergence of Weighting Matrices) $\mathcal{W}_N \to \mathcal{W}_0$ *almost surely for some constant symmetric positive-definite matrix $\mathcal{W}_0$.*

Under Assumption 3.3 and the conditions of Proposition 3.1, the criterion function $C_N(\vartheta) = G_N(\vartheta)^\top \mathcal{W}_N G_N(\vartheta)$ converges almost surely to the asymptotic criterion function $C : \Theta \to \mathbb{R}$ defined by $C(\vartheta) = G_\infty(\vartheta)^\top \mathcal{W}_0 G_\infty(\vartheta)$. In particular, we have $C(\vartheta_0) = 0$, given the moment condition (3.2), the pointwise-SLLN portion of Proposition 3.1, and the fact that $V_n^{\vartheta_0} = V_n^\Delta$.

Assumption 3.4 (Uniqueness of Minimizer) $C(\vartheta_0) \neq C(\vartheta)$, $\vartheta \in \Theta$, $\vartheta \neq \vartheta_0$.

Theorem 3.1 (Strong Consistency) *Under Assumptions 3.1–3.4, the IS-GMM $\{\vartheta_N\}$ estimator converges to ϑ_0 almost surely as $N \to \infty$.*

Given the Uniform SLLN (Proposition 3.1), the proof is standard and omitted. (See, for example, the proof of Theorem 3.3 in Gallant and White [1988].)

[17] Independent-sampling strong laws for more general processes can be found in Glynn and Sigman [1998].

3.2.3 Asymptotic Normality

Next, we establish asymptotic normality for the IS-GMM estimator, allowing for time-varying contract variables. Because $\nu(v, \vartheta_0, r, q, \tau, k) = v$, the sample moment $G_N(\vartheta)$ evaluated at the true parameter ϑ_0 does not depend on the contract variables $\{\tau_n, k_n\}$. Given the consistency result above, the asymptotic normality of $\sqrt{N} G_N(\vartheta_0)$ therefore depends only on the properties of (y, V) and h via a standard form of Central Limit Theorem (CLT).

Assumption 3.5 (CLT) *$\sqrt{N} G_N(\vartheta_0)$ converges in distribution as $N \to \infty$ to a normal random vector with mean zero and some covariance matrix Σ_0.*

This assumption follows immediately from the geometric ergodicity of (y, V) and an assumption of integrability of $||h\left(y_{(n,n_y)}, V_{(n,n_v)}\right)||^{2+\delta}$, for some $\delta > 0$, over the stationary distribution of $\left(y_{(n,n_y)}, V_{(n,n_v)}\right)$. (See, for example, Theorem 7.5 of Doob [1953] and the proof of Theorem 4 of Duffie and Singleton [1993].)

The asymptotic normality of $\sqrt{N}(\vartheta_N - \vartheta)$ depends further on the local behavior of the observation functions in a neighborhood of ϑ_0, and is influenced by the contract variables $\{\tau_n, k_n\}$. For this, we consider the derivative $d(\vartheta, X_n, Y_n)$ of $H(X_n, \vartheta, Y_n)$ with respect to ϑ, defined by

$$d(\vartheta, X_n, Y_n) = \frac{\partial}{\partial \vartheta} h\left(y_{(n,n_y)}, V^{\vartheta}_{(n,n_v)}, \vartheta\right) + \sum_{i=n-n_v+1}^{n} \frac{\partial}{\partial v_i} h\left(y_{(n,n_y)}, V^{\vartheta}_{(n,n_v)}, \vartheta\right) g_{\vartheta}\left(c_i, \vartheta, r^{\Delta}_i, q^{\Delta}_i, \tau_i, k_i\right), \tag{3.12}$$

where $g_{\vartheta}(c, \vartheta, r, q, \tau, k) = \partial g(c, \vartheta, r, q, \tau, k)/\partial \vartheta$, with g defined by (3.6), and where $c_i = f(V^{\Delta}_i, \vartheta_0, r^{\Delta}_i, q^{\Delta}_i, \tau_i, k_i)$. The first term on the right-hand side of (3.12) arises from the explicit dependence of h on ϑ, while the second term arises from the dependence of h on V_i and the dependence of $V^{\vartheta}_i = g(c_i, \vartheta, r^{\Delta}_i, q^{\Delta}_i, \tau_i, k_i)$ on ϑ, for $i \in \{n - n_v + 1, \ldots, n\}$. This second term is important in identifying risk-premium parameters such as η^v. Intuitively, such parameters are identified by exploring the option-pricing relation through V^{ϑ}.

Assumption 3.6 (Convergence of "Jacobian Estimator") *For some constant $(n_h \times n_{\vartheta})$ matrix d_0 of rank n_{ϑ}: (i) $\frac{1}{N}\sum_{n \le N} d(\vartheta_0, X_n, Y_n)$ converges in probability as $N \to \infty$ to d_0. (ii) For any $\{\vartheta_n\}$ converging in probability as $n \to \infty$ to ϑ_0, $\frac{1}{N}\sum_{n \le N} d(\vartheta_n, X_n, Y_n)$ converges in probability as $N \to \infty$ to d_0.*

Part (i) of Assumption 3.6 follows from geometric ergodicity of X, independence and time-stationary of Y, and integrability (over the stationary distribution of X) of $d(\vartheta, X_n, i)$, for each i. Given that part (i) holds, part (ii) follows from assuming first-moment continuity (as in Hansen [1982]) of $d(\vartheta, X_n, Y_n)$ at ϑ_0.

Theorem 3.2 (Asymptotic Normality) *Under Assumptions 3.1–3.6, $\sqrt{N}\left(\vartheta_N - \vartheta_0\right)$ converges in distribution as $N \to \infty$ to a normal random vector with mean zero and covariance matrix*

$$\Lambda = (d_0^{\top} \mathcal{W}_0\, d_0)^{-1}\, d_0^{\top} \mathcal{W}_0\, \Sigma_0\, \mathcal{W}_0\, d_0\, (d_0^{\top} \mathcal{W}_0\, d_0)^{-1}. \tag{3.13}$$

The proof is a standard application of the mean-value theorem (for example, Hamilton [1994]), and omitted. The asymptotic covariance matrix Λ differs from its GMM counterpart in that d_0 is affected by the dependence of V^ϑ on ϑ and $\{\tau_n, k_n\}$.

For the usual two-step GMM of Hansen [1982], under which the distance matrices are chosen so that $\mathcal{W}_0 = \Sigma_0^{-1}$, we have $\Lambda = (d_0^\top \Sigma_0^{-1} d_0)^{-1}$. Our setting is that of an exactly-identified GMM estimator ($n_h = n_\vartheta$, d_0 is of rank n_ϑ, and $\mathcal{W}_0$ is the identify matrix), so $\Lambda = d_0^{-1}\Sigma_0(d_0^\top)^{-1}$.

3.3 "Optimal" Moment Selection

In this section, we construct a set of "optimal" moment conditions by taking advantage of the explicitly known date-n conditional moment-generating function of (y_{n+1}, V_{n+1}). (Throughout this subsection, we denote V_n^Δ by V_n for notational simplicity.) Under certain integrability conditions (Duffie, Pan, and Singleton [1999]), one can show that, for any u_y and u_v in $\mathbb{R}$,

$$E_n\Big[\exp\left(u_y\, y_{n+1} + u_v\, V_{n+1}\right)\Big] = \phi(u_y, u_v, V_n)\,,$$

where $\phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_+$ is defined by[18]

$$\phi(u_y, u_v, v) = \exp\left(A\left(u_y, u_v\right) + B\left(u_y, u_v\right) v\right)\,, \tag{3.14}$$

where A and B are shown explicitly in Appendix D.

With the conditional moment-generating function $\phi(\,\cdot\,)$, one can in principle perform full-information estimation that is asymptotically equivalent to MLE. For example, Singleton [1999a] develops a characteristic-function-based estimator for general affine diffusions that is computationally tractable and asymptotically efficient. Liu [1997] develops a GMM-based approach for affine diffusions, and shows that MLE efficiency can be achieved by increasing the number of moment conditions. These two approaches offer a natural framework under which the optimal instruments of Hansen [1985] can be implemented.

Our approach is closely related to those of Liu [1997] and Singleton [1999a]. We first select a set of moment conditions implied by the conditional moments of (y, V), and then construct "optimal" instruments for the selected moment conditions. In both steps, we rely on the fact that

$$E_n\left(y_{n+1}^i V_{n+1}^j\right) = \left.\frac{\partial^{(i+j)}\phi(u_y, u_v, V_n)}{\partial^i u_y\, \partial^j u_v}\right|_{u_y=0\,,\,u_v=0}\,, \quad i, j \in \{0, 1, \dots\}\,. \tag{3.15}$$

Direct computation of the derivatives in (3.15), although straightforward, can be cumbersome for higher orders of i and j.[19] Appendix F offers an easy-to-implement method for calculating $E_n(y_{n+1}^i V_{n+1}^j)$, recursively in i and j, up to arbitrary orders.

[18]See also Heston [1993], Bates [1997], and Das and Sundaram [1999].

[19]Das and Sundaram [1999] derive the first four central moments of y for the Heston [1993] model of (S, V), extended to include jumps at a constant intensity.

We let $M : \mathbb{R}_+ \times \Theta \to \mathbb{R}^7$ denote the list of conditional moments of (y, V) defined by $M_1(V_n, \vartheta) = E_n^\vartheta(y_{n+1})$, $M_2(V_n, \vartheta) = E_n^\vartheta\left(y_{n+1}^2\right)$, $M_3(V_n, \vartheta) = E_n^\vartheta\left(y_{n+1}^3\right)$, $M_4(V_n, \vartheta) = E_n^\vartheta\left(y_{n+1}^4\right)$, $M_5(V_n, \vartheta) = E_n^\vartheta(V_{n+1})$, $M_6(V_n, \vartheta) = E_n^\vartheta\left(V_{n+1}^2\right)$, and $M_7(V_n, \vartheta) = E_n^\vartheta(y_{n+1}V_{n+1})$. For this choice of conditional moments, we construct the "fundamental moment conditions" by

$$E_{n-1}^\vartheta(\epsilon_n) = 0\,, \quad \epsilon_n = [\epsilon_n^{y1}, \epsilon_n^{y2}, \epsilon_n^{y3}, \epsilon_n^{y4}, \epsilon_n^{v1}, \epsilon_n^{v1}, \epsilon_n^{yv}]^\top\,, \tag{3.16}$$

where

$$\begin{aligned} \epsilon_n^{y1} &= y_n - M_1(V_{n-1}, \vartheta)\,, \quad \epsilon_n^{y2} = y_n^2 - M_2(V_{n-1}, \vartheta) \\ \epsilon_n^{y3} &= y_n^3 - M_3(V_{n-1}, \vartheta)\,, \quad \epsilon_n^{y4} = y_n^4 - M_4(V_{n-1}, \vartheta) \\ \epsilon_n^{v1} &= V_n - M_5(V_{n-1}, \vartheta)\,, \quad \epsilon_n^{v2} = V_n^2 - M_6(V_{n-1}, \vartheta) \\ \epsilon_n^{yv} &= y_n V_n - M_7(V_{n-1}, \vartheta)\,. \end{aligned} \tag{3.17}$$

This choice (3.16) of moment conditions is intuitive, and provides some natural and testable conditions on certain lower moments and cross moments of y and V. Relative to full-information MLE, however, this approach sacrifices some efficiency by exploiting only a limited portion of the distributional information contained in the moment-generating function.

Next, we construct "optimal" instruments for the fundamental moment conditions of (3.16). In the spirit of Hansen [1985], we define the "optimal" moment conditions by

$$\mathcal{H}_{n+1} = \mathcal{Z}_n \epsilon_{n+1}\,, \quad \text{with} \quad \mathcal{Z}_n = \mathcal{D}_n^\top \times \left(\mathrm{Cov}_n^\vartheta(\epsilon_{n+1})\right)^{-1}\,, \tag{3.18}$$

where $\mathrm{Cov}_n^\vartheta(\epsilon_{n+1})$ denotes the date-n conditional covariance matrix of ϵ_{n+1} associated with the parameter ϑ, and $\mathcal{D}_n$ is the $(7 \times n_\vartheta)$ matrix with i-th row $\mathcal{D}_n^i$ defined by

$$\begin{aligned} \mathcal{D}_n^i &= -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta} - g_\vartheta(c_n, \vartheta)\, \frac{\partial M_i(v, \vartheta)}{\partial v}\bigg|_{v=V_n}\,, \qquad i = 1, 2, 3, 4\,, \\ \mathcal{D}_n^i &= -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta}\,, \qquad i = 5, 6, 7\,, \end{aligned} \tag{3.19}$$

where $c_n = f(V_n, \vartheta_0)$ with f given by the option-pricing formula (2.11), and where $g_\vartheta(c, \vartheta) = \partial g(c, \vartheta)/\partial\vartheta$ with g defined by (3.6). (For notational simplicity, the dependence of f, g, and g_ϑ on (r, q, τ, k) is not shown.) The component of $\mathcal{D}$ associated with $g_\vartheta(c_n, \vartheta)$ is specific only to our implied state variable setting. Intuitively, $g_\vartheta(c_n, \vartheta)$ measures the sensitivity of the date-n option-implied volatility $V_n^\vartheta = g(c_n, \vartheta)$ to ϑ. We can calculate g_ϑ by using the implicit function theorem, in that

$$g_\vartheta(c_n, \vartheta) = -f_v^{-1}(V_n, \vartheta)\, f_\vartheta(V_n, \vartheta)\,, \tag{3.20}$$

where $f_v = \partial f(v, \vartheta)/\partial v$, and $f_\vartheta = \partial f(v, \vartheta)/\partial \vartheta$, with the option pricing formula f defined by (2.11). The partial derivatives f_v and f_ϑ can be calculated explicitly, up to numerical integration, by differentiating through the integrals in (2.12).

Each element $\mathcal{H}_{n+1}^j$ of the "optimal" observations $\mathcal{H}_{n+1} = \left(\mathcal{H}_{n+1}^1, \ldots, \mathcal{H}_{n+1}^{n_\vartheta}\right)$ is associated with an element ϑ_j of the parameter vector ϑ. Intuitively, $\mathcal{H}_{n+1}^j$ is the weighted sum of the 7 fundamental observations ϵ_{n+1}, normalized by the covariance matrix $\text{Cov}_n^{\vartheta}(\epsilon_{n+1})$, with weights proportional to the date-n "conditional sensitivity" of ϵ_{n+1} to ϑ_j. Given this set $\mathcal{H}$ of "optimal" observations, we can apply our implied-state-variable approach outlined in Section 3.4 by replacing the unobserved stochastic volatility V_n with the option-implied stochastic volatility V_n^{ϑ}.

It should be noted[20] that the efficiency of this "optimal-instrument" scheme is limited in that the Jacobian $\mathcal{D}$ constructed in (3.19) differs from that, denoted $(\mathcal{D})^{\text{H}}$, of Hansen [1985]. Specifically, we have $(\mathcal{D}_n^i)^{\text{H}} = \mathcal{D}_n^i$, for $i \in \{1, \ldots, 4\}$, but

$$
\begin{aligned}
(\mathcal{D}_n^5)^{\text{H}} &= \mathcal{D}_n^5 + E_n^{\vartheta}\left[g_{\vartheta}(c_{n+1}, \vartheta)\right] - g_{\vartheta}(c_n, \vartheta) \left.\frac{\partial M_5(v, \vartheta)}{\partial v}\right|_{v=V_n} \\
(\mathcal{D}_n^6)^{\text{H}} &= \mathcal{D}_n^6 + E_n^{\vartheta}\left[2V_{n+1}^{\vartheta} g_{\vartheta}(c_{n+1}, \vartheta)\right] - g_{\vartheta}(c_n, \vartheta) \left.\frac{\partial M_6(v, \vartheta)}{\partial v}\right|_{v=V_n} \\
(\mathcal{D}_n^7)^{\text{H}} &= \mathcal{D}_n^7 + E_n^{\vartheta}\left[y_{n+1} g_{\vartheta}(c_{n+1}, \vartheta)\right] - g_{\vartheta}(c_n, \vartheta) \left.\frac{\partial M_7(v, \vartheta)}{\partial v}\right|_{v=V_n},
\end{aligned}
\tag{3.21}
$$

where $c_{n+1} = f(V_{n+1}, \vartheta_0)$. Effectively, in constructing $\mathcal{D}^5$, $\mathcal{D}^6$, and $\mathcal{D}^7$, we sacrifice efficiency by ignoring the dependence of V_{ϑ} on ϑ. We do, however, gain analytic tractability, as calculations of the form $E_n^{\vartheta}[g_{\vartheta}(c_{n+1}, \vartheta)]$, $E_n^{\vartheta}[V_{n+1}\, g_{\vartheta}(c_{n+1}, \vartheta)]$, and $E_n^{\vartheta}[y_{n+1}\, g_{\vartheta}(c_{n+1}, \vartheta)]$ would be challenging.

4 Empirical Results

In this section, we find that jump-risk premia, especially their ability to respond quickly to market volatility, are critical in reconciling the dynamics implied by spot and option prices. On their own, premia for return risk and volatility risk, cannot accommodate the joint data. Further diagnostic analyses of the models considered in this section provide evidence of misspecification in term structure of volatility, and the possibility of jumps in volatility. This section then extrapolates the time-series estimation results to cross-sectional options data, finding that jump-risk premia also play an important role in explaining volatility "smiles" and "smirks."

4.1 Data

The joint spot and option data are from the Berkeley Options Data Base (BODB), a complete record of trading activity on the floor of the Chicago Board Options Exchange (CBOE). We construct a time-series $\{S_n, C_n\}$ of S&P 500 index and near-the-money short-dated option

[20]When the state is observed directly, the optimality of this choice of instruments (without the extra parameter dependence associated with the "implied" state) for our conditional moment restrictions follows immediately from Hansen [1985] and Hansen, Heaton, and Ogaki [1988]. In our case of implied states, an analogous optimality result obtains for a class of estimators based on the same moment equations evaluated at the implied states. This can be seen by adapting the analysis in Hansen [1985] to the case of implied states; see also Singleton [1999b] for details.

prices, from January 1989 to December 1996, with "weekly" frequency (every 5 trading days), as follows.

For each observation day, we collect all of the bid-ask quotes (on both calls and puts) that are time-stamped in a pre-determined sampling window. The sampling window, lying always between 10:00am to 10:30am, varies from year to year. For example, it is set at 10:07am–10:23am for all trading days in 1989; for 1996, it is set at 10:14am–10:16am. Such adjustments in the length of the sampling window accommodate significant changes from year to year in the trading volume of S&P 500 options. Our objective is to have an adequate pool of options with a spectrum of expirations and strike prices. For the n-th observation day, we first sort the options by time to expiration. Among all available options, we select those with a time τ_n to expiration that is larger than 15 calendar days and as close as possible to 30 calendar days.[21] From the pool of options with the chosen time τ_n to expiration, we next select all options with a strike price K_n that is nearest to the date-n average of the S&P 500 index. If the remaining pool of options, with the chosen τ_n and K_n, contains multiple calls, we select one of these call options at random. Otherwise, a put option is selected at random.[22] By repeating this strategy for each date n, we obtain a time-series $\{C_n\}$ of option prices, using the average of bid and ask prices. One nice feature of the CBOE data set is that, for each option price C_n, we have a record of the contemporaneous S&P 500 index price S_n. The combined time series $\{S_n, C_n\}$ is therefore synchronized. The sample mean of $\{\tau_n\}$ is 31 days, with a sample standard deviation of 9 days. The sample mean of the strike-to-spot ratio $\{k_n = K_n/S_n\}$ is 1.0002, with a sample standard deviation of 0.0067. The time series $\{\tau_n, k_n\}$ is illustrated in Figure 1.

4.2 Interest Rates r and Dividend Yields q

For the purpose of estimating the respective parameter vectors θ_r and θ_q of the short-rate process r and the dividend-rate process q defined by (2.2), we use, from Datastream, weekly time-series of 3-month LIBOR rates and S&P 500 composite dividend yields from January 1987 to December 1996.

Fixing a sampling interval Δ, and taking advantage of the fact that the conditional density of q_n^Δ given q_{n-1}^Δ is that of a non-central χ^2 (Feller [1951] and Cox, Ingersoll, and Ross [1985]), we estimate θ_q using MLE. The time series of S&P 500 composite dividend yields is used as a proxy for $\{q_n^\Delta\}$. The observed T-year LIBOR rates $\{R_n\}$ (converted to continuous compounding rates) can be expressed in terms of r_n^Δ by (Cox, Ingersoll, and Ross [1985])

$$R_n = -\frac{1}{T}\left(\alpha_r\left(0, T, \theta_r^0\right) + \beta_r\left(0, T, \theta_r^0\right)\, r_n^\Delta\right),$$

where θ_r^0 denotes the true parameter vector, and where α_r and β_r are as defined in (D.3).

[21] Both time to expiration τ_n and sampling interval Δ are annualized, using a 365-calendar-day year and a 252-business-day year, respectively.

[22] One can either use the put-call parity to convert the observed put price to that of a call option, or treat the mixture of call and put options using an additional contract variable. These two approaches are equivalent, for our estimation strategy.

The one-period conditional density $p^R(\,\cdot\,|R_{n-1}\,;\,\theta_r)$ of R_n given R_{n-1} is therefore given by

$$p^R\left(x \,\middle|\, R_{n-1}\,;\,\theta_r\right) = \frac{T}{|\beta(0,T,\theta_r)|}\, p^r\left(-\frac{xT+\alpha_r(0,T,\theta_r)}{\beta_r(0,T,\theta_r)} \,\middle|\, r^{\Delta}_{n-1}\,;\,\theta_r\right), \quad x \in \mathbb{R}_+\,,$$

where, as with the dividend-rate process q, the one-step conditional density $p^r(\,\cdot\,|\,r^{\Delta}_{n-1}\,;\,\theta_r)$ of the short-rate r^{Δ}_n is that of a non-central χ^2.

Table 1: ML Estimates of Interest Rates r and Dividend Yields q.

κ_r	$\bar{r}$	σ_r	κ_q	$\bar{q}$	σ_q
0.20	0.058	0.0415	0.24	0.025	0.0269
(0.15)	(0.016)	(0.0009)	(0.33)	(0.011)	(0.0004)

Data: Weekly 3-month LIBOR rates and S&P 500 dividend yields, Jan. 1987 to Dec. 1996.

The ML estimates of θ_r and θ_q are summarized in Table 1. The long-run means of r and q are 5.8% and 2.5%, respectively. Both processes exhibit high persistence with relatively slow mean reversions.

4.3 Estimation Results and Goodness-of-Fit Tests

In this section, we use the time series $\{S_n, C_n\}$ of S&P 500 index and option prices to estimate the parameters of the SVJ and nested models, and examine how well the models accommodate the joint time-series data.

Throughout this section we maintain the assumption that $\lambda_0 = 0$, so that the jump-arrival intensity is $\lambda_1 V$. We choose this formulation so as to reduce the number of free parameters that are important in explaining the joint distribution of spot and option prices. This hypothesis of $\lambda_0 = 0$ is tested formally in Section 4.3.3. We also consider the following three nested models:

- The SVJ0 model: $\eta^v = 0$ (no volatility-risk premia).
- The SV model: $\lambda_1 = 0$ (no jumps).
- The SV0 model, $\lambda_1 = 0$ and $\eta^v = 0$ (no jumps, no volatility-risk premia).

Each of the four models (SVJ, SVJ0, SV, and SV0) is estimated within an exactly-identified IS-GMM setting, with the number of "optimal" moment conditions equaling the number of unknown model parameters. Table 2 reports the IS-GMM estimates and asymptotic standard errors. Table 3 summarizes the results of goodness-of-fit tests for the four models. These tests focus on how well the respective models satisfy the fundamental moment conditions (3.16), and are constructed directly from the heteroskedasticity-corrected version $\tilde{\epsilon}$ of ϵ, defined by

$$\tilde{\epsilon}^{\,i}_n = \frac{\epsilon^{\,i}_n}{\sqrt{E_{(n-1)}(\epsilon^{\,i}_n)^2}}\,, \quad i \in \{1,\ldots,7\}. \tag{4.1}$$

We test the 7 moment conditions, $E_{n-1}(\tilde{\epsilon}_n) = 0$, both individually and jointly. The large-sample distribution of the test statistics is standard normal for the individual tests, and, for any n, χ^2 with n degrees of freedom for a joint test on n moment conditions.[23] We let $\tilde{\epsilon}^y = [\tilde{\epsilon}^{y1}, \tilde{\epsilon}^{y2}, \tilde{\epsilon}^{y3}, \tilde{\epsilon}^{y4}]^\top$ denote the y-related moments, and let $\tilde{\epsilon}^v = [\tilde{\epsilon}^{v1}, \tilde{\epsilon}^{v2}]^\top$ denote the V-related moments. In addition to the joint test on $E(\tilde{\epsilon}_n) = 0$, results for joint tests on $E(\tilde{\epsilon}_n^y) = 0$ and $E(\tilde{\epsilon}_n^v) = 0$ are reported in Table 3. We next interpret the results of these goodness-of-fit tests, as well as the estimation results.

4.3.1 The Pure-Diffusion Models: SV0 and SV

The SV0 model (Heston [1993] with no volatility-risk premia) is strongly rejected by the joint time-series data. Table 3 shows that the joint test on $E(\tilde{\epsilon}_n) = 0$ is rejected in the SV0 model with a p-value of 10^{-10}. Results from the individual tests reveal that a key source of this violation is the moment condition $E(\tilde{\epsilon}_n^{y2}) = 0$, which connects the realized squared return y_n^2 to its conditional expectation $M_2(V_{n-1}^{\vartheta}, \vartheta)$. By a Taylor-expansion of $M_2(v, \vartheta)$ over small Δ, we have $M_2(V_{n-1}^{\vartheta}, \vartheta) \approx V_{n-1}^{\vartheta}\Delta$, which leads to $\epsilon_n^{y2} \approx y_n^2 - V_{n-1}^{\vartheta}\Delta$. The significantly negative test statistics associated with $E(\epsilon_n^{y2})$ therefore indicate that the volatility realized in the spot market is significantly less than that observed from the options market through the SV0 model. In other words, given the level of volatility observed in the underlying spot market, the SV0 model under-prices options.[24] One possible explanation for this under-pricing is that the SV0 model does not incorporate investors' aversion toward the risk of changes in volatility.

Indeed, moving to the SV model (Heston [1993]), we investigate the possibility of a volatility-risk premium by introducing a volatility-risk premium coefficient η^v. As reported in Table 2, the SV model estimate of η^v is positive and significantly different from zero. (Here and below, we use a conventional p-value of 5% to judge "significance.") Let $\mathcal{H}_{n+1}(\eta^v)$ denote the "optimal" moment associated with η^v, as described in Section 3.3. We further perform a Lagrange-multiplier test of the SV0 model against the SV model, using the moment condition $E_n[\mathcal{H}_{n+1}(\eta^v)] = 0$. This test is of the Lagrange-multiplier style, in the sense that the moment condition $E_n[\mathcal{H}_{n+1}(\eta^v)] = 0$, which is true under the alternative (the SV model), is tested using the parameter estimates associated with the null (the SV0 model). The SV0 model (that with $\eta^v = 0$) is rejected against the SV model, with a p-value of 0.0002. We also see from Table 3 that the moment condition $E(\epsilon_n^{y2}) = 0$ is no longer strongly violated for the SV model. These findings suggest that introducing a volatility-risk premium partially reconciles the tension between spot and option prices that arises in the SV0 setting.

The role of volatility-risk premia is also examined in Guo [1998], Benzoni [1998], Poteshman [1998], Kapadia [1998], and Chernov and Ghysels [1999]. As volatility-risk premia are

[23] Appendix G provides more details on large-sample distributions of such test statistics. Our tests of moment conditions follow from the tests of orthogonality conditions developed in Eichenbaum, Hansen, and Singleton [1988], and are also closely related to the Hansen [1982] test of over-identifying restrictions. More details can be found in Singleton [1999b].

[24] Similar findings under the Black-Scholes setting are reported in Jackwerth and Rubinstein [1996]. The SV0 model considered in this paper extends the Black-Scholes model by allowing stochastic volatility, but it still has the same risk-premium structure as the Black-Scholes model. It is therefore not surprising that the SV0 model is found to "under-price" near-the-money short-dated options.

the only form of risk premia examined in these studies, the evidence in support of volatility-risk premia is not conclusive. For example, these studies ignore any premia for jump risk, an issue to be examined shortly.

Table 2: IS-GMM Estimates of the SVJ and Nested Models.

	κ_v	$\bar{v}$	σ_v	ρ	η^s	η^v	λ_1	μ	σ_J	μ^*
SV0	5.3	0.0242	0.38	−0.57	4.4	$\equiv 0$	$\equiv 0$			
	(1.9)	(0.0044)	(0.04)	(0.05)	(1.8)					
SV	7.1	0.0137	0.32	−0.53	8.6	7.6	$\equiv 0$			
	(2.1)	(0.0023)	(0.03)	(0.06)	(2.3)	(2.0)				
SVJ0	7.1	0.0134	0.28	−0.52	3.1	$\equiv 0$	27.1	−0.3%	3.25%	−18.0%
	(1.9)	(0.0029)	(0.04)	(0.07)	(2.9)		(11.8)	(1.7%)	(0.64%)	(1.6%)
SVJ	7.5	0.0135	0.28	−0.49	−1.5	−7.5	55.8	−0.5%	2.70%	−17.1%
	(2.0)	(0.0026)	(0.04)	(0.09)	(6.9)	(11.6)	(54.3)	(1.3%)	(0.73%)	(1.6%)

Data: Weekly spot and options data, S&P 500 index, Jan. 1989 to Dec. 1996.

Table 3 shows that, although allowing for a volatility-risk premium leads to a significant improvement of the overall goodness of fit, the joint test on $E(\tilde{\epsilon}_n) = 0$ is still strongly rejected (with a p-value of 10^{-5}) in the SV model.

Indeed, a disconcerting feature of the SV model is that its parameter estimates imply an explosive stochastic-volatility process under the risk-neutral measure,[25] with an estimated risk-neutral mean-reversion rate of $\hat{\kappa}_v^* = \hat{\kappa}_v - \hat{\eta}^v < 0$. Although such behavior is not explicitly ruled out by arbitrage arguments, it leads the SV model to severely over-price long-dated options, as will be shown in Section 4.6.

Figure 2 plots the time series of option-implied volatility V^{SV0} and V^{SV}, backed out using the IS-GMM estimates $\hat{\vartheta}^{SV0}$ and $\hat{\vartheta}^{SV}$ of the SV0 and SV models, respectively. While the observed patterns of V^{SV0} and V^{SV} are similar, V^{SV0} runs at a relatively higher level. The same finding is reflected in estimated long-run mean $\bar{v}$ of the stochastic-volatility. We find the estimate $\sqrt{\bar{v}} = 15.6\%$ for the SV0 model, and $\sqrt{\bar{v}} = 11.7\%$ for the SV model.[26] During the same period, the annualized sample standard deviation of the S&P 500 weekly return is 11.4%. This is consistent with our previous finding that, assuming the SV0 model at its estimated parameters, the volatility implied by option prices is higher than that observed in the spot market.

Both the SV0 and SV models show evidence of "volatility asymmetry." That is, the correlation coefficient ρ between the short-run returns and changes in volatility is estimated

[25] This is largely due to an overstated volatility-risk premium associated with the large value of $\hat{\eta}^v$, which also causes condition (C.2) to be violated. As a result, a Novikov-like sufficient condition (laid out in Appendix C) for the martingality of π is not satisfied, and there is no guarantee of the existence of a risk-neutral measure Q.

[26] We report these numbers in the form of $\sqrt{\bar{v}}$, since $\sqrt{V}$ is the "volatility process" in the conventional sense of the relative "instantaneous" standard deviation.

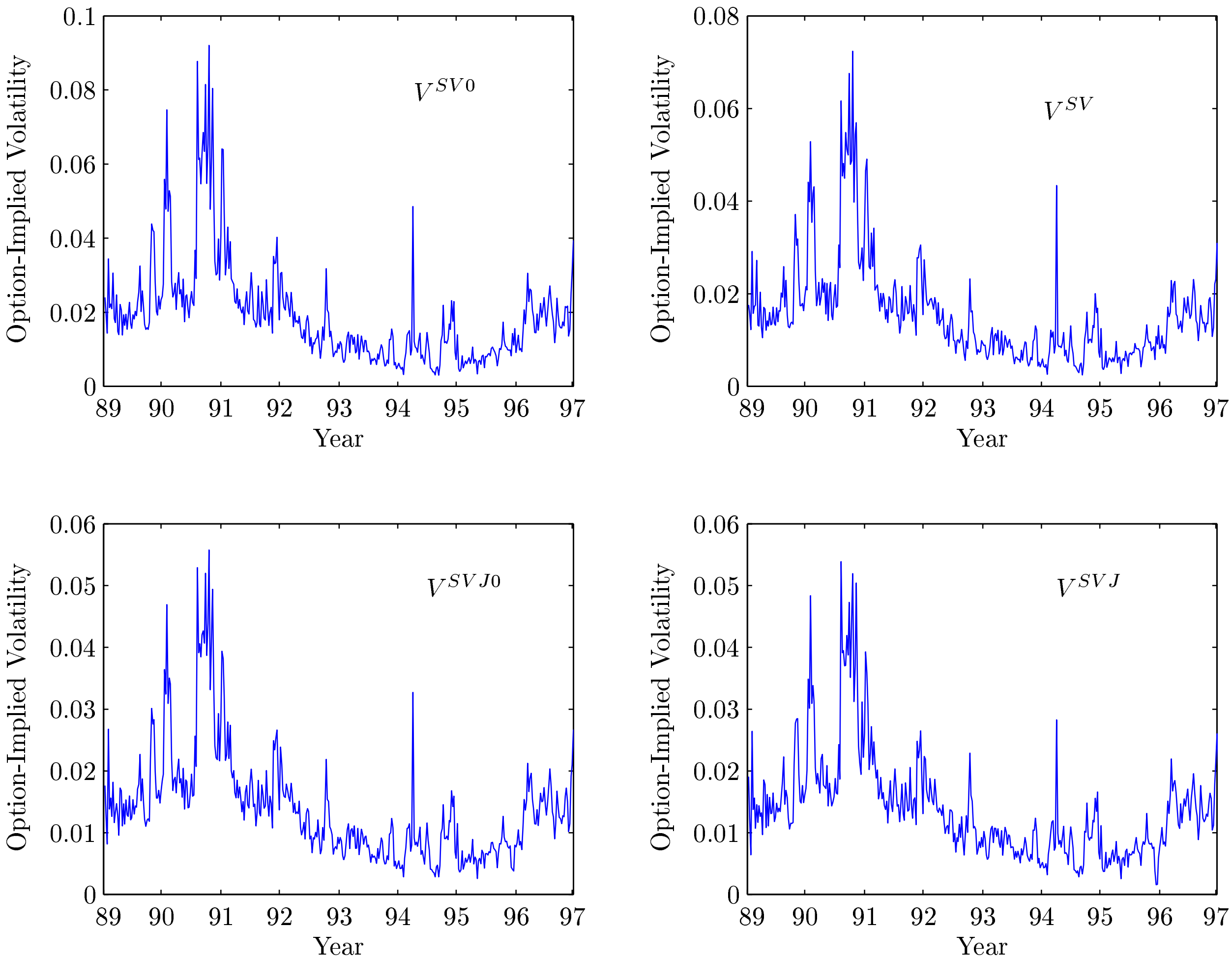

Figure 2: Time-series of option-implied volatility $\{V_n^{SV0}\}$, $\{V_n^{SV}\}$, $\{V_n^{SVJ0}\}$, and $\{V_n^{SVJ}\}$, backed out from the joint time-series $\{S_n, C_n\}$ of spot and option prices, using the IS-GMM estimates of the SV0, SV, SVJ0, and SVJ models.

Table 3: Goodness-of-Fit Tests on Fundamental Moment Conditions

	$\tilde{\epsilon}$	SV0	SV	SVJ0	SVJ
Individual Tests	$y1$	1.46	-0.59	0.27	0.46
	$y2$	-3.98^{**}	-1.56	-0.60	-0.71
	$y3$	0.77	-0.29	-0.65	-0.50
	$y4$	-1.45	0.27	-0.36	-0.21
	$v1$	-1.91	1.80	0.95	0.92
	$v2$	-2.28^{*}	1.24	0.59	0.52
	yv	2.47^{*}	-0.10	0.60	0.74
Joint Tests	all y	28.2^{**}	8.1	1.8	1.4
	$\chi^2(4)$	(10^{-5})	(0.09)	(0.77)	(0.84)
	all v	9.1^{*}	11.4^{**}	3.2	3.1
	$\chi^2(2)$	(0.01)	(0.003)	(0.20)	(0.21)
	all	59.9^{**}	31.6^{**}	7.6	7.1
	$\chi^2(7)$	(10^{-10})	(10^{-5})	(0.37)	(0.42)

$*$ and $**$ indicate significance under a 5% and 1% test, respectively. For individual tests, only the test statistics (standard normal in large sample) are reported. The p-values for the χ^2 joint tests are reported in parentheses.

to be significantly negative for both models.[27] Although the point estimates for ρ vary from model to model, the order of magnitude of these estimates is the same for all the models considered here, even for those with jumps. Similar results are reported by Bakshi, Cao, and Chen [1997] for the options market, and by Andersen, Benzoni, and Lund [1998] for the spot market.[28]

Moving from the SV0 model to the SV model, we also notice a decrease in the estimated volatility coefficient σ_v of the volatility process, and an improvement in the goodness of fit associated with ϵ^{v2}, which tests how well the volatility of volatility is fit. (As one moves to the jump models of SVJ0 and SVJ, this decrease in σ_v and improvement in the fit of ϵ^{v2} are more evident.) Both Bates [1997] and Bakshi, Cao, and Chen [1997] report that in order to explain the volatility "smiles" and "smirks" found in the cross-sectional options data, this volatility parameter σ_v has to be set to a level that is too high to be consistent with the time-series property of the volatility process. Our SV0 estimate of σ_v is very close to that found in Bakshi, Cao, and Chen [1997], and the inconsistency reported in both

[27] The economic mechanism behind this negative correlation is a subject of growing interest. Possible explanations considered in recent studies include volatility feedback (Campbell and Hentschel [1992] and Wu [1998]), differences of opinion among investors (Hong and Stein [1999]), and investor uncertainty over the true drift rate of dividends (David and Veronesi [1999]).

[28] Using joint spot and options data on the S&P 500 market Chernov and Ghysels [1999] report estimates for ρ that is one order of magnitude smaller (in absolute value). Using time-series data of the S&P 500 index, Eraker, Johannes, and Polson [1999] report positive but insignificant estimates for ρ, a puzzle that remains to be resolved.

Bates [1997] and Bakshi, Cao, and Chen [1997] is also reflected in our goodness-of-fit test associated with ϵ^{v2}. In this paper, however, we uncover this inconsistency from a time-series investigation of spot and near-the-money short-dated option prices. Another related point has recently been raised by Jones [1999], who reports that the volatility of volatility is higher during more volatile markets, a phenomenon the Heston [1993] model cannot accommodate. In order to allow a more relaxed volatility structure of volatility, Jones [1999] suggests a stochastic-volatility model in the class of constant elasticity of variance.[29] It is unclear how such a model will perform in the presence of jump and jump-risk premia. In particular, as we include jump and jump-risk premia in the SVJ0 and SVJ models, we do not find any evidence of mis-specification for volatility of the volatility process.

4.3.2 Introducing Jumps in Returns: the SVJ0 and SVJ Models

Moving from the pure-diffusion models, SV0 and SV, to the jump-diffusion models, SVJ0 and SVJ, our first observation is a much improved overall goodness of fit. As reported in Table 3, p-values of the joint test of the hypothesis that $E(\tilde{\epsilon}_n) = 0$ are 0.34 and 0.42 for the SVJ0 and SVJ models, respectively. In contrast to the pure-diffusion models considered earlier, the SVJ0 and SVJ models are not rejected by the joint time-series data.

The most interesting aspect of the SVJ0 and SVJ models is the premia for jump risk. Focusing first on the SVJ0 model, we see from Table 2 that the risk-neutral mean μ^* of the relative jump size is estimated to be -18%, while its counterpart μ for the data-generating process is estimated to be -0.3%. This implies that, when weighted by aversion to large price movements, negative jumps are perceived to be more negative ($\mu^* - \mu$ is estimated at -17.6%, with a standard error of 2.2%). Actual daily returns of comparable magnitude occurred only once, when the market jumped -23% on October 19, 1987. It seems, however, that fear of such adverse price movements is reflected in option prices, through a large jump-risk premium.

It should be noted that because we have set the jump-timing risk premium $\lambda_1^* - \lambda_1$ to zero, it is likely that the estimated premium for jump-size risk, measured in terms of $\mu^* - \mu$, has absorbed some risk regarding timing risk. Alternatively, there may be some other aspects of the market price of jump risk that we have mis-specified, such as assuming that the risk-neutral and actual variance of jump sizes is the same. Consequently, in interpreting the above result, the reader should be cautioned that such a risk premium, measured in terms of $\mu^* - \mu$, reflects not only investors' fear regarding the size of jumps, but also their fear regarding other types of jump risks. In particular, aversions to both jump size and jump timing are realistic and potentially important. Our approach, however, does not provide a clear picture of how investors' jump-risk aversion is split between jump size and jump timing. We adopt this approach only out of the concern over our ability to separately identify $\lambda_1^* - \lambda_1$ and $\mu^* - \mu$, since both μ and λ_1 are known to be difficult to pin down from the S&P 500 index data with the usual GMM approach.

The time-t instantaneous mean rate of return associated with the jump-risk premium is $\lambda_1(\mu - \mu^*)V_t$, while that associated with the premium for usual return risk is $\eta^s V_t$. Measuring the average volatility level by its long-run mean $\sqrt{\bar{v}}$, the SVJ0 model estimates imply that

[29] Alternatively, one can introduce a second volatility factor.

the average mean excess rate of (cum-dividend) return demanded for jump risk is 6.4% with a standard deviation of 1.6%, while that demanded for usual return risk is 4.2% with a standard deviation of 3.8%. This shows that a significant portion of the equity risk premium is assigned as a premium to compensate for investors' aversion towards jumps. Again, this premium could be a "lumped" effect including investors' aversion toward jump-size risk, as well as jump-timing risk.

The SVJ model relaxes the constraint $\eta^v = 0$ from the SVJ0 model to accommodate a premium for volatility risk. Table 2 shows that while the point estimates of μ^* and μ differ only slightly from their SVJ0 counterparts, the estimate for λ_1 is roughly twice that for the SVJ0 model. The SVJ estimates of the volatility parameters $\bar{v}$, κ_v, and σ_v stay close to their respective SVJ0 counterparts, indicating similar volatility dynamics for these two models.[30] Combining the evidence, we conclude that, moving from the SVJ0 model to the SVJ model, the estimated "instantaneous" jump-risk premium $\lambda_1(\mu - \mu^*)V_t$ roughly doubles in magnitude. From the negative SVJ-model estimates of the risk-premium coefficients η^s and η^v, we see that the SVJ model compensates for this "overstated" (relative to the SVJ0 model) jump-risk premium by negative premia for volatility and conventional return risks.[31] As we extend our analysis to cross-sectional options data in Section 4.6, this "overstated" jump-risk premium shows up in the form of exaggerated volatility "smirks" implied by the estimated SVJ model. It is plausible that the relaxed parameter η^v is compensating for some form of mis-specification, although this is difficult to interpret.

The goodness-of-fit tests reported in Table 3 do not show a significant preference between the SVJ0 and SVJ models. Let $\mathcal{H}_n(\eta^v)$ denote the "optimal" moment associated with η^v, as described in Section 3.3. Using the moment condition $E[\mathcal{H}_n(\eta^v)] = 0$, which is true under the SVJ model, we perform a Lagrange-multiplier test of SVJ0 against SVJ, using the SVJ0-model estimates. The SVJ0 model (that with $\eta^v = 0$) is not rejected against the SVJ model at traditional confidence levels. (The p-value is 0.55.)

4.3.3 Testing the Constant Component of the Jump Arrival Intensity

In order to further investigate the possible presence of a constant component λ_0 of the jump-arrival intensity, we test the SVJ0 model against the alternative that the jump-arrival intensity is $\lambda_0 + \lambda_1 V$. The test is of the Lagrange-multiplier style, and based on the moment condition $E\left[\mathcal{H}_n(\lambda_0)\right] = 0$, where $\mathcal{H}_n(\lambda_0)$ is the "optimal" moment associated with λ_0. The hypothesis that $\lambda_0 = 0$ is not rejected at traditional confidence levels. (The p-value is 0.12.)

Because of our earlier constraint that the "risk-neutral" jump-arrival intensity coefficients λ_0^* and λ_1^* be the same as their respective data-generating counterparts λ_0 and λ_1, the jump-arrival intensity $\lambda_0 + \lambda_1 V_t$ plays two different roles: (1) it dictates the jump arrival times under the data-generating measure;[32] (2) it reconciles the dynamics implied by spot and option prices by introducing a premium for jump risk. The LM test performed here hinges more on

[30]Similarly, the option-implied volatilities V^{SVJ0} and V^{SVJ} are close in magnitude and pattern, as shown in Figure 2.

[31]Similar to the case for the SV model, a disconcerting feature of the SVJ model is that the relatively large value of $|\eta^v|$ (relative to κ^v) causes the Novikov-like condition (C.2) to be violated.

[32]For evidence of state-dependent jump times under the data-generating measure, see Johannes, Kumar, and Polson [1998].

its latter role as the jump-risk "pricer," showing that the missing constant component λ_0 in the SVJ0 model does not result in any significant tension in the system.

4.4 A Comparative Study of Premia for Jump and Volatility Risks

The previous analysis makes it clear that, in order to fit the joint time-series of spot and option prices, certain types of risk premia, beyond simple compensation for short-run return uncertainty, must be introduced. The SVJ0 model incorporates jump-risk premia, while the SV model considers volatility-risk premia. This section offers a comparative examination of these two risk premia.

We introduce a concept of *relative risk premia*, which measures the relative difference in prices between an option priced with and without risk premia. Specifically, letting $\hat{\vartheta}^{SVJ0}$ denote the IS-GMM estimate of ϑ for the SVJ0 model, we "turn off" the jump-risk premium component of $\hat{\vartheta}^{SVJ0}$ by letting ϖ^{SVJ0} denote the parameter vector $\hat{\vartheta}^{SVJ0}$, changed only by letting $\mu^* = \mu$. Similarly, letting $\hat{\vartheta}^{SV}$ denote the estimate of ϑ for the SV model, we "turn off" the volatility-risk premium component of $\hat{\vartheta}^{SV}$ by letting ϖ^{SV} be a copy of $\hat{\vartheta}^{SV}$, changed only by letting $\eta^v = 0$. For an option with time $\overline{\tau}$ to expiration and strike-to-spot ratio $\overline{k}$, the date-n relative jump-risk premium $\Pi_n^J(\overline{\tau}, \overline{k})$ and the relative volatility-risk premium $\Pi_n^V(\overline{\tau}, \overline{k})$ are defined by

$$\Pi_n^J(\overline{\tau}, \overline{k}) = \frac{c_n^{SVJ0} - \overline{c}_n^{SVJ0}}{c_n^{SVJ0}}, \quad \Pi_n^V(\overline{\tau}, \overline{k}) = \frac{c_n^{SV} - \overline{c}_n^{SV}}{c_n^{SV}}, \tag{4.2}$$

where, letting f denote the option-pricing function defined by (2.11), and letting V_n^{SVJ0} and V_n^{SV} denote the SVJ0 and SV model estimates of the date-n option-implied volatilities, we define

$$\begin{aligned} c_n^{SVJ0} &= f(V_n^{SVJ0}, \hat{\vartheta}^{SVJ0}, r_n^{\Delta}, q_n^{\Delta}, \overline{\tau}, \overline{k}), \quad & c_n^{SV} &= f(V_n^{SV}, \hat{\vartheta}^{SV}, r_n^{\Delta}, q_n^{\Delta}, \overline{\tau}, \overline{k}), \\ \overline{c}_n^{SVJ0} &= f(V_n^{SVJ0}, \varpi^{SVJ0}, r_n^{\Delta}, q_n^{\Delta}, \overline{\tau}, \overline{k}), \quad & \overline{c}_n^{SV} &= f(V_n^{SV}, \varpi^{SV}, r_n^{\Delta}, q_n^{\Delta}, \overline{\tau}, \overline{k}). \end{aligned}$$

Intuitively, $\overline{c}_n^{SVJ0}$ and $\overline{c}_n^{SV}$ are the respective "risk-premium-free" versions of c_n^{SVJ0} and c_n^{SV}.

Fixing strike-to-spot ratio $\overline{k} = 1$, Figure 3 plots time-series of $\{\Pi_n^J(\overline{\tau}, \overline{k}), \Pi_n^V(\overline{\tau}, \overline{k})\}$ for a list of times $\overline{\tau}$ to expiration. Figure 3 shows that jump-risk and volatility-risk premia respond quite differently to market volatility. The relative jump-risk premium is highly responsive to market volatility for short-dated options, becoming less responsive for long-dated options. On the other hand, the relative volatility-risk premium is relatively unresponsive to market volatility for short-dated options, reaching maximum responsiveness for medium-dated options, then becoming less responsive for long-dated options.

Figure 3 also shows that the relative volatility-risk premium increases much faster with maturity than does the relative jump-risk premium. Moreover, these two risk premia are considerably different in magnitude for short-dated and long-dated options. In particular, for long-dated options, the relative volatility-risk premia are significantly higher them their jump counterparts, which is consistent with our finding (see Section 4.6) that the SV model severely over-prices long-dated options. Here we see that the over-pricing is due to an overstated volatility-risk premium.

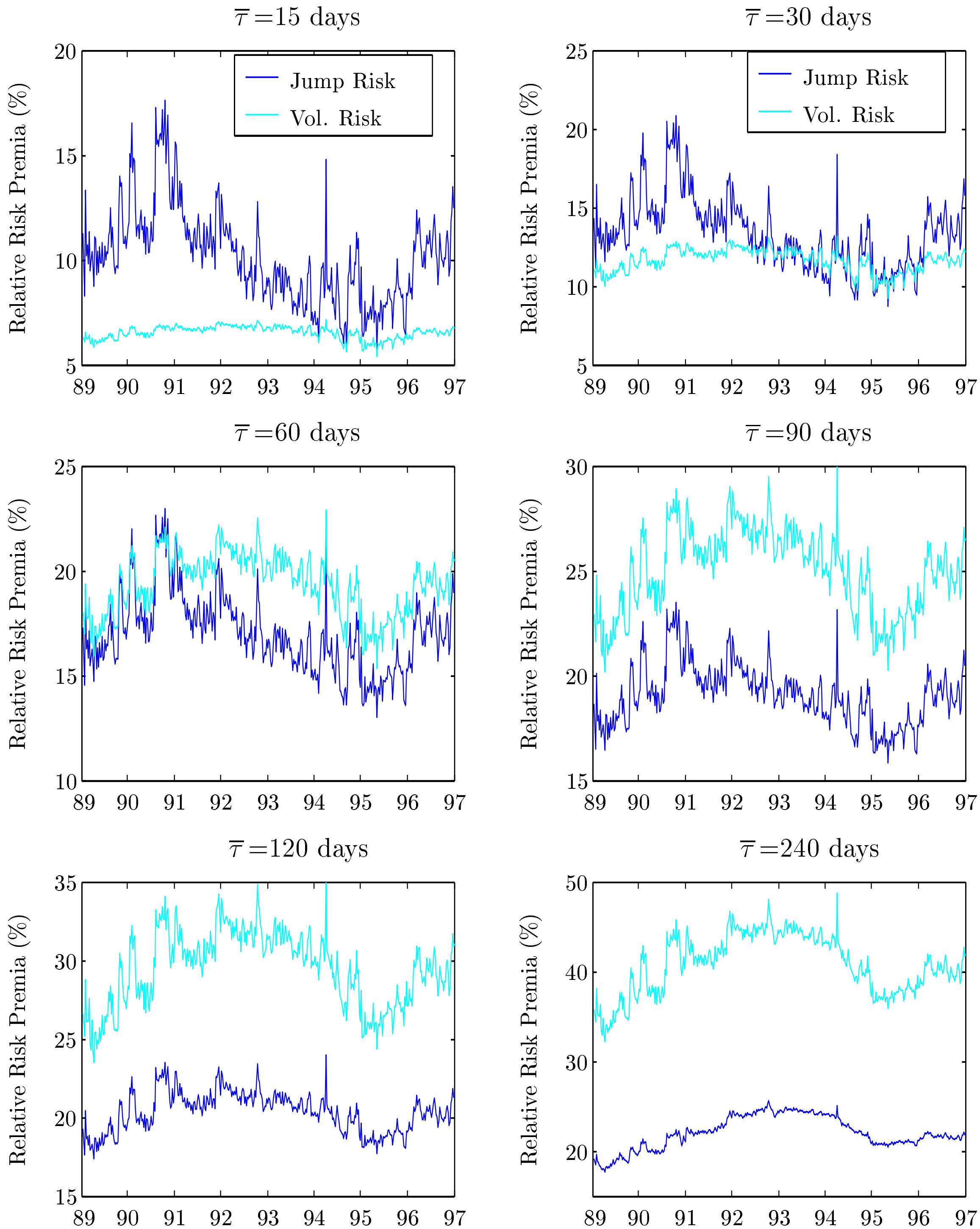

Figure 3: The relative jump-risk premium $\Pi_n^J(\overline{\tau}, \overline{k})$ implied by the SVJ0 model, and the relative volatility-risk premium $\Pi_n^V(\overline{\tau}, \overline{k})$ implied by the SV model, for options with strike-to-spot ratio $\overline{k} = 1$, and times to expiration of 15, 30, 60, 90, 120, and 240 days. See Section 4.4 for a definition of relative risk premia.

While it may be dangerous to draw conclusions based on this sort of extrapolation to "out-of-sample" options with substantially longer times to expiration, the risk-premium behavior implied by the SV model appears to be counter-intuitive. On the other hand, the risk-premium behavior implied by the SVJ0 model seems more intuitive.

4.5 Diagnostic Tests

In the spirit of goodness-of-fit tests of fundamental moment conditions, this section reports the results of further diagnostic tests of the SVJ and nested models. Section 4.5.1 uses extra conditioning information to test the fundamental moment conditions, and reports some evidence of model mis-specification for term structure of volatility. Section 4.5.2 focuses on the third and fourth moments of stochastic volatility, and finds evidence of jumps in stochastic volatility.

4.5.1 Tests of $E_n(\epsilon_{n+1}) = 0$ with Conditional Information

This section tests the moment conditions $E_n(y_n \epsilon_{n+1}) = 0$ (lag-y tests), and $E_n(\epsilon_n \epsilon_{n+1}) = 0$ (lag-ϵ tests). The test results are reported in Table 4.

The "lag-y" tests are designed mainly to look for any impact of excess returns y on the dynamics of V, possibly missed by our model specification. As reported in Table 4, the lag-y tests do not indicate mis-specification of this sort.

Table 4: Diagnostic tests of fundamental moment conditions

		SV0	SV	SVJ0	SVJ	SV0	SV	SVJ0	SVJ
	ϵ	lag-ϵ $E(\epsilon_n\, \epsilon_{n+1}) = 0$				lag-y $E(y_n\, \epsilon_{n+1}) = 0$			
Individual Tests	$y1$	-0.63	-0.34	-0.58	-0.62	-0.70	-0.66	-0.73	-0.74
	$y2$	1.38	0.38	0.43	0.51	0.39	0.22	-0.11	-0.19
	$y3$	-0.35	0.21	-0.48	-0.59	-0.60	-0.42	-0.75	-0.79
	$y4$	1.72	0.25	-0.26	-0.26	1.47	1.27	0.68	0.51
	$v1$	-2.37^*	-2.62^{**}	-2.04^*	-2.19^*	1.58	1.45	1.44	1.39
	$v2$	-1.68	-1.73	-1.21	-1.25	0.81	0.70	0.74	0.69
	yv	-0.48	-0.13	-0.35	-0.31	-0.22	-0.01	0.22	-0.28
Joint Tests	all y $\chi^2(4)$	3.4 (0.49)	0.49 (0.97)	1.1 (0.89)	1.4 (0.83)	4.2 (0.38)	3.8 (0.43)	3.1 (0.54)	2.8 (0.59)
	all v $\chi^2(2)$	7.3^* (0.03)	7.9^* (0.02)	7.4^* (0.02)	9.4^{**} (0.01)	5.3 (0.07)	4.8 (0.09)	4.9 (0.09)	5.2 (0.07)
	all $\chi^2(7)$	14.3^* (0.05)	9.9 (0.20)	9.4 (0.23)	10.9 (0.14)	10.7 (0.15)	8.5 (0.29)	8.0 (0.33)	8.0 (0.34)

$*$ and $**$ indicate significance under a 5% and 1% test, respectively. For the individual tests, only the test statistics (standard normal in large sample) are reported. The p-values for the joint χ^2 tests are reported in parentheses.

Using the moment condition $E_n(\epsilon_n \, \epsilon_{n+1}) = 0$, "lag-$\epsilon$" tests investigate whether or not the fundamental moments ϵ are indeed serially uncorrelated. The 7 individual lag-ϵ tests reject the null that $E_n(\epsilon_n^{v1} \epsilon_{n+1}^{v1}) = 0$, but do not reject the other 6 moment conditions. The sample estimate of $\mathrm{corr}(\epsilon_n^{v1}, \epsilon_{n+1}^{v1})$ is negative and significant. To see the implication of this finding, we recall that $\epsilon_n^{v1} = V_n^{\vartheta} - M_5(V_{n-1}^{\vartheta}, \vartheta)$, where, for any $v \in \mathbb{R}_+$,

$$M_5(v, \vartheta) = \exp(-\kappa_v \Delta t) v + (1 - \exp(-\kappa_v \Delta t)) \, \bar{v} \, .$$

It then follows that the unconditional stationary correlation can be calculated as[33]

$$\mathrm{corr}(\epsilon_n^{v1}, \epsilon_{n-1}^{v1}) = \mathrm{corr}\left(V_{n+1} - V_n e^{-\kappa_v \Delta t} \, , \quad V_n - V_{n-1} e^{-\kappa_v \Delta t} \right) \, .$$

Using the fact that the one-step auto-correlation $\mathrm{corr}(V_n, V_{n-1})$ is $\exp(-\kappa_v \Delta)$, we have $\mathrm{corr}(\epsilon_n^{v1}, \epsilon_{n-1}^{v1}) = e^{-\kappa_v \Delta t} \left[e^{-2\kappa_v \Delta t} - \mathrm{corr}\left(V_{n+1}, V_{n-1}\right) \right]$. A negative and significant sample estimate of $\mathrm{corr}(\epsilon_n^{v1}, \epsilon_{n-1}^{v1})$ therefore indicates that the data call for $\mathrm{corr}(V_{n+1}, V_{n-1}) > \exp(-2\kappa_v \Delta t)$. This is in contrast to the model-prescribed two-step auto-correlation $\mathrm{corr}(V_{n+1}, V_{n-1}) = \exp(-2\kappa_v \Delta t)$. In other words, the stochastic volatility model is not capable of fitting simultaneously one- and two-step auto-correlations. The reason is quite simple: The model prescribes a term-structure of volatility, $\mathrm{corr}(V_n, V_{n+m}) = \exp(-m \kappa_v \Delta t)$, which "dies" too quickly, relative to the data. Multiple volatility factors with different rates of mean-reversion could generate a richer term-structure of volatility. Some examples include the two-factor square-root model of Bates [1997] and a stochastic-volatility model with stochastic long-run mean suggested by Duffie, Pan, and Singleton [1999].

4.5.2 Tests of Higher Moments of Volatility

In this section, we test the third and fourth moments of the stochastic-volatility process, looking for evidence of jumps in volatility, as conjectured by Bates [1997]. The test statistics are constructed from

$$\begin{aligned} E_n(\epsilon_{n+1}^{v3}) = 0 \, , \quad \epsilon_n^{v3} &= (V_n^{\vartheta})^3 - M_8(V_{n-1}^{\vartheta}, \vartheta) \, , \\ E_n(\epsilon_{n+1}^{v4}) = 0 \, , \quad \epsilon_n^{v4} &= (V_n^{\vartheta})^4 - M_9(V_{n-1}^{\vartheta}, \vartheta) \, , \end{aligned} \tag{4.3}$$

where $M_8(V_n, \vartheta) = E_n(V_{n+1}^3)$ and $M_9(V_n, \vartheta) = E_n(V_{n+1}^4)$ are the third and fourth conditional moments of V.

Table 5 reports test results on the conditions $E_n(\epsilon_{n+1}^{v3}) = 0$ and $E_n(\epsilon_{n+1}^{v4}) = 0$ and their respective heteroskedasticity-corrected versions, $E_(\tilde{\epsilon}_{n+1}^{v3}) = 0$ and $E_n(\tilde{\epsilon}_{n+1}^{v4}) = 0$. The heteroskedasticity-corrected ($\tilde{\epsilon}$) tests evidently have more asymptotic power than their respective uncorrected (ϵ) counterparts, although both sets of tests build on the same moment conditions (4.3). In particular, the sample estimates of the moment conditions $E_n(\tilde{\epsilon}_{n+1}^{v3}) = 0$ and $E_n(\tilde{\epsilon}_{n+1}^{v4}) = 0$ are found to be positive and significantly different from zero for SVJ0 and SVJ models, indicating the possibility of jumps (with positive mean jump size) in the stochastic-volatility process.[34]

[33] Here, correlation is with respect to the stationary distribution. That is, the volatility process is assumed to start from its ergodic distribution, as opposed to the Dirac measure (with $V_0 = v$) that has been assumed in our empirical setting. For a large sample, this difference does not affect the discussion that follows.

[34] Examples of jump in stochastic volatility can be found in Duffie, Pan, and Singleton [1999]. Empirical findings with respect to such jumps-in-volatility models can be found in Eraker, Johannes, and Polson [1999].

Table 5: Tests on the Higher Moments of Stochastic Volatility

	SV0	SV	SVJ0	SVJ
$E(\epsilon_n^{v3}) = 0$	-0.69	0.85	0.70	0.54
$E(\epsilon_n^{v4}) = 0$	-0.38	0.65	0.62	0.53
joint $\chi^2(2)$	4.3 (0.12)	2.4 (0.31)	0.6 (0.72)	0.3 (0.86)
$E(\tilde{\epsilon}_n^{v3}) = 0$	-0.58	1.46	2.17^*	2.59^{**}
$E(\tilde{\epsilon}_n^{v4}) = 0$	0.95	1.68	2.67^{**}	3.75^{**}
joint $\chi^2(2)$	37.1^{**} (10^{-10})	4.0 (0.14)	7.3^* (0.03)	15.6^{**} (0.0004)

$*$ and $**$ indicate significance under a 5% and 1% test, respectively. For the individual tests, only the test statistics (standard normal in large sample) are reported. The p-values for the joint χ^2 tests are reported in parentheses.

Jumps in stochastic volatility provides a plausible explanation of how the model "undershoots" the data in terms of the third and fourth moments of V. Such a jump explanation is quite intuitive given the "spikes" in option-implied volatility shown in Figure 2. This, however, is not the only possible explanation. Other promising specifications are regime-switching in the stochastic volatility, and two-factor stochastic-volatility models, both of which generate fat-tailed transition distributions for stochastic volatility.

4.6 Option-Implied Volatility "Smiles"

This section extends the time-series results to cross-sectional option data. Using the estimates $\hat{\vartheta}^{SV0}$, $\hat{\vartheta}^{SV}$, $\hat{\vartheta}^{SVJ0}$, and $\hat{\vartheta}^{SVJ}$ of the respective SV0, SV, SVJ0, and SVJ model parameters, and constructing the corresponding date-n option-implied stochastic volatility estimates V_n^{SV0}, V_n^{SV}, V_n^{SVJ0}, and V_n^{SVJ}, we price the same cross-sectional set of options that is observed in the market on each date n. We compare the model-implied and market-observed option prices, examining the extent to which the parametric models, estimated exclusively from the joint time-series $\{S_n, C_n\}$, correctly price the entire cross-section of options.

A typical feature of the cross-sectional option data is the so-called volatility "smiles" or "smirks,"[35] whose pattern and shape vary from day to day. In the framework of our parametric models, a plausible explanation for such time variation in smile curves is that stochastic volatility varies from day to day. For this reason, we divide our sample days into three groups — days of high, medium, and low volatilities.

We select the 10 most and 10 least volatile days from the weekly sample between January 1989 to December 1996, as measured by the Black-Scholes implied volatility of $\{C_n\}$. For

[35] See the pioneer work of Rubinstein [1994] for further details. A more recent examination of the S&P 500 options market can be found in Derman [1999].

a comparison group of medium-volatility days, we select the ten successive days (at weekly intervals) between September 20, 1996 and November 22, 1996. The average Black-Scholes implied volatility (*BS vol*) for the days of high, medium, and low volatilities are 25.1%, 13.6%, and 8.7%, respectively. On each date n, we collect all bid and ask quotes of those call and put options that are time-stamped between 10:00am to 11:00am. For 1996, the time-window is reduced to 10:10am–10:20am, due to a surge in trading volume in 1996. Options with less than 15 days to expiration are discarded. This set of cross-sectional data is then filtered through the Black-Scholes option-pricing formula to obtain the corresponding *BS vol*, discarding any observation from which the *BS vol* cannot be obtained.[36]

Table 6 summarizes the cross-sectional pricing errors implied by the SV0, SV, SVJ0, and SVJ models. The pricing errors are measured as the absolute differences between the model-implied and market-observed Black-Scholes implied volatilities. (This avoids placing undue weight on expensive options, such as deep-in-the-money or longer-dated options.) As a reference to the degree of mis-pricing, we also report the average bid/ask spreads (difference between offer and ask prices, each measured in terms of *BS vol*). The positive and negative signs in the parentheses indicate whether, on average, the model over-prices or under-prices, respectively. The pricing errors and bid/ask spreads reported in Table 6 are obtained from 11,434 (high vol), 33,919 (medium vol), and 19,589 (low vol) sets of option data (bid/ask quotes on both call and put options). Figures 4 and 5 plot the volatility smiles the most and the least volatile days of our sample, respectively. Figure 6 plots the volatility smile on a medium-volatility day.

4.6.1 Smiles with SVJ0

Among the four parametric models, the SVJ0 model best captures the "smirkiness" of the cross-sectional data.

- On medium-volatility days, the SVJ0 model fits the volatility "smirks" remarkably well, across all times to expiration.

- The SVJ0 model under- and over-prices long-dated options on high- and low-volatility days, indicating a stochastic-volatility process that mean-reverts back to its long-run mean more slowly than suggested by the model. This is consistent with the time-series finding of "long memory" in volatility persistence, and again calls for a two-factor volatility model. Similar findings on the long memory of stochastic volatility implied by options data can be found in Stein [1989] and Bollerslev and Mikkelsen [1996].

- On high- and low-volatility days, the SVJ0 model under-prices the right side (with $k > 1.03$) of the smile curve for short-dated options. As Figure 4 and Figure 5 clarify, this mis-pricing corresponds to the "tipping-at-the-end" behavior, contributed mostly by deep in-the-money puts. This "tipping-at-the-end" behavior seems to require more randomness on the right tail of the underlying return distribution under the risk-neutral measure than suggested by the estimated SVJ0 model. A possible solution is to allow jumps in volatility, with jump arrivals that are more frequent, or with larger jump amplitudes, when volatility is low.

[36]For example, this happens when the quoted option price is less than its intrinsic value.

Table 6: The Model-Implied Pricing Errors for the S&P 500 Call and Put Options

		Days of High Volatility			Days of Medium Volatility			Days of Low Volatility		
		$k < 0.97$	[0.97,1.03]	$k > 1.03$	$k < 0.97$	[0.97,1.03]	$k > 1.03$	$k < 0.97$	[0.97,1.03]	$k > 1.03$
SV0	$\tau < 60$	5.4 (−)	2.1 (−)	5.0 (−)	2.9 (−)	0.8 (−)	0.8 (+)	2.9 (−)	0.9 (−)	4.6 (−)
SV		5.1 (−)	1.8 (−)	5.0 (−)	2.8 (−)	0.6 (−)	1.0 (+)	2.8 (−)	0.9 (−)	4.6 (−)
SVJ0		2.3 (−)	1.8 (−)	5.7 (−)	1.6 (+)	0.4 (−)	0.8 (−)	1.2 (+)	0.7 (−)	4.9 (−)
SVJ		3.1 (+)	2.4 (−)	7.0 (−)	3.2 (+)	0.8 (−)	1.6 (−)	2.3 (+)	0.7 (−)	5.1 (−)
bid/ask		3.2	1.3	3.9	2.7	1.0	2.3	1.1	0.5	2.8
SV0	$60 \le \tau < 180$	5.3 (−)	4.2 (−)	3.0 (−)	2.1 (−)	0.9 (−)	0.5 (+)	0.7 (−)	0.6 (+)	2.3 (−)
SV		2.8 (+)	2.6 (+)	4.2 (+)	1.5 (+)	2.1 (+)	3.1 (+)	1.5 (+)	2.2 (+)	3.2 (+)
SVJ0		3.4 (−)	3.6 (−)	3.9 (−)	0.6 (+)	0.3 (+)	0.4 (−)	2.1 (+)	1.1 (+)	2.5 (−)
SVJ		5.0 (−)	5.9 (−)	7.3 (−)	0.8 (+)	1.0 (−)	1.8 (−)	2.2 (+)	0.7 (+)	2.8 (−)
bid/ask		1.9	1.0	1.8	0.9	0.6	0.9	1.0	0.4	1.4
SV0	$\tau > 180$	5.5 (−)	3.9 (−)	3.3 (−)	1.7 (−)	0.9 (−)	0.3 (−)	0.8 (−)	0.6 (+)	1.6 (+)
SV		7.8 (+)	9.1 (+)	9.7 (+)	7.2 (+)	7.9 (+)	8.7 (+)	6.3 (+)	7.3 (+)	8.1 (+)
SVJ0		4.2 (−)	3.0 (−)	3.4 (−)	0.3 (+)	0.5 (+)	0.7 (+)	1.6 (+)	1.6 (+)	1.8 (+)
SVJ		6.4 (−)	5.6 (−)	6.6 (−)	1.4 (−)	1.2 (−)	1.5 (−)	0.7 (+)	0.5 (+)	0.5 (−)
bid/ask		1.7	1.0	1.3	0.8	0.7	0.6	0.5	0.4	0.6

All Prices are measured in 100 times the Black-Scholes implied volatility. Pricing errors are measured as the absolute differences between the model-implied and market-observed prices. "+" indicates that on average the model over-prices, and "−" indicates under-pricing. The bid/ask spreads are measured as the differences between the offer and ask prices.

4.6.2 Smiles with SVJ

- The SVJ model exaggerates the "smirkiness" of short- and medium-dated options. This exaggerated smirkiness implied by the SVJ model can be explained by its "over-stated" jump-risk premium, which is roughly twice that estimated for the SVJ0 model. (See Section 4.3.2.)

- On medium-volatility days, the SVJ model under-prices long-dated options. This can be explained by the negative volatility-risk premium estimated by the SVJ model to compensate for the "over-stated" jump-risk premium. For long-dated ones, the jump-risk premium becomes less important while the volatility-risk premium becomes more prominent. The negative volatility-risk premium estimated for the SVJ model therefore dominates at the long end, and under-prices long-dated options.

- The SVJ model suffers from the same mis-specification for term structure of volatility as the SVJ0 model. It therefore has the tendency to under-price and over-price long-dated options on high- and low-volatility days. This effect, however, is mixed the above two effects.

4.6.3 Smiles with SV

The SV model severely over-prices long-dated options. This over-pricing can be attributed to an overstated volatility-risk premium that "blows up" the stochastic-volatility process under the "risk-neutral" measure. This "explosive" nature of the "risk-neutral" volatility process is reported for the SV model in Section 4.3.1. Although, there seems to be no theoretical reason to rule out such behavior, its implication for long-dated options clearly rules out the SV model as a sensible candidate.

5 Concluding Remarks

We have uncovered, from the joint time-series of the S&P 500 index $\{S_t\}$ and of near-the-money short-dated option prices $\{C_t\}$, a jump-risk premium that is highly correlated with market volatility. This jump-risk premium is important, not only in reconciling the dynamics implied by $\{S_t, C_t\}$, but also in explaining changes over time in volatility "smiles" and "smirks."

We have examined four arbitrage-free stochastic-volatility models, two with jumps in returns and two without jumps. The joint time-series data $\{S_t, C_t\}$ strongly reject the pure-diffusion model of Heston [1993], but do not reject either of the two jump-diffusion models. Key to this goodness of fit of the jump-diffusion models is the presence of non-trivial jump-risk premia that respond quickly to market volatility. Our findings for the four models can be summarized as follows:

The SV0 Model: The Heston [1993] model, *without* volatility-risk premia, fails to accommodate the dynamics of $\{S_t\}$ and $\{C_t\}$. In particular, given the level of volatility observed in the spot market, this model severely under-prices options.

The SV Model: The Heston [1993] model, *with* volatility-risk premia, is still strongly rejected by the joint time series of spot and option prices. By introducing volatility-risk premia, this model partially resolves the "under-pricing" problem of the SV0 model, but with an unrealistically high level of volatility-risk premia. Given the estimated volatility-risk premia, the stochastic volatility is explosive (negative mean-reversion) under the "risk-neutral" measure. Consequently, long-dated options are severely overpriced by this model.

The SVJ0 Model: Among the four models examined, this variant of the Bates [1997] model, incorporating premia for jump risk, is found to best characterize both the transition behavior of the joint time series $\{S_t, C_t\}$ and the behavior of "smiles" and "smirks." Even for this model, our diagnostic tests indicate mis-specification for the term structure of volatility, as well as poor fits for the third and fourth conditional moments of volatility. Both findings appear to call for a stochastic-volatility model with two factors — one strongly persistent, the other quickly mean-reverting and highly volatile. The second finding also indicates the possibility of jumps in volatility.

The SVJ Model: This model nests of all the model specifications considered, but does not perform as well as the SVJ0 model. In particular, the SVJ model overstates the level of jump-risk premia (relative to that estimated for the SVJ0 model), compensating for its "overstated" jump-risk premia with negative volatility-risk premia. Such overstated jump-risk premia result in exaggerated volatility "smirks" for short- and medium-dated options, and in a tendency for the SVJ model, through its negative volatility-risk premia, to under-price long-dated options. Like the SVJ0 model, the SVJ model suffers from its mis-specification of the term structure of volatility, and provides poor fits for the higher conditional moments of volatility.

Finally, this empirical study also suggests that an important component of the equity risk premium can be attributed to investors' aversion to jump risk. Examining the S&P 500 index and options markets jointly, our estimates of the SVJ0 model imply that, in sample and on average, the excess mean rate of return demanded for jump risk is 6.4% per year, while that demanded for the usual "diffusive" return risk is 4.2% per year. As shown in Mehra and Prescott [1985], the overall equity risk premium is too high to be assigned by a power-utility maximizer using aggregated consumption data. (See Rietz [1988] for a "peso" explanation.) In order to capture the large premium demanded for jump risk, it may be fruitful to explore utility models showing potentially extreme aversion to big losses (as in, for example, Gul [1991]).

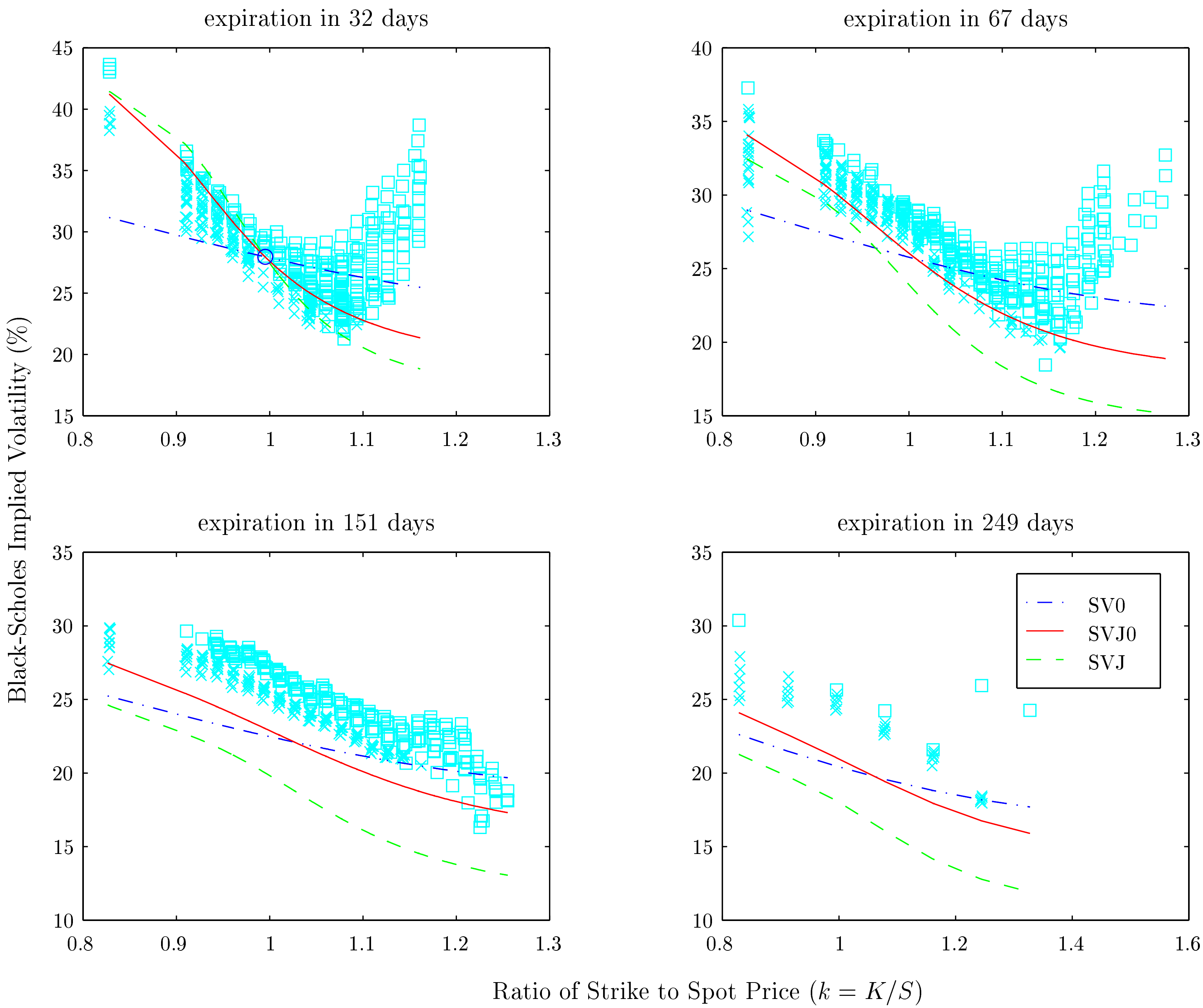

Figure 4: Smiles curves on a "high-volatility" day. All observations are observed between 10:00am to 11:00am on October 16, 1990. The call options are marked by '×,' and the put options by '□'.

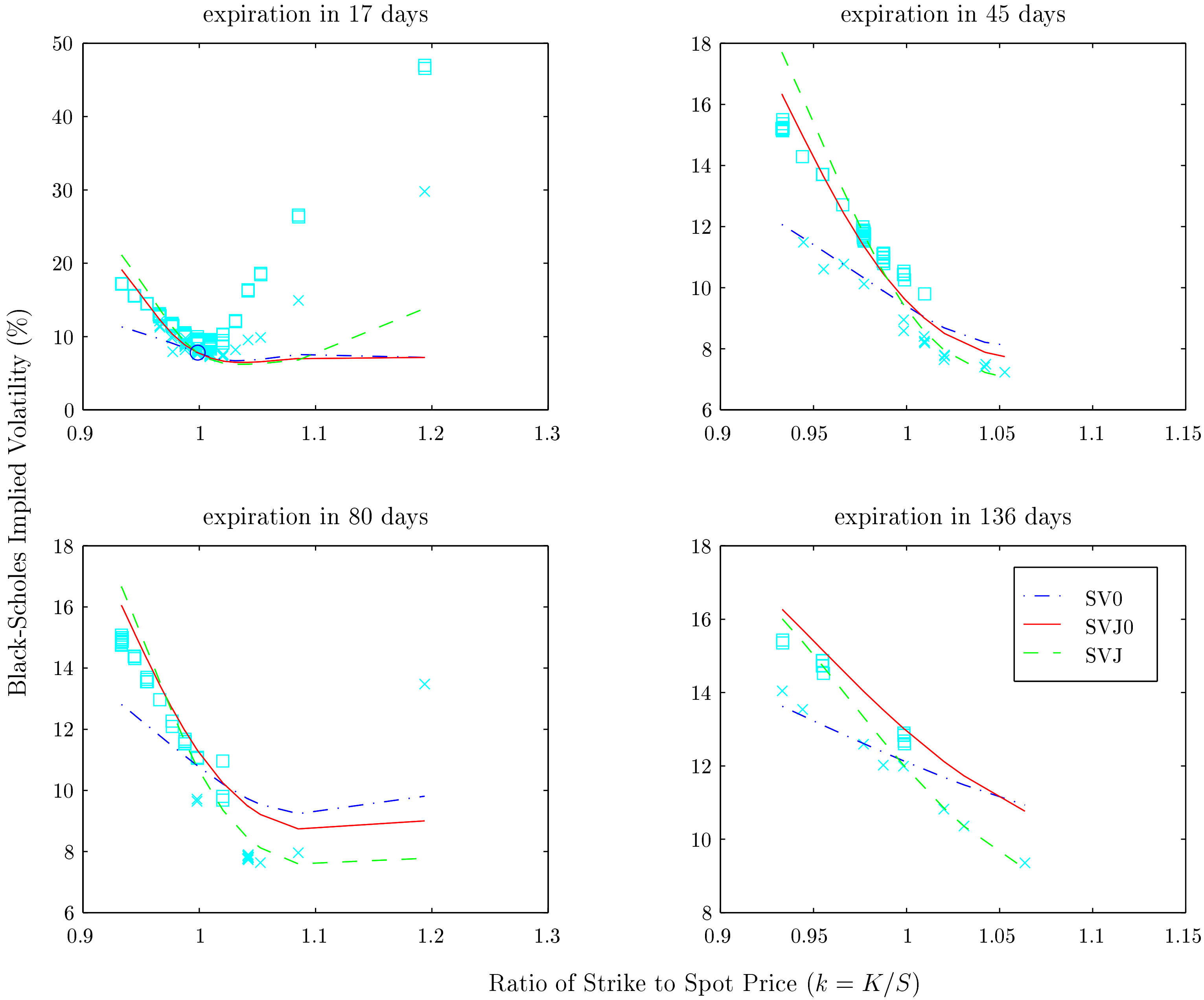

Figure 5: Smiles curves on the least volatile day of the sample. All observations are observed between 10:00am to 11:00am on August 3, 1994. The call options are marked by '$\times$,' and the put options by '$\Box$'.

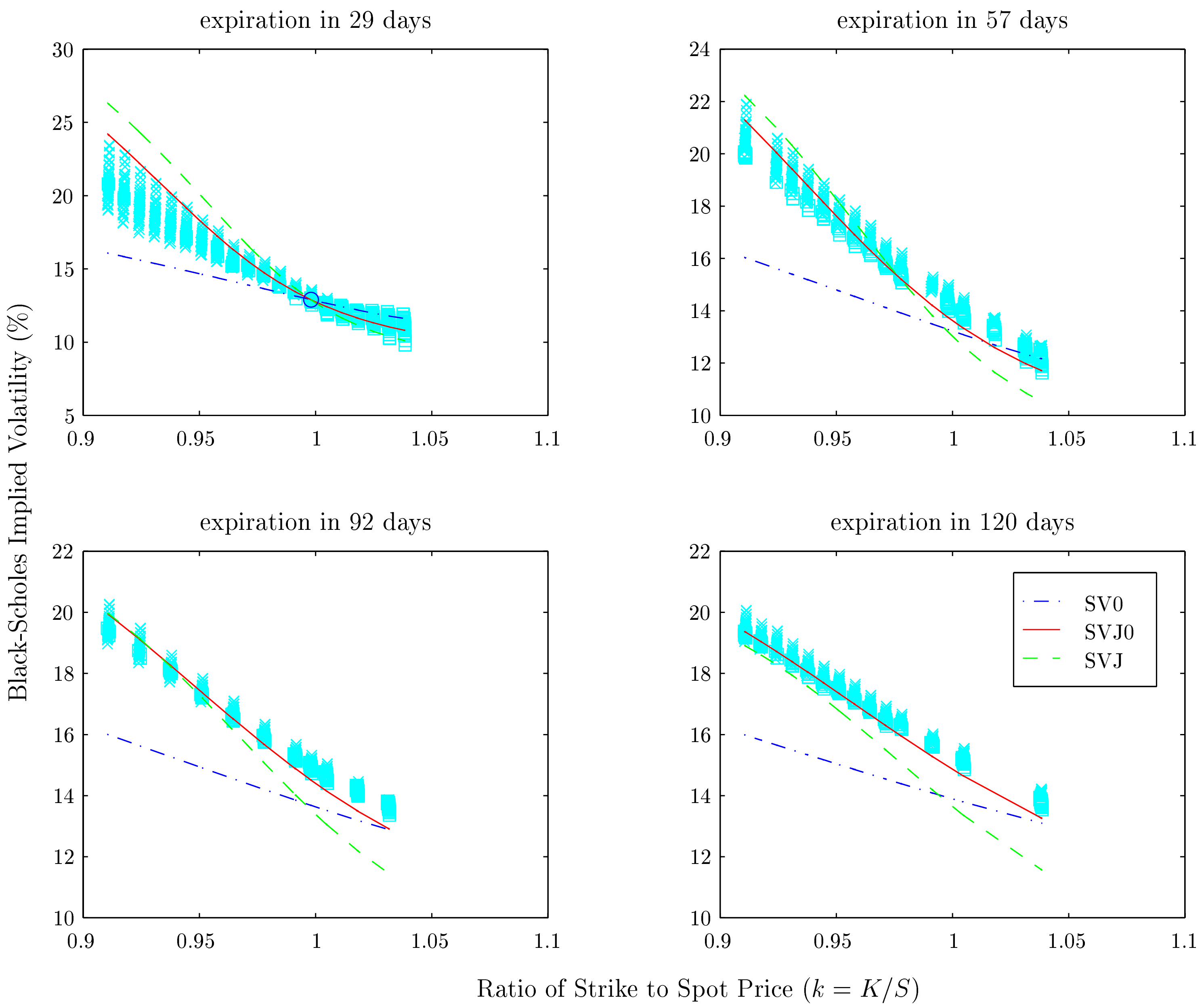

Figure 6: Smiles curves on a "medium-volatility" day. All observations are observed between 10:10am to 10:20am on November 22, 1996. The call options are marked by '×,' and the put options by '□'.

Appendices

A Technical Specification of the SVJ model

A more precise specification of the state process $X = (\ln S, V, r, q)$ defined by (2.1) and (2.2) is via its infinitesimal generator $\mathcal{D}$. For any smooth bounded function $f : \mathbb{R}\times\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}_+ \to \mathbb{R}$ with bounded derivatives, $\{f(X_t) - \int_0^t \mathcal{D}f(X_u)\,du \;:\; t \geq 0\}$ is a martingale with $\mathcal{D}f(x)$ defined, at any $x = [\ln s, v, r, q]^\top$, by

$$\begin{aligned}\mathcal{D}f(x) &= f_x(x)\,(\mathcal{K}_0 + \mathcal{K}_1\,x) + \frac{1}{2}\mathrm{tr}\left[f_{xx}(x)\,\mathcal{K}_2\,\right] \\ &\quad + (\lambda_0 + \lambda_1 v)\int_{\mathbb{R}} \left[f(x + [z,0,0,0]^\top) - f(x)\right]\,d\nu(z)\,,\end{aligned}$$

where $\nu(\cdot)$ is the cumulative distribution of a normal random variable with mean μ_J and variance σ_J^2, and where

$$\mathcal{K}_0 = \begin{pmatrix} -\lambda_0\mu^* \\ \kappa_v\bar{v} \\ \kappa_r\bar{r} \\ \kappa_q\bar{q}\end{pmatrix}, \quad \mathcal{K}_1 = \begin{pmatrix} 0 & \eta^s - \frac{1}{2} - \lambda_1\mu^* & 1 & -1 \\ 0 & -\kappa_v & 0 & 0 \\ 0 & 0 & -\kappa_r & 0 \\ 0 & 0 & 0 & -\kappa_q \end{pmatrix}, \quad \mathcal{K}_2 = \begin{pmatrix} v & \sigma_v\rho\,v & 0 & 0 \\ \sigma_v\rho\,v & \sigma_v^2\,v & 0 & 0 \\ 0 & 0 & \sigma_r^2\,r & 0 \\ 0 & 0 & 0 & \sigma_q^2\,q \end{pmatrix}.$$

Similarly, for the state process $X = (\ln S, V, r, q)$ under the "risk-neutral" measure Q defined in Section 2.3, an alternative specification can be achieved through its infinitesimal generator $\mathcal{D}^*$ defined by

$$\begin{aligned}\mathcal{D}^*f(x) &= f_x(x)\,(\mathcal{K}_0 + \mathcal{K}_1^*\,x) + \frac{1}{2}\mathrm{tr}\left[f_{xx}(x)\,\mathcal{K}_2\,\right] \\ &\quad + (\lambda_0 + \lambda_1 v)\int_{\mathbb{R}} \left[f(x + [z,0,0,0]^\top) - f(x)\right]\,d\nu^*(z)\,,\end{aligned}$$

where $\nu^*(\cdot)$ is the cumulative distribution of a normal random variable with mean μ_J^* and variance σ_J^2, and where

$$\mathcal{K}_1^* = \begin{pmatrix} 0 & -\frac{1}{2} - \lambda_1\mu^* & 1 & -1 \\ 0 & -\kappa_v + \eta^v & 0 & 0 \\ 0 & 0 & -\kappa_r & 0 \\ 0 & 0 & 0 & -\kappa_q \end{pmatrix}.$$

B The State-Price Density and No Arbitrage

This appendix focuses on the relationship between the state-price density and no arbitrage, and shows that our model is arbitrage free. Let $\mathcal{S}$ and $\mathcal{B}$ be the respective gain processes of the underlying security and the bank account, defined in Section 2.2. Let π be the state-price density process defined by (2.3), chosen so that the deflated gain processes $\mathcal{S}^\pi$ and $\mathcal{B}^\pi$

are local martingales. Now let $C = \left(C^{(1)}, \ldots, C^{(n)}\right)$ denote the process process of any n options, priced by the state-price density π, as in (2.9). The deflated option price process $C^\pi = \pi C$ is also a local martingale; indeed, it is a martingale. Letting θ be a self-financing trading strategy with respect to $X = \left(\mathcal{S}, \mathcal{B}, C^{(1)}, \ldots, C^{(n)}\right)$, we show in this appendix that if θ satisfies the full-collateralization (no-credit) constraint $\theta_t \cdot X_t \geq 0$, for all t, then θ is not an arbitrage.[37]

A *deflater* is a strictly positive semi-martingale (RCLL). For example, the state-price density π is a deflater. We first establish in Lemma B1 the result of numeraire invariance for an arbitrary deflater Y. This result extends from the Ito processes considered in Duffie [1996] to general semi-martingales, so that jumps can be considered. Similar results can be found in Huang [1985] and, independently, Protter [1999]. As usual, a trading strategy is a predictable process θ such that $\int \theta \, dX$ is well defined as a stochastic integral, and is self-financing if, for all t, $\theta_t \cdot X_t = \theta_0 \cdot X_0 + \int_0^t \theta_s \, dX_s$. (See Harrison and Pliska [1981].)

Lemma B.1 (Numeraire Invariance) *Suppose Y is a deflater. A trading strategy θ is self-financing with respect to X if and only if θ is self-financing with respect to the deflated process $X^Y = XY$.*

Proof: Let $w_t = \theta_0 \cdot X_0 + \int_0^t \theta_s \, dX_s$, $t \in [0, T]$. Let w^Y be the π-deflated process defined by $w_t^Y = w_t Y_t$. We note that $w_t = \theta_t \cdot X_t$ is implied by either: (1) assuming that θ is X self-financing, or (2) assuming that θ is X^Y self-financing. Using Ito's Formula for semi-martingales (Protter [1990]),

$$dw_t^Y = Y_{t-} \, dw_t + w_{t-} \, dY_t + d[w, Y]_t^c + \Delta w_t \, \Delta Y_t \,,$$

where $[w, Y]^c$ denotes the continuous part of the cross-variation process $[w, Y]$. Using the result that $\Delta w_t = \theta_t \cdot \Delta X_t$ (Protter [1990], Theorem 18, page 135), we have $w_{t-} = w_t - \Delta w_t = w_t - \theta_t \Delta X_t$. It then follows that $w_{t-} = \theta_t \cdot X_{t-}$, and that

$$\begin{aligned} dw_t^Y &= Y_{t-} \theta_t \, dX_t + \theta_t \cdot X_{t-} \, dY_t + \theta_t \cdot d[X, Y]_t^c + \theta_t \cdot \Delta X_t \, \Delta Y_t \\ &= \theta_t \cdot \left(Y_{t-} \, dX_t + X_{t-} \, dY_t + d[X, Y]_t^c + \Delta X_t \, \Delta Y_t \right) = \theta_t \, dX_t^Y \,. \end{aligned}$$

Thus, $\theta_t \cdot X_t^Y = \theta_0 \cdot X_0^Y + \int_0^t \theta_s \, dX_s^Y$ if and only if $\theta_t \cdot X_t = \theta_0 \cdot X_0 + \int_0^t \theta_s \, dX_s$, establishing the result.

Proposition B.1 (No Arbitrage) *If θ is self-financing with respect to X and $\theta_t \cdot X_t \geq 0$, and if π is a deflator such that X^π is local martingale, then θ is not an arbitrage.*

Proof: By Lemma B.1, since θ is self-financing with respect to X, θ is also self-financing with respect to X^π. That is, $\theta_t \cdot X_t^\pi = \theta_0 \cdot X_0^\pi + \int_0^t \theta_s \, dX_s^\pi$. We know that $\int \theta \, dX^\pi$ is local martingale because X^π is. It then follows from the self-financing property of θ that $\{\theta_t \cdot X_t^\pi\}$ is also a local martingale. Moreover, $\{\theta_t \cdot X_t^\pi\}$ is a nonnegative local martingale, as $\theta_t \cdot X_t \geq 0$

[37] A self-financing strategy ϑ is an arbitrage if $\theta_0 \cdot X_0 < 0$ and $\theta_T \cdot X_T \geq 0$, or $\theta_0 \cdot X_0 \leq 0$ and $\theta_T \cdot X_T > 0$.

implies $\theta_t \cdot X_t^\pi \geq 0$. Since a local martingale that is bounded below is a supermartingale (Revuz and Yor [1991], page 117), we know that $E(\theta_T \cdot X_T^\pi) \leq \theta_0 \cdot X_0^\pi$. Therefore θ is not an arbitrage with respect to X^π. By Lemma B.1 and the definition of an arbitrage, θ is not an arbitrage with respect to X.

C A Sufficient Condition for the Martingality of ξ

We provide a Novikov-like sufficient condition for the exponential local martingale ξ define by (2.5) to be a martingale.

PROPOSITION C.1 *(i) A sufficient condition for ξ, defined by (2.5), to be a martingale is that*

$$E_x\left[\exp\left(\int_0^T \zeta_t \cdot \zeta_t \, dt\right)\right] < 0\,, \tag{C.1}$$

where E_x denotes expectation with respect to initial condition $x \in \mathbb{R}_+$ for V, and ζ is the market price of Brownian shocks defined by (2.4). (ii) Condition (C.1) holds if

$$\frac{1}{1-\rho^2}\left((\eta^s)^2 + 2\rho\,\eta^s\,\frac{\eta^v}{\sigma_v} + \left(\frac{\eta^v}{\sigma_v}\right)^2\right) < \left(\frac{\kappa_v}{\sigma_v}\right)^2. \tag{C.2}$$

Proof: We first show (i). Using the fact that U_i^π and U_j^π are independent for $i \neq j$, that, for any i, U_i^π is independent of W and of the jump times $\{\mathcal{T}_i\}$, and that $E\left(\exp\left(U_i^\pi\right) - 1\right) = 0$, we have, for $0 \leq t \leq s \leq T$,

$$E_t\left(\xi_s\right) = E_t\left[\mathcal{E}\left(-\int_0^s \zeta_u\, dW_u\right)\exp\left(\sum_{i,\mathcal{T}_i \leq s} U_i^\pi\right)\right] = \xi_t\, E_t\left[\mathcal{E}\left(-\int_t^s \zeta_u\, dW_u\right)\right],$$

where E_t denotes $\mathcal{F}_t$-conditional expectation, and $\mathcal{E}$ denotes the stochastic exponential. Using Novikov [1972], under (C.1), $\left\{\mathcal{E}\left(-\int_0^t \zeta_u\, dW_u\right) : 0 \leq t \leq T\right\}$ is a martingale. Then (i) follows immediately from the fact that $E_t\left[\mathcal{E}\left(-\int_t^s \zeta_u\, dW_u\right)\right] = 1$ for any $0 \leq t \leq s \leq T$.

Next, we show that (ii) holds. Letting

$$L^2 = \frac{1}{1-\rho^2}\left((\eta^s)^2 + 2\rho\,\eta^s\,\frac{\eta^v}{\sigma_v} + \left(\frac{\eta^v}{\sigma_v}\right)^2\right),$$

(C.1) is equivalent to

$$E_x\left[\exp\left(\int_0^T \frac{1}{2}L^2 V_t\, dt\right)\right] < \infty.$$

Under (C.2), $\gamma = \sqrt{1 - L^2\,\sigma_v^2/\kappa_v^2}$ is real-valued and $0 < \gamma \leq 1$. We conjecture (based on the affine structure of V) that,

$$E_x\left[\exp\left(\int_0^T \frac{1}{2}L^2 V_t\, dt\right)\right] = \exp\left[\alpha(T) + \beta(T)v\right], \tag{C.3}$$

where

$$\beta(T) = \frac{\kappa_v}{\sigma_v^2} \frac{(1-\gamma^2)(1-\exp(-\gamma\kappa_v T))}{(1+\gamma)-(1-\gamma)\exp(-\gamma\kappa_v T)}, \quad \alpha(T) = \kappa_v \theta \int_0^T \beta(t)\, dt\,.$$

We are done if we can show this conjecture holds. A sufficient condition is that[38]

$$E_x\left(\int_0^T e^{2\beta(t)V_t} V_t\, dt\right) = \int_0^T E_x(e^{2\beta(t)V_t} V_t)\, dt < \infty\,. \tag{C.4}$$

The equality holds, by Fubini, assuming that $\sup_t E^x(e^{2\beta(t)V_t}V_t) < \infty$. To show this, we consider the probability density $p^x(t,v) = P^x(V_t \in dv)$ of V_t, found in Feller [1951]. One can show, that for large $v \in \mathbb{R}_+$, the asymptotic behavior of p^x is

$$p^x(t,v) \sim \exp\left(-\omega(t)\, v\right)\,, \quad \omega(t) = \frac{2\kappa_v}{\sigma_v^2}\left(1-\exp\left(-\kappa_v t\right)\right)^{-1}.$$

The integrability condition (C.4) follows from the fact that $2\beta(t) < 2\kappa_v/\sigma_v^2 < \omega(t)$, and the fact that $\{\omega(t) : 0 \le t \le T\}$ is bounded.

D Explicit Formulae for A, B, α, β_v, β_r, and β_q

The coefficients B and A of (3.14) are defined by

$$\begin{aligned}
B(u_y,u_v) &= -\frac{a\left(1-\exp(-\gamma\Delta)\right) - u_v\left[2\gamma - (\gamma-b)(1-\exp\left(-\gamma\Delta\right)\right]}{2\gamma - (\gamma+b)\left(1-\exp(-\gamma\Delta)\right) - u_v\sigma_v^2\left(1-\exp(-\gamma\Delta)\right)}\,,\\
A(u_y,u_v) &= -\frac{\kappa_v\bar{v}}{\sigma_v^2}\left((\gamma+b)\Delta + 2\ln\left[1 - \frac{\gamma+b+\sigma_v^2\, u_v}{2\gamma}(1-e^{-\gamma\Delta})\right]\right)\\
&\quad + \left(\exp\left(u_y\mu_J + \frac{u_y^2\sigma_J^2}{2}\right) - 1 - u_y\mu^*\right)\lambda_0\Delta\,,
\end{aligned} \tag{D.1}$$

and where $b = \sigma_v\rho u_y - \kappa_v$, $a = -u_y^2 - 2u_y\left[\eta^s - 1/2 - \lambda_1\mu^*\right] - 2\lambda_1\left(\exp(u_y\mu_J + u_y^2\sigma_J^2/2) - 1\right)$, and $\gamma = \sqrt{b^2 + a\sigma_v^2}$.

The coefficients α_v and β_v of (2.10) are defined by

$$\begin{aligned}
\beta_v(c,t,\vartheta) &= -\frac{a\left(1-\exp(-\gamma_v t)\right)}{2\gamma_v - (\gamma_v+b)\left(1-\exp(-\gamma_v t)\right)}\\
\alpha_v(c,t,\vartheta) &= -\frac{\kappa_v^*\bar{v}^*}{\sigma_v^2}\left((\gamma_v+b)\,\tau + 2\ln\left[1 - \frac{\gamma_v+b}{2\gamma_v}\,\left(1-e^{-\gamma_v\tau}\right)\right]\right)\\
&\quad + \lambda_0\, t\left(\exp\left(c\mu_J^* + \frac{c^2\sigma_J^2}{2}\right) - 1 - c\mu^*\right)\,,
\end{aligned} \tag{D.2}$$

where $b = \sigma_v\rho c - \kappa_v^*$, $a = c(1-c) - 2\lambda_1\left[\exp(c\mu_J^* + c^2\sigma_J^2/2) - 1 - c\mu^*\right]$, and $\gamma_v = \sqrt{b^2 + a\sigma_v^2}$. The parameters superscripted by $*$ denote the risk-neutral counterparts of those under the

[38]See, for example, Duffie, Pan, and Singleton [1999] for details.

data-generating measure P. For example, $\kappa_v^* = \kappa_v - \eta^v$ and $\bar{v}^* = \kappa_v \bar{v}/\kappa_v^*$ are the risk-neutral mean-reversion rate and long-term mean, respectively, and $\mu_J^* = \ln(1+\mu^*) - \sigma_J^2/2$ is the risk-neutral counterpart of μ_J. While the square root and logarithm of a complex number z are not uniquely defined, for notational simplicity the results are presented as if we are dealing with real numbers. To be more specific, we define, $\sqrt{z} = |z|^{1/2} \exp\left(\frac{i \arg(z)}{2}\right)$ and $\ln(z) = \ln|z| + i \arg(z)$, where for any $z \in \mathbb{C}$, $\arg(z)$ is defined such that $z = |z| \exp(i \arg(z))$, with $-\pi < \arg(z) \leq \pi$.

The coefficients α_r and β_r are defined by

$$\begin{aligned} \beta_r(c,t,\theta_r) &= -\frac{2(1-c)\,(1-\exp(-\gamma_r t))}{2\gamma_r - (\gamma_r - \kappa_r)\,(1-\exp(-\gamma_r t))} \\ \alpha_r(c,t,\theta_r) &= -\frac{\kappa_r \bar{r}}{\sigma_r^2}\left((\gamma_r - \kappa_r)\,\tau + 2\ln\left[1 - \frac{\gamma_r - \kappa_r}{2\gamma_r}\,\left(1 - e^{-\gamma_r \tau}\right)\right]\right)\,, \end{aligned} \tag{D.3}$$

with $\gamma_r^2 = \kappa_r^2 + 2(1-c)\sigma_r^2$.

Next, α_q and β_q are defined by

$$\begin{aligned} \beta_q(c,t,\theta_q) &= -\frac{2c\,(1-\exp(-\gamma_q t))}{2\gamma_q - (\gamma_q - \kappa_q)\,(1-\exp(-\gamma_q t))} \\ \alpha_q(c,t,\theta_q) &= -\frac{\kappa_q \bar{q}}{\sigma_q^2}\left((\gamma_q - \kappa_q)\,\tau + 2\ln\left[1 - \frac{\gamma_q - \kappa_q}{2\gamma_q}\,\left(1 - e^{-\gamma_q \tau}\right)\right]\right)\,, \end{aligned} \tag{D.4}$$

with $\gamma_q^2 = \kappa_q^2 + 2c\sigma_q^2$.

Finally, we let $\alpha(c,t,\vartheta,\theta_r,\theta_q) = \alpha_v(c,t,\vartheta) + \alpha_r(c,t,\theta_r) + \alpha_q(c,t,\theta_q)$.

E Option Pricing via Numerical Integration

The improper probabilities $\mathcal{P}_1$ and $\mathcal{P}_2$ defined by (2.12) are key to determining the time-t price C_t of an option with time τ to expiration and strike-to-spot ratio k. This appendix provides a fast numerical scheme, with error analysis, for the inversion (2.12), assuming that the transform $\psi(c,v,r,q,\tau)$ defined by (2.10) is explicitly known. It should be noted that, whenever applicable, all of expectations and probability calculations in this appendix are taken with respect to the risk-neutral measure Q.

Fixing today at time t, we write, for national simplicity, $\psi(c) = \psi(c, V_t, r_t, q_t, \tau)$, where V_t, r_t, and q_t are today's volatility, risk-free short rate, and dividend yield. First, consider $\mathcal{P}_1 = \psi(1)\,\tilde{\mathcal{P}}_1$, where as can be seen from the CIR discount formula,

$$\psi(1) = E_t\left[\exp\left(-\int_t^{t+\tau} q_s\,ds\right)\right] = \exp\left(\alpha_q(1,\tau,\theta_q) + \beta_q(1,\tau,\theta_q)\,q_t\right)\,, \tag{E.1}$$

where α_q and β_q are as defined in (D.4). Effectively, $\psi(1)$ is the dividend analogue of a "τ-period bond price." Thus defined $\tilde{\mathcal{P}}_1$ is a real probability that can be calculated through the standard Lévy inversion formula

$$\tilde{\mathcal{P}}_1 = P\left(\tilde{X}_1 \leq \bar{x}\right) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\tilde{\psi}_1(u)\exp\left(-iu\bar{x}\right)\right)}{u}\,du\,, \tag{E.2}$$

where $\bar{x} = (r_t - q_t)\tau - \ln k$, and where the random variable $\tilde{X}_1$ is uniquely defined by its characteristic function $\tilde{\psi}_1(u)$ via

$$\tilde{\psi}_1(u) = \frac{\psi(1-iu)\exp(iu(r_t-q_t)\tau)}{\psi(1)}\,. \tag{E.3}$$

In practice, the Lévy inversion (E.2) is carried out via some form of numerical integration. Letting $I_1(u) = \mathrm{Im}\left(\tilde{\psi}_1(u)\exp(-iu\bar{x})\right)$ denote the integrand, we approximate by

$$\tilde{\mathcal{P}}_1 \approx \frac{1}{2} - \frac{1}{\pi}\sum_{n=0}^{[U_1/\Delta u_1]} \frac{I_1((n+1/2)\Delta u_1)}{n+1/2}, \tag{E.4}$$

where $[x]$ is an integer such that $[x]-1 < x \leq [x]$. Two types of errors are introduced by this numerical scheme. For any $U_1 < \infty$, there is a *truncation error*. For any $\Delta u_1 > 0$, there is a *discretization error*. To achieve any desired precision δ for $\tilde{\mathcal{P}}_1$, we can select a cutoff level U_1 such that

$$\textit{truncation error} = \left|\frac{1}{\pi}\int_{U_1}^{\infty}\frac{I_1(u)}{u}\,du\right| \leq \delta\,. \tag{E.5}$$

We can select a step size Δu_1 such that

$$\textit{discretization error} \leq \max\left[P\left(\tilde{X}_1 < \bar{x} - \frac{2\pi}{\Delta u_1}\right), P\left(\tilde{X}_1 > \bar{x} + \frac{2\pi}{\Delta u_1}\right)\right] \leq \delta\,, \tag{E.6}$$

where the first inequality follows from a Fourier analysis. See, for example, Davies [1973].

To control for the *truncation error*, we take advantage of the fact that $I(u)$ is explicit, and study its asymptotic behavior for large u. In particular, we can show that, for large enough u, $|I_1(u)| \leq \exp\left(-uA_1 + A_0\right)$ where $A_1 = (v + \bar{v}^*\kappa^*\tau)\sqrt{1-\rho^2}/\sigma_v$, and $A_0 = (v + \bar{v}^*\kappa^*\tau)\left(\kappa^* - \sigma_v\rho\right)/\sigma_v^2 + \ln\left(4(1-\rho^2)\right)\kappa^*\bar{v}^*/\sigma_v^2$. For the desired accuracy δ, we can therefore choose U_1 such that

$$\frac{1}{\pi A_1 U_1}\exp\left(-A_1U_1 + A_0\right) \leq \delta.$$

To control for the *discretization error*, we focus on the probabilities $P\left(\tilde{X}_1 < \bar{x} - 2\pi/\Delta u_1\right)$ and $P\left(\tilde{X}_1 > \bar{x} + 2\pi/\Delta u_1\right)$, which sample further into the left and right tails as Δu_1 approaches to zero. Given that the mean μ_{X1} and variance $\sigma^2_{X_1}$ of $\tilde{X}_1$ are finite, the tail probabilities can be controlled by Chebyshev's inequality:

$$P\left(|\tilde{X}_1 - \mu_{X_1}| > \frac{\sigma_{X_1}}{\sqrt{\delta}}\right) < \delta. \tag{E.7}$$

We can therefore establish an upper bound in probability for the two tail events $\{\tilde{X}_1 - \mu_{X_1} > \sigma_{X_1}/\sqrt{\delta}\}$ and $\{\tilde{X}_1 - \mu_{X_1} < -\sigma_{X_1}/\sqrt{\delta}\}$. The discretization step Δu_1 can be chosen such that

$$\frac{2\pi}{\Delta u_1} = \max\left(\bar{x} - \mu_{X_1}, \mu_{X_1} - \bar{x}\right) + \frac{\sigma_{X_1}}{\sqrt{\delta}}. \tag{E.8}$$

To calculate the mean and variance of $\tilde{X}_1$, we again take advantage of its explicitly known characteristic function $\tilde{\psi}_1(\cdot)$. Specifically, for any $u \in \mathbb{R}$, the moment-generating function of $\tilde{X}_1$ is $E\left[\exp\left(u\tilde{X}_1\right)\right] = \tilde{\psi}_1(-iu)$, from which its mean and variance can be derived accordingly.

The numerical integration scheme used for $\mathcal{P}_2$ is similar. Details are omitted, and are available upon request.

F Conditional Moments of the SVJ Model

Let (S, V, r, q) be the state process defined by (2.1) and (2.2). For a fixed time horizon Δ, and for some arbitrary non-negative integers i and j, this appendix provides a computational method for $\mathcal{F}_t$-conditional moments of the form $E_t\left(y^i_{t+\Delta}V^j_{t+\Delta}\right)$, where $y_t = \ln S_t - \ln S_{t-\Delta} - \int_{t-\Delta}^t (r_u - q_u)\,du$ is the time-t Δ-period excess return.[39] The joint moments of y and V can be calculated recursively by[40]

$$E_t\left(y^0_{t+\Delta}V^m_{t+\Delta}\right) = \sum_{j=0}^{m-1} C^j_{m-1}\, E_t\left(y^0_{t+\Delta}V^j_{t+\Delta}\right)\; p^{(0,m-j)}_{y,v}(V_t)\,, \qquad m \geq 1\,,$$
$$E_t\left(y^n_{t+\Delta}V^m_{t+\Delta}\right) = \sum_{i=0}^{n-1}\sum_{j=0}^{m} C^i_{n-1}C^j_m\, E_t\left(y^i_{t+\Delta}V^j_{t+\Delta}\right)\; p^{(n-i,m-j)}_{y,v}(V_t)\,, \qquad n \geq 1,\ m \geq 0\,, \tag{F.1}$$

where, for any $n \geq 0$ and $0 \leq i \leq n$, $C^i_n = n!/i!(n-i)!$, and where

$$p^{(i,j)}_{y,v}(V_t) = A^{(i,j)}_{y,v} + B^{(i,j)}_{y,v}V_t\,, \tag{F.2}$$

where $A^{(i,j)}_{y,v}$ and $B^{(i,j)}_{y,v}$ are constants that can be derived in a recursive fashion, as follows.

We first derive $B^{(i,j)}_{y,v}$ for $i \geq 0$ and $j \geq 0$. With "initial" values of $B^{(0,1)}_{y,v} = \exp(-\kappa\Delta)$, $B^{(1,0)}_{y,v} = \left(\eta^s - \frac{1}{2} + \lambda_1(J_1 - \mu^*)\right) f_0$, $B^{(2,0)}_{y,v} = (1 + \lambda_1 J_2)f_0 - f_1 B^{(1,0)}_{y,v}$, and $B^{(1,1)}_{y,v} = \sigma_v\rho f_0 + \frac{1}{2}\kappa f_0 f_1 + \frac{1}{2}\sigma_v^2 f_0 B^{(1,0)}_{y,v}$, the following formulas enable us to calculate $B^{(i,j)}_{y,v}$ recursively up to any order. We have

$$B^{(0,m)}_{y,v} = \frac{m}{2} f_0 \sigma_v^2 B^{(0,m-1)}_{y,v}\,, \quad m \geq 2\,,$$
$$B^{(n,0)}_{y,v} = \lambda_1 J_n f_0 - \frac{1}{2}\sum_{i=1}^{n-1} C^i_n B^{(i,0)}_{y,v} f_{n-i}\,, \quad n \geq 3\,,$$
$$B^{(n,1)}_{y,v} = \frac{1}{2}\kappa f_0 f_n + \frac{1}{2}\sigma_v^2 f_0 B^{(n,0)}_{y,v} - \frac{1}{2}\sum_{i=1}^{n-1} C^i_n f_i B^{(n-i,1)}_{y,v}\,, \quad n \geq 2\,,$$
$$B^{(n,m)}_{y,v} = \frac{m}{2}\sigma_v^2 f_0 B^{(n,m-1)}_{y,v} - \frac{1}{2}\sum_{i=1}^{n} C^i_n f_i B^{(n-i,m)}_{y,v}\,, \quad n \geq 1\,, \quad m \geq 2\,,$$

[39] This is a slight abuse of notation, as the Δ-period excess return y defined by (3.1) is indexed by integer n, while here y is indexed by time t.

[40] For pure affine diffusions, an alternative approach can be found in Liu [1997]. Das and Sundaram [1999] provide central moments of y, up to the fourth order, for the special case of $\lambda_1 = 0$.

where: $J_1 = \mu_J$, $J_2 = \sigma_J^2 + \mu_J^2$, and $J_n = J_{n-1}\mu_J + (n-1)J_{n-2}\sigma_J^2$ (for $n \geq 3$) are the moments of the jump amplitude. The coefficients f_i and g_i are given by

$$g_0 = 2\,, \quad g_n = 2\gamma_n + \frac{1}{f_0}\sum_{i=1}^{n-1}\Gamma_{n-i}\,g_i\,, \quad n \geq 1\,,$$

$$f_0 = \frac{1-\exp(-\kappa\Delta)}{\kappa}\,, \quad f_1 = g_1 - (\kappa\gamma_1 + \sigma_v\rho)f_0\,, \quad f_n = g_n - \kappa\gamma_n f_0\,, \quad n \geq 2\,,$$

with

$$\gamma_1 = -\left(\frac{\sigma_v}{\kappa}\right)^2\left(\rho\frac{\kappa}{\sigma_v} + \eta^s - \frac{1}{2} + \lambda_1(J_1 - \mu^*)\right)\,, \quad \gamma_2 = -\gamma_1^2 - \left(\frac{\sigma_v}{\kappa}\right)^2\left(1 - \rho^2 + \lambda_1 J_2\right)\,,$$

$$\gamma_n = -\sum_{i=1}^{n-1}\gamma_i\gamma_{n-i}C_{n-1}^i - \left(\frac{\sigma_v}{\kappa}\right)^2\lambda_1 J_n\,, \quad n \geq 3$$

$$\Gamma_0 = \frac{\exp(-\kappa\Delta)}{\kappa}\,, \quad \Gamma_n = -\kappa\Delta\sum_{i=0}^{n-1}C_{n-1}^i\gamma_{n-i}\Gamma_i\,, \quad n \geq 1\,.$$

Next, we derive $A_{y,v}^{(i,j)}$ for $i \geq 0$ and $j \geq 0$. Again, with "initial" values of $A_{y,v}^{(0,1)} = \kappa\bar{v}f_0$ and $A_{y,v}^{(1,0)} = (-\lambda_0\mu^* + \lambda_0 J_1)\Delta - (\kappa\gamma_1 + \sigma_v\rho)\,(\Delta - f_0)\,\kappa\bar{v}/\sigma_v^2$, the following formulas enable us to calculate $A_{y,v}^{(i,j)}$ recursively up to any order. We have

$$A_{y,v}^{(0,n)} = \frac{n-1}{2}\sigma_v^2 f_0 A_{y,v}^{(0,n-1)}\,, \quad n \geq 2\,,$$

$$A_{y,v}^{(n,0)} = \lambda_0 J_n^0\Delta - \frac{\kappa\bar{v}}{\sigma_v^2}\left(\kappa\gamma_n\Delta + \hat{f}_n - \hat{g}_n\right)\,, \quad n \geq 2\,,$$

$$A_{y,v}^{(n,1)} = -\frac{\kappa\bar{v}}{2}f_0 f_n - \frac{1}{2}\sum_{i=1}^{n-1}C_n^i f_i A_{y,v}^{(n-i,1)}\,, \quad n \geq 1\,,$$

$$A_{y,v}^{(n,m)} = -\kappa\bar{v}\,\sigma_v^{2(m-1)}m!f_n\left(\frac{f_0}{2}\right)^m - \frac{1}{2}\sum_{i=1}^{n-1}\hat{C}_n^i(m)f_i A_{y,v}^{(n-i,m)}\,, \quad n \geq 1\,, m \geq 2\,,$$

where for $n \geq 1$, $\hat{C}_n^0(m) = m$, $\hat{C}_n^n(m) = 1$, and, for $0 < i < n$, $C_n^i(m) = C_{n-1}^i(m) + C_{n-1}^{i-1}(m)$. (Notice that, $C_n^i = n!/i!(n-i)!$ defined previously, is a special case of $C_n^i(m)$, with $m = 1$.) The coefficients $\hat{f}$ and $\hat{g}$ are defined by

$$\hat{f}_1 = f_1\,, \quad \hat{f}_n = f_n - \frac{1}{2}\sum_{i=1}^{n-1}C_{n-1}^{n-i}\hat{f}_i f_{n-i}\,, \quad \hat{g}_1 = g_1\,, \quad \hat{g}_n = g_n - \frac{1}{2}\sum_{i=1}^{n-1}C_{n-1}^{n-i}\hat{g}_i g_{n-i}\,.$$

G Tests of Moment Conditions

This appendix is closely related to the test of over-identifying restrictions developed by Hansen [1982] (in particular, Lemma 4.1 of Hansen [1982]). Let $E_n(\epsilon_{n+1}) = 0$ be the $m = 7$

moment conditions under consideration, and let $\hat{\vartheta}_N$ be the exactly-identified IS-GMM estimators, obtained from the "optimal" moment condition $E_n(\mathcal{H}_{n+1}) = 0$. To test $E_n(\epsilon_{n+1}) = 0$ we construct its sample analogue by

$$\mathcal{G}_N(\hat{\vartheta}_N) = \frac{1}{N} \sum_{n \leq N} \epsilon_n(\hat{\vartheta}_N), \tag{G.1}$$

where $\epsilon_n(\hat{\vartheta}_N)$ denotes evaluating the moments ϵ at the IS-GMM estimator $\hat{\vartheta}_N$. Using arguments similar to those following Assumption 3.5 in Section 3, one can show that, under typical technical regularity conditions, $\sqrt{N}\, \mathcal{G}_N(\vartheta_0)$ is asymptotically normal. Applying a standard mean-value expansion,

$$\mathcal{G}_N(\hat{\vartheta}_N) = \mathcal{G}_N(\vartheta_0) + \left.\frac{\partial \mathcal{G}_N(\vartheta)}{\partial \vartheta}\right|_{\bar{\vartheta}_N} \left(\hat{\vartheta}_N - \vartheta_0\right), \tag{G.2}$$

where $\bar{\vartheta}_N^j$ is can be shown between ϑ_0^j and ϑ_N^j, for $j \in \{1, \ldots, n_\vartheta\}$. Moreover, for sufficiently large N and with probability arbitrarily close to one, we can write

$$\hat{\vartheta}_N - \vartheta_0 = -\left(\left.\frac{\partial G_N(\vartheta)}{\partial \vartheta}\right|_{\bar{\vartheta}_N}\right)^{-1} G_N(\vartheta_0), \tag{G.3}$$

where $G_N = (N)^{-1} \sum_n \mathcal{H}_n$ is the sample analogue of the "optimal" moments. We know that $\partial G_N(\bar{\vartheta}_N)/\partial\vartheta$ converges to a constant full-rank matrix d_0 in probability, under Assumption 3.6, using the fact that $\hat{\vartheta}_N$ is estimated under an exactly identified IS-GMM setting.

Substituting (G.3) into (G.2), we obtain

$$\sqrt{N}\mathcal{G}_N(\hat{\vartheta}_N) \overset{a}{\approx} \sqrt{N} \left(\mathcal{G}_N(\vartheta_0) - \left.\frac{\partial \mathcal{G}_N(\vartheta)}{\partial \vartheta}\right|_{\bar{\vartheta}_N} \left(\left.\frac{\partial G_N(\vartheta)}{\partial \vartheta}\right|_{\bar{\vartheta}_N}\right)^{-1} G_N(\vartheta_0) \right), \tag{G.4}$$

where $\overset{a}{\approx}$ means "asymptotically equivalent in distribution to." Thus, $\mathcal{G}_N(\hat{\vartheta}_N)$ is asymptotically normal with some covariance matrix Ω. An estimator Ω_N of Ω can be obtained by estimating the covariance matrix of the right hand side of (G.4).

The m moment conditions can be tested either individually or jointly. We can test the i-th moment condition by using the fact that $\sqrt{N}\mathcal{G}_N^i(\hat{\vartheta}_N)/\sqrt{(\Omega_N)_{ii}}$ is asymptotically standard normal. We can test any subgroup of moment conditions, indexed by I, by using the fact that, in large sample, $N(\mathcal{G}_N^I(\hat{\vartheta}_N))^\top((\Omega_N)_I)^{-1}\mathcal{G}_N^I(\hat{\vartheta}_N)$ is distributed as a χ^2 random variable with $\#(I)$ degrees of freedom.

H Proof of Proposition 3.1

Fixing some $\vartheta \in \Theta$,

$$\begin{aligned}\lim_N G_N(\vartheta) &= \lim \frac{1}{N} \sum_{n \leq N} H(X_n, \vartheta, Y_n) = \lim_N \frac{1}{N} \sum_i \sum_{n \in A_N^{(i)}} H(X_n, \vartheta, Y_n) \\ &= \lim_N \sum_i \frac{\#A_N^{(i)}}{N} G_N^{(i)}(\vartheta) \sum_i \lim_N \frac{\#A_N^{(i)}}{N} \lim_N G_N^{(i)}(\vartheta) = \sum_i w_i G_\infty^{(i)}(\vartheta).\end{aligned}$$

Letting $G_\infty(\vartheta) = \sum_i w_i G_\infty^{(i)}(\vartheta)$, we have shown that $G_N(\vartheta)$ converges a.s. to $G_\infty(\vartheta)$ as $N \to \infty$. Next, we show that this convergence is uniform in ϑ. We have

$$\begin{aligned}
\sup_\vartheta |G_N(\vartheta) - G_\infty(\vartheta)| &= \sup_\vartheta \left| \sum_i \frac{\# A_N^{(i)}}{N} G_N^{(i)}(\vartheta) - \sum_i w_i G_\infty^{(i)}(\vartheta) \right| \\
&\le \sum_i \sup_\vartheta \left| \frac{\# A_N^{(i)}}{N} G_N^{(i)}(\vartheta) - w_i G_\infty^{(i)}(\vartheta) \right| \\
&\le \sum_i \sup_\vartheta \left(\left| \frac{\# A^{(i)}}{N} G_N^{(i)}(\vartheta) - w_i G_N^{(i)}(\vartheta) \right| + |w_i G_N^{(i)}(\vartheta) - w_i G_\infty^{(i)}(\vartheta)| \right) \\
&\le \sum_i \left(\sup_\vartheta |G_N^{(i)}(\vartheta)| \left| \frac{A_N^{(i)}}{N} - w_i \right| + \sup_\vartheta w_i \, |G_N^{(i)}(\vartheta) - G_\infty^{(i)}(\vartheta)| \right)
\end{aligned}$$

which converges almost surely to zero, by (3.10) and (3.11).

I Geometric Ergodicity of (y, V, r, q)

In this appendix, we show that $\{y_n, V_n, r_n, q_n\}$ is geometrically ergodic. We first show that any square-root process satisfying the Feller condition is geometrically ergodic. This result applies to V, r, and q separately. Next, we show that $\{y_n, V_n\}$ is geometrically ergodic under certain parameter restrictions. As r and q are independent, and independent of $\{y_n, V_n\}$, steps one and two therefore establish the geometric ergodicity of (y, V, r, q).

The *geometric drift condition* of Meyn and Tweedie [1993] plays a central role in establishing geometric ergodicity in both steps one and two. Letting $\mathbb{X}$ denote the measurable state space of a Markov chain whose one-step transition probability is $P(x, \mathcal{A})$, (defined for any $x \in \mathbb{X}$ and any $\mathcal{A}$ in the σ-field of subsets of $\mathbb{X}$) the geometric drift condition is stated as following.

Geometric Drift Towards C (Meyn and Tweedie [1993] page 367) *There exists an extended-real valued function $f : \mathbb{X} \to [1, \infty]$, a measurable set C, and constants $\beta > 0$ and $b < \infty$ such that*

$$\Delta f(x) \le -\beta f(x) + b \mathbf{1}_C(x) \,, \quad x \in \mathbb{X} \,, \tag{I.1}$$

where $\Delta f(x) = \int P(x, dy) f(y) - f(x)$ denotes the "drift operator."

Intuitively, (I.1) requires that, when the process wanders far out into the state space, it is pulled back fast enough.

Proposition I.1 (Square-Root Process) *Let X be a square-root process defined by*

$$dX_t = \kappa(\bar{x} - X_t)\, dt + \sigma \sqrt{X_t}\, dW_t \,,$$

where κ, $\bar{x}$, and σ are positive constants, and W is a standard Brownian motion. Suppose that X satisfies the Feller condition $\kappa\theta > \sigma^2/2$. Then the Markov chain $\{X_n\}$ sampled from X with an arbitrary discrete time interval Δt is geometrically ergodic.

Proof Letting $f(x) = 1 + x$ and $\beta = (1 - \exp(-\kappa\Delta t))/2 > 0$, we have

$$\begin{aligned}\Delta f(x) + \beta f(x) &= E\left(X_{n+1} | X_n = x\right) - x + \beta\,(1+x) \\ &= x\exp\left(-\kappa\Delta t\right) + \bar{x}\,\left(1 - \exp\left(-\kappa\Delta t\right)\right) - x + \beta\,(1+x) \\ &= (2\bar{x}+1)\beta - x\beta\,.\end{aligned}$$

Letting $C = (0\,, 2\bar{x}+1]$ and $b = (2\bar{x}+1)\beta$, we have (I.1). Using Theorem 15.0.1 of Meyn and Tweedie [1993], the geometric ergodicity of $\{X_n\}$ follows from the fact that $\{X_n\}$ is irreducible and aperiodic, and C is a compact set (thus a petite set for this chain).

Proposition I.2 (The Joint Process of Return and Volatility) *For a fixed time interval $\Delta t > 0$, let $\{y_n\}$ be the Δt-period return defined in (3.1). Suppose that the volatility process V satisfies the Feller condition, and, moreover, that*

$$\kappa^2\,\frac{1+\exp(-\kappa\Delta t)}{1-\exp\left(-\kappa\Delta t\right)} - \left(\eta^s - \frac{1}{2} + \lambda\left(\mu_J - \mu^*\right)\right)^2 > 0\,. \tag{I.2}$$

Then the bi-variate Markov chain $\{y_n, V_n\}$ is geometrically ergodic.

Proof Letting $f(y,v) = y^2 + v^2 + 1$, we have

$$\Delta f(y_n, V_n) = E\left(y_{n+1}^2 \,|\, V_n\right) - y_n^2 + E\left(V_{n+1}^2 \,|\, V_n\right) - V_n^2\,,$$

where, using the results in Appendix F,

$$\begin{aligned}E\left(y_{n+1}^2 \,|\, V_n\right) &= A_y^{(2)} + B_y^{(2)}\,V_n + \left(A_y^{(1)} + B_y^{(1)}\,V_n\right)^2 \\ E\left(V_{n+1}^2 \,|\, V_n\right) &= A_v^{(2)} + B_v^{(2)}\,V_n + \left(A_v^{(1)} + B_v^{(1)}\,V_n\right)^2\,,\end{aligned}$$

where $B_v^{(1)} = \exp(-\kappa\Delta)$ and $B_y^{(1)} = \left(\eta^s - \frac{1}{2} + \lambda(\mu_J - \mu^*)\right)\left(1 - \exp(-\kappa\Delta t)\right)/\kappa$. Letting $\delta = \kappa^2\left(1+\exp(-\kappa\Delta t)\right)/(1-\exp\left(-\kappa\Delta t\right)) - \left(\eta^s - \frac{1}{2} + \lambda\left(\mu_J - \mu^*\right)\right)^2$, we further define

$$\begin{aligned}a_2 &= 1 - \left(B_y^{(1)}\right)^2 - \left(B_v^{(1)}\right)^2 = \left(\frac{1-\exp\left(-\kappa\Delta t\right)}{\kappa}\right)^2\delta\,, \\ a_1 &= \frac{1}{2}\left(B_y^{(2)} + B_v^{(2)}\right) + A_y^{(1)}\,B_y^{(1)} + A_v^{(1)}\,B_v^{(1)}\,, \quad a_0 = A_y^{(2)} + A_v^{(2)} + \left(A_y^{(1)}\right)^2 + \left(A_v^{(1)}\right)^2\,,\end{aligned}$$

where $a_2 > 0$ follows from (I.2). Letting $\beta = \min(a_2/2, 1/2)$, so that $0 < \beta \le 1/2$ and $\beta \le a_2/2$, we have

$$\begin{aligned}\Delta f(y,v) + \beta f(y,v) &= -(a_2 - \beta)v^2 + 2a_1 v + a_0 - (1-\beta)\,y^2 + \beta \\ &< -\beta v^2 + 2a_1 v + a_0 \\ &\le -\beta(v + a_1)^2 + \beta a_1^2 + a_0\,.\end{aligned}$$

It is easy to check[41] that $a_0 > 0$, we therefore have $\beta a_1^2 + a_0 > 0$. Letting

$$C = \{0\} \times \left(0, \sqrt{a_1^2 + \frac{a_0}{\beta}} - a_1\right]$$

[41] In order for $A_y^{(2)} + B_y^{(2)} V_n = \mathrm{Var}(y_{n+1}|V_n) \ge 0$ to hold almost surely, it must be that $A_y^{(2)} \ge 0$. Moreover $A_v^{(1)} > 0$ and $A_v^{(2)} > 0$.

and $b = \beta a_1^2 + a_0$, we see that $\Delta f(y,v) + \beta f(y,v) \leq b$ for $(y,v) \in C$, and $\Delta f(y,v) + \beta f(y,v) \leq 0$ for $y \neq 0$ (trivially) and $v > \sqrt{a_1^2 + a_0/\beta} - a_1$. We have therefore shown that (I.1) holds for $\{y_n, V_n\}$. Using Theorem 15.0.1 of Meyn and Tweedie [1993], the geometric ergodicity of $\{y_n, V_n\}$ follows from the fact that it is irreducible and aperiodic, and C is a compact set (thus a petite set for this chain).

References

Aït-Sahalia, Y., Y. Wang, and F. Yared (1998). Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset? Working Paper, Graduate School of Business, University of Chicago.

Andersen, T., L. Benzoni, and J. Lund (1998). Estimating Jump-Diffusions for Equity Returns. Working Paper, Kellogg Graduate School of Management, Northwestern University.

Andrews, D. W. K. (1987). Consistency in Nonlinear Econometric Models: A Generic Uniform Law of Large Numbers. *Econometrica* **55**, 1465–1471.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical Performance of Alternative Option Pricing Models. *Journal of Finance* **52**, 2003–49.

Bakshi, G. and D. Madan (1999). Spanning and Derivative-Security Valuation. *Journal of Financial Economics, forthcoming*.

Bates, D. (1997). Post-'87 Crash fears in S&P 500 Futures Options. Working Paper 5894, National Bureau of Economic Research.

Benzoni, L. (1998). Pricing Options under Stochastic Volatility: An Econnometric Analysis. Working Paper, J.L. Kellog Graduate School of Management, Northwestern University.

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81**, 637–654.

Bollerslev, T. and H. O. Mikkelsen (1996). Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics* **73**, 151–184.

Brandt, M. and P. Santa-Clara (1999). Simulated Likelihood Estimation of Multivariate Diffusions with an Application to Interest Rates and Exchange Rates with Stochastic Volatility. Working Paper, The Wharton School, University of Pennsylvania, and The Anderson School, University of California, Los Angeles.

Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics*. New York: Springer-Verlag.

Campbell, J. Y. and L. Hentschel (1992). No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics* **31**, 281–318.

Chernov, M. and E. Ghysels (1999). A Study towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation. *Journal of Financial Economics, forthcoming*.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53*(2), 385–407.

Dai, Q. and K. Singleton (1997). Specification Analysis of Affine Term Structure Models. Working Paper, Graduate School of Business, Stanford Unversity.

Das, S. R. and R. K. Sundaram (1999). Of Smiles and Smirks: A Term Structure Perspective. *Journal of Financial and Quantitative Analysis* **34**, 211–240.

David, A. and P. Veronesi (1999). Option Prices with Uncertain Fundamentals, Theory and Evidence on the Dynamics of Implied Volatilities. Working Paper, Graduate School of Business, University of Chicago.

Davies, R. (1973). Numerical Inversion of a Characteristic Function. *Biometrika* **60**, 415–417.

Derman, E. (1999, January). Regimes of Volatility. *Goldman Sachs Quantitative Strategies Research Notes*.

Detemple, J. and L. Selden (1991). A General Equilibrium Analysis of Option and Stock Market Interactions. *International Economic Review* **32**, 279–303.

Doob, J. (1953). *Stochastic Processes*. New York: John Wiley and Sons.

Duffee, G. (1999). Estimating the Price of Default Risk. *Review of Financial Studies* **12**, 197–226.

Duffie, D. (1996). *Dynamic Asset Pricing Theory* (second edition). Princeton University Press.

Duffie, D., J. Pan, and K. Singleton (1999). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica, forthcoming*.

Duffie, D. and K. Singleton (1993). Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* **61**, 929–952.

Duffie, D. and K. Singleton (1999). Modeling Term Structure of Defaultable Bonds. *Review of Financial Studies* **12**, 687–720.

Eichenbaum, M. S., L. P. Hansen, and K. J. Singleton (1988). A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertanity. *Quarterly Journal of Economics* **103**, 51–78.

Eraker, B., M. S. Johannes, and N. G. Polson (1999). Asset Return Dynamics with Jumps, Stochastic Volatility and Jumps to Volatility. Working Paper, University of Chicago.

Feller, W. (1951). Two Singular Diffusion Problems. *Annals of Mathematics* **54**, 173–182.

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.

Gallant, A. R. and H. White (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell.

Glynn, P. (1999). “Written Communication,” the Department of Engineering Economic Systems-Operations Research, Stanford University.

Glynn, P. and K. Sigman (1998). Independent Sampling of a Stochastic Process. *Stochastic Processes and their Applications* **74**, 151–164.

Gul, F. (1991). A Theory of Disappointment Aversion. *Econometrica* **59**, 667–686.

Guo, D. (1998). The Risk Premium of Volatility Implicit in Currentcy Options. *Journal of Business and Economics Statistics* **16**, 498–507.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.

Hansen, L., J. Heaton, and M. Ogaki (1988). Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions. *Journal of the American Statistical Association* **83**, 863–871.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* **50**, 1029–1054.

Hansen, L. P. (1985). A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators. *Journal of Econometrics* **30**, 203–238.

Harrison, M. and D. Kreps (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* **20**, 381–408.

Harrison, M. and S. Pliska (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and Their Applications* **11**, 215–260.

Heston, S. (1993). A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* **6**, 327–343.

Hong, H. and J. Stein (1999). Differences of Opinion, Rational Arbitrage, and Market Crashes. Working Paper, Graduate School of Business, Stanford University.

Huang, C. (1985). Information Structures and Viable Price Systems. *Journal of Mathematical Economics* **14**, 215–240.

Hull, J. and A. White (1987). The pricing of options on Assets with Stochastic Volatilities. *Journal of Finance* **42**, 281–300.

Jackwerth, J. C. and M. Rubinstein (1996). Recovering Probability Distributions from Option Prices. *Journal of Finance* **51**, 1611–1631.

Johannes, M., R. Kumar, and N. Polson (1998). Jump Diffusion Models: Extracting Jumps from Equity Indices. Working Paper, Graduate School of Business, University of Chicago.

Jones, C. (1999). The Dynamics of Stochastic Volatility. Working Paper, Simon School of Business, University of Rochester.

Kapadia, N. (1998). Do Equity Options Price Volatility Risk? Empirical Evidence. Working Paper, University of Massachusetts, Amherst.

Karatzas, I. and S. Shreve (1991). *Brownian Motion and Stochastic Calculus* (second edition). Springer-Verlag New York.

Liu, J. (1997). Generalized Method of Moments Estimation of Affine Diffusion Processes. Working Paper, Graduate School of Business, Stanford University.

Lucas, R. E. (1978). Asset Prices in an Exchange Economy. *Econometrica* **46**, 1429–1445.

Mehra, R. and E. Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics* **15**, 145–161.

Merton, R. (1976). Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* **3**, 125–44.

Meyn, S. and R. Tweedie (1993). *Markov Chains and Stochastic Stability*. New York: Springer-Verlag.

Naik, V. and M. Lee (1990). General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns. *Review of Financial Studies* **3**, 493–521.

Novikov, A. (1972). On an Identity for Stochastic Integrals. *Theory of Probability and Its Applications* **17**, 538–541.

Pastrorello, S., E. Renault, and N. Touzi (1996). Statistical Inference for Random Variance Option Pricing. Working Paper, Dipartimento di Scienze Economiche, Università di Bologna, Italy.

Pham, H. and N. Touzi (1996). Equilibrium State Prices in a Stochastic Volatility Model. *Mathematical Finance* **6**, 215–236.

Piazzesi, M. (1999). An Econometric Model of the Yield Curve with Macroeconomic Jump Effects. Working Paper, Stanford University.

Poteshman, A. M. (1998). Estimating a Genral Stochastic Variance Model from Options Prices. Working Paper, Graduate School of Business, University of Chicago.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

Protter, P. (1999). A Partial Introduction to Finance. In *Fields Institute Symposium on Numerical Stochastics in Finance*.

Renault, E. and N. Touzi (1996). Option Hedging and Implied Volatilities in a Stochastic Volatility Model. *Mathematical Finance* **6**, 279–302.

Revuz, D. and M. Yor (1991). *Continuous Martingales and Brownian Motion*. New York: Springer.

Rietz, T. (1988). The Eequity Risk Premium: A Solution. *Journal of Monetary Economics* **22**, 117–131.

Rubinstein, M. (1994). Implied Binomial Trees. *Journal of Finance* **49**, 771–818.

Singleton, K. (1999a). Estimation of Affine Diffusion Models Based on the Empirical Characteristic Function. Working Paper, Graduate School of Business, Stanford University.

Singleton, K. (1999b). *Lecture Notes on Econometric Analysis of Dynamic Asset Pricing Models*. Graduate School of Business, Stanford University.

Stein, E. and J. Stein (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies* **4**, 725–52.

Stein, J. (1989). Overreactions in the Options Market. *Journal of Finance* **44**, 1011–1023.

Wu, G. (1998). The Determinants of Asymmetric Volatility. Working Paper, University of Michigan Business School.