

Integrated Time-Series Analysis of Spot and Option Prices

Jun Pan*

MIT Sloan School of Management

First draft : August 20, 1999

This draft : August 8, 2000

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.

*MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142. Email: junpan@mit.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 comments from Andrew Ang, Joe Chen, Mark Ferguson, Peter Glynn, Harrison Hong, Ming Huang, George Papanicolaou, Monika Piazzesi, Tom Sargent, and seminar participants at Stanford, Berkeley, MIT, UCLA, HBS, Rochester, Wisconsin Madison, Michigan, Duke, Kellogg, Chicago, and Columbia. This paper can be downloaded at <http://www.mit.edu/~junpan/>.

1 Introduction

Empirical evidence in support of stochastic volatility and jumps is documented by a number of studies. For example, both risk factors are shown to be important in explaining the cross-sectional behavior of option prices (Bakshi, Cao, and Chen [1997], Bates [2000]), as well as in characterizing the time-series behavior of the underlying stock prices (Jorion [1989], Andersen, Benzoni, and Lund [1998], Chernov, Gallant, Ghysels, and Tauchen [1999]).

To what extent are such risk factors priced? This question (left unanswered in these studies) is important as at the heart of investors' decision making is how such risk factors are priced in the market, and central to modern finance is how investors react to different types of uncertainty.

Motivated by this question, this paper is an empirical investigation of how volatility and jump risks are priced in S&P 500 index options. It contributes to the existing literature by providing compelling evidence of a jump-risk premium that is highly correlated with the market volatility. We show that this jump-risk premium is important not only in reconciling the dynamics implied by the joint time series of spot and option prices, but also in explaining changes over time of the "smiles" and "smirks" found in cross-sectional options data.

As an illustration on how we uncover risk premia from the joint time series of spot and option prices, we offer a simple example. We allow the volatility of index returns to be stochastic (via the Heston [1993] model), but assume that volatility risks are not priced. Fitting this model to the bi-variate time series of the S&P 500 index and near-the-money short-dated option prices, we find significant inconsistency between the level of volatility observed in the spot market and that implied, through the model, by the options market. In particular, option-implied volatility is too high to be rationalized by the ex-post realized volatility observed in the spot market.

While such an "upward bias" of option prices can potentially be explained by a number of reasons,¹ we focus our attention on the role of risk premium. In essence, this paper is about finding the right type of risk premium to explain the joint time-series behavior of spot and option prices. For this, our estimation strategy is that of an integrated treatment of the joint time series of spot and option prices, and our framework is that of a dynamic arbitrage-free option pricing model. We adopt the Bates [2000] model to characterize the underlying price dynamics, allowing for both stochastic volatility (nesting the Heston model) and jumps in returns. Our focus is on how these two types of risk factors are priced in the options market. For this, we introduce a parametric pricing kernel to price all of the risk factors, including the volatility risk and the jump risk. One important feature of the jump-risk premium considered in this paper is that it is allowed to depend on the market volatility.

Evidence in support of volatility-risk premium is reported under pure-diffusion settings by Guo [1998], Benzoni [1998], Chernov and Ghysels [2000], Poteshman [1998] and Kapadia [1998]. The role of jump-risk premium and the relatively importance of premia for jump and

¹For example, observing this upward bias under the setting of the Black-Scholes model, Jackwerth and Rubinstein [1996] point out that it is perhaps a combination of flaws in the Black-Scholes model coupled with the market's correction for the fact that extreme rare events are possible, but not usually present. Longstaff [1995] points out the possibility of market friction in explaining this phenomenon. Along with liquidity stories, the specific character of the market making in the options market can also be a factor.

volatility risks are largely unknown.

The integrated time-series analysis performed in this paper offers some new empirical evidence. First, allowing volatility-risk premium in the Heston model does partially reconcile the joint dynamics implied by spot and option prices, but overall the model is still rejected by the joint data. Moreover, the volatility-risk premium thus estimated from the joint time-series data implies an explosive volatility process under the “risk-neutral” measure, leading to severely over-priced long-dated options. Second, introducing jump and jump-risk premium to the Heston model greatly improves the fit to the joint time-series data. In particular, the jump-risk premium is capable of reconciling the joint time-series behavior of spot and short-dated option prices without any distortion of long-dated option prices. One important feature of this jump-risk premium is that it is highly correlated with the market volatility. Finally, when we allow both types of risk premia simultaneously to reconcile the spot and option dynamics, we find that the jump-risk premium by far dominates the volatility-risk premium.

The intuition behind our empirical results is quite simple. Given its diffusive nature, the uncertainty associated with volatility risk diminishes as one shrinks the time horizon of concern. Jump risk, on the other hand, has a quite different impact on investors. While the jump-timing uncertainty decreases with time horizon, the jump-size uncertainty does not. As a result, relative to the risk premium associated with the diffusive volatility risk, the premium for jump-size uncertainty could be quite pronounced for short-dated options. Our empirical results clearly indicate the importance of this risk premium component in short-dated options that are close to the money.

In documenting the importance of the jump-risk premium, we rely almost exclusively on the joint time series of one index and one option (that is near the money with an average maturity of one month), leaving out the cross sections of options with different moneyness and maturities. One important question for us to ask is how well our model, with parameters estimated exclusively from the joint time series of one spot and one option, explains the cross sections of option prices. This is complementary to the studies of Bakshi, Cao, and Chen [1997] and Bates [2000], who focus mainly on uncovering the risk-neutral dynamics using cross-sectional option prices, while leaving out the time-series information embedded in the spot and options dynamics. Our cross-sectional investigation shows that the Bates [2000] model, estimated using the joint time series of one spot and one option, can explain quite well the changes over time of the “smirk” patterns. Using model estimates of both the actual and risk-neutral dynamics, we point out that such pronounced “smirk” patterns are largely due to investors’ aversion to jump risks. Moreover, given that we use near-the-money option prices in model estimation, our results show that investors’ fear for such jump risk is reflected not only in deep out-of-the-money puts, but also in near-the-money options. Our results also indicate that the one-factor stochastic volatility model implies a term structure of volatility that dies out too fast to be consistent with the data. Consequently, the Bates model is found to overprice (underprice) long-dated options on days of low (high) volatilities.

For the purpose of understanding the pricing kernel that links the two markets, the spot and options data can be exploited in alternative ways.² Under a consumption-based asset

²The linkage between spot and option prices is exploited in a somewhat different setting by Cochrane and Saá-Requejo [2000] and Bernardo and Ledoit [2000]. Taking the underlying price distribution as given,

pricing setting, Jackwerth [1999] and Ait-Sahalia and Lo [2000] uncover the risk-aversion coefficient for a representative agent by comparing the risk-neutral distribution implied by index options with the actual distribution estimated from the time series of index. In the same spirit, Rosenberg and Engle [1999] estimate an empirical pricing kernel. In this paper, we rely instead on an arbitrage-free asset pricing model with a parametric pricing kernel that prices diffusive return shocks, volatility shocks, and jump risks. Different risk factors would have different impact on option prices, and the compensation for their uncertainty could also be reflected in very different ways through option pricing. By allowing investors to have different risk attitudes toward the three types of risk factors, our parametric approach to the pricing kernel allows us to investigate such differences. Indeed, our empirical results indicate that investors might have quite different risk attitudes toward jump risk and diffusive risk.

In reality, apart from the pricing-kernel story, there could very well be other factors contributing to the option prices that we observe in the market. For example, Longstaff [1995] points out the possibility of market frictions. Jackwerth and Rubinstein [1996] raise the possibility of a “peso” component in option prices that reflects extreme events, which are possible but are rarely present in the underlying market. This “peso” explanation is pursued in Ait-Sahalia, Wang, and Yared [1998]. Comparing the option-implied state-price density (SPD) with the index-implied SPD, they report that the option-implied SPD is more negatively skewed than the index-implied SPD. In order to reconcile this difference, they introduce Poisson jumps with negative and deterministic jump sizes. Such jumps are allowed to be present in estimating the option-implied SPD, but are assumed to be absent in estimating the index-implied SPD, therefore capturing the spirit of “peso” events. It should be noted that one cannot interpret this result as evidence of a jump-risk premium, as by definition the effect of a risk premium shows up in both SPDs with equal footing, and one cannot use a risk premium story to reconcile the difference between the two SPDs.

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 outlines the integrated approach adopted in this paper and details on the estimation strategy. Section 4 describes the data, Section 5 summarizes the empirical findings, and Section 6 concludes the paper. Technical details are provided in appendices.

2 The Model

We adopt the Bates [2000] model to characterize the dynamics of the underlying spot prices. For completeness of the paper, a brief description of the Bates model is given in Section 2.1. This model introduces three sources of uncertainty to the underlying price movement: diffusive return shocks, volatility shocks, and jump risk. In order to price all three types of uncertainty, we introduce in Section 2.2 a parametric pricing kernel. Equivalently, one could specify the risk-neutral dynamics, which is given in Section 2.3. Option pricing based on this class of risk-neutral dynamics is well-established and involves Levy inversion of the

they look for option-pricing bounds by imposing different restrictions on the pricing kernel.

characteristic function of the state variables.³ A brief description is given in Section 2.3 for completeness, and a new numerical scheme for the Levy inversion is introduced in Appendix E.

2.1 The Data-Generating Process

We fix a probability space (Ω, \mathcal{F}, P) and an information filtration (\mathcal{F}_t) satisfying the usual conditions (see, for example, Protter [1990]), 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

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

where $W = [W^{(1)}, W^{(2)}]^\top$ is an adapted standard Brownian motion in \mathbb{R}^2 , $\rho \in (-1, 1)$ 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 , 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

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

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

³See, for example, Stein and Stein [1991], Heston [1993], Scott [1997], Bates [2000], Bakshi, Cao, and Chen [1997], Bakshi and Madan [2000], Duffie, Pan, and Singleton [1999].

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 [2000], which in turn extends the stochastic-volatility model (SV) of Heston [1993] (alternative stochastic volatility models are Hull and White [1987] and Stein and Stein [1991]). Compared with the Merton [1976] jump extension (see also Cox and Ross [1976]) 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 depends on the stochastic volatility.

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 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 [1996] and references therein.) In our setting of random jump size, markets are incomplete with respect to the underlying and the finite number of options contracts, and the state-price density is not unique. Our approach⁵ 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. The interest-rate and dividend-rate risks, however, are assumed to be not priced, given their mild effects on short-dated option prices.

Consider a candidate state-price density π of the form

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

where $\mathcal{E}(\cdot)$ denotes the stochastic exponential,⁶ and where ζ and U_i^π are defined as follows.

The Brownian shocks in price and volatility are priced using their respective market

⁴In 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 .

⁵An 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.

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

prices of risk ζ , defined by

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

where η^s and η^v are constant coefficients. For this specification of market price of risk, the time- t instantaneous risk premium associated with the diffusive price shock is $\eta^s V_t$, while that associated with the volatility shock is $\eta^v V_t$.

The jump risks are priced by the jump component in the pricing kernel — whenever the underlying price jumps, the pricing kernel also jumps. The jump sizes U_i^π are assumed to be *i.i.d.* normal with mean μ_π and variance σ_π^2 , and are assumed to be independent of W , W^r , W^q , and inter-jump times. The random jump sizes U_i^π and U_i^s are allowed to be correlated (with constant ρ_π), but are assumed to be independent across different jump times. 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, $\mu_\pi + \sigma_\pi^2/2 = 0$. As we will see in the next section, this constraint is translated to a zero jump-timing risk premium. (Similarly, if we were to turn off the correlation between U_i^π and U_i^s by letting $\rho_\pi = 0$, the jump-size risk premium would be zero.)

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 to be local martingales.⁷ To see that \mathcal{S}^π and \mathcal{B}^π are indeed local martingales, we apply Ito's Formula,

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

where \mathcal{N}_t is the number of price jumps by time t , and where $\mu^* = \exp(\mu_J + \sigma_J \sigma_\pi \rho_\pi + \sigma_J^2/2) - 1$. We see that \mathcal{S}^π and \mathcal{B}^π 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

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

⁷Appendix B shows that this indeed rules out arbitrage opportunities involving \mathcal{S} and \mathcal{B} , under natural conditions on dynamic trading strategies.

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 $E_t(\pi_T(S_T - K)^+)$ 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.” (Cox and Ross [1976], Harrison and Kreps [1979].) For this, we define a density process $\xi_t = \pi_t \exp\left(\int_0^t r_s ds\right)$. Applying Ito’s Formula, one can show that ξ is a local martingale. If ξ is actually a martingale,⁸ 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

$$\begin{aligned} dS_t &= [r_t - q_t - (\lambda_0 + \lambda_1 V_t)\mu^*] S_t dt + \sqrt{V_t} S_t dW_t^{(1)}(Q) + dZ_t^Q, \\ dV_t &= [\kappa_v(\bar{v} - V_t) + \eta^v V_t] dt + \sigma_v \sqrt{V_t} \left(\rho dW_t^{(1)}(Q) + \sqrt{1 - \rho^2} dW_t^{(2)}(Q) \right), \end{aligned} \quad (2.5)$$

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

$$W_t(Q) = W_t + \int_0^t \zeta_s ds, \quad 0 \leq t \leq T. \quad (2.6)$$

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 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 $\mu^* = E^Q(\exp(U^s) - 1) = \exp(\mu_j^* + \sigma_j^2/2) - 1$.

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 supposing⁹ 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

⁸Appendix C gives a sufficient Novikov-like condition on the model parameters, for ξ to be a martingale.

⁹This 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 $\lambda_0^* = \lambda_0 \exp(\mu_\pi + \sigma_\pi^2/2)$ and $\lambda_1^* = \lambda_1 \exp(\mu_\pi + \sigma_\pi^2/2)$. 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 .

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^s . 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 (2.5). 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 $\eta_0 + \eta_1 V_t + \eta_2 V_t^2 + \dots + \eta_l V_t^l$, for some constant coefficients $\eta_0, \eta_1, \eta_2, \dots, \eta_l$. Our specification, however, rules out the possibility that $\eta_0 \neq 0$, as it could imply non-diminishing risk premia as the volatility approaches to zero. 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

$$\vartheta = [\kappa_v, \bar{v}, \sigma_v, \rho, \eta^s, \eta^v, \lambda_0, \lambda_1, \mu, \sigma_J, \mu^*]^\top, \quad (2.7)$$

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 $n_\vartheta = 11$.

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

$$C_t = \frac{1}{\pi_t} E_t [\pi_T (S_T - K)^+] = E_t^Q \left[\exp \left(- \int_t^T r_u du \right) (S_T - K)^+ \right]. \quad (2.8)$$

In order to calculate the expectation in (2.8), we adopt a transform-based approach. (See, for example, Stein and Stein [1991], Heston [1993], Scott [1997], Bates [2000], Bakshi, Cao, and Chen [1997], Bakshi and Madan [2000], 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]),

$$\psi^\vartheta(c, v, r, q, \tau) = \exp \left(\alpha(c, \tau, \vartheta, \theta_r, \theta_q) + \beta_v(c, \tau, \vartheta) v + \beta_r(c, \tau, \theta_r) r + \beta_q(c, \tau, \theta_q) q \right), \quad (2.9)$$

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

$$f(v, \vartheta, r, q, \tau, k) = \mathcal{P}_1 - k \mathcal{P}_2, \quad (2.10)$$

with

$$\begin{aligned} \mathcal{P}_1 &= \frac{\psi(1, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}(\psi(1 - iu, v, r, q, \tau) e^{iu(\ln k)})}{u} du \\ \mathcal{P}_2 &= \frac{\psi(0, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}(\psi(-iu, v, r, q, \tau) e^{iu(\ln k)})}{u} du, \end{aligned} \quad (2.11)$$

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

The integrations in (2.11) 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

Empirical analyses of spot and option prices are challenged by the richness of the data. On any given day, in addition to observing the underlying index,¹⁰ one observes a cross section of option prices with different maturities and moneyness. Such a rich structure of spot and 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.

Combining the two sources of information from spot and option prices has long been advocated in the literature (see, for example, Melino [1994], Renault [1997], Chernov and Ghysels [2000], and references therein). While there has been a number of empirical studies using both spot and option prices, most of them examine the dynamic information of these two dataset separately, and then draw inferences by comparing the two separate estimation results.¹¹ Chernov and Ghysels [2000] is one notable exception, who use time-series data of spot and option prices simultaneously to estimate the Heston Model. (See also Jones [1999].)

¹⁰Parallel to the empirical studies that use options data to learn about the underlying dynamics, there is a vast literature on time-series estimation of stochastic volatility models using the cash market data. See, for example, Taylor [1986], Melino and Turnbull [1990], Harvey, Ruiz, and Shephard [1994], Jacquier, Polson, and Rossi [1994], Danielsson [1994], Gallant, Hsieh, and Tauchen [1997], Fouque, Papanicolaou, and Sircar [2000], and others. More recent empirical studies that focus on the role of stochastic volatility as well as jumps include Andersen, Benzoni, and Lund [1998], Johannes, Kumar, and Polson [1998], Eraker, Johannes, and Polson [1999], Chacko and Viceira [1999], and Chernov, Gallant, Ghysels, and Tauchen [1999].

¹¹Benzoni [1998], Pastorello, Renault, and Touzi [1996] and others. See also the literature on the forecastability of option prices for future stock volatility: Day and Lewis [1992], Canina and Figlewski [1993], Lamoureux and Lastrapes [1993], Christensen and Prabhala [1998], Figlewski [1997], and Poteshman [2000].

Given that our main objective is to recover the information about the risk premium implicit in option prices, this approach of using spot and options data simultaneously is indeed well motivated. In an arbitrage-free option-pricing model such as the one adopted in this paper, the market observed security prices, both for the underlying and for the options, are driven by the same set of state variables and pricing kernel. The most efficient way to learn about the pricing kernel is therefore to examine the time series of spot and option prices simultaneously.

In this paper, we adopt an estimation strategy that is along the line of the generalized method of moments (GMM) of Hansen [1982], and the efficient GMM of Hansen [1985]. The basic idea of our GMM approach is as follows. On any date n , we observe the spot price S_n , and an option price C_n . Assuming S_n and C_n are measured without error, we invert, for a given set of model parameters ϑ , a proxy V_n^ϑ for the unobserved volatility V_n , by solving V_n^ϑ for $C_n = S_n f(V_n^\vartheta, \vartheta)$. Quite intuitively, the accuracy of this proxy V_n^ϑ depends on how close ϑ is to the true model parameter ϑ_0 . Equipped with V_n^ϑ , we can now examine the joint distribution of spot and option prices by focusing directly on the dynamic structure of the state variables (S, V) , which fall into the class of affine jump-diffusions. This affine structure of the state variables (S, V) allows us to calculate their conditional moment-generating function in closed-form, which in turn provides a rich set of moment conditions. Replacing V_n by V_n^ϑ in the moment conditions, we can perform the usual GMM estimation — the only difference is that one of the state variables V^ϑ is parameter-dependent. (Hence the term “implied-state” GMM.)

This IS-GMM approach falls into a group of estimation strategies for state variables that can only be observed up to unknown model parameters.¹² Renault and Touzi [1996] develop an MLE-based two-step iterative procedure that can be applied to the time series of option prices. (See Patilea and Renault [1997] for a more general treatment in the spirit of this iterative and recursive approach.) Pastorello, Renault, and Touzi [1996] apply SMM to time series of spot and option prices separately. Chernov and Ghysels [2000] adopt the SNP/EMM empirical strategy developed by Gallant and Tauchen [1998] and apply a simultaneous time-series estimation of spot and option prices. Singleton [2000] develops a conditional-characteristic-function-based estimation method for the general class of affine jump-diffusions.

The main motivation for us to take on this GMM approach is to exploit the affine structure of our setting. Using the explicitly known conditional moment-generating function, we can construct moment conditions, as well as “efficient” instruments (in the spirit of Hansen [1985]) for our moment conditions. Moreover, the conditional moments of the state variables also provide a rich set of diagnostic tests, allowing us to examine explicitly how well various model constraints (for example, constraints on the risk premium) fit with the joint time-series

¹²This econometric setting arises in 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. 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. (See, for example, Duffie and Singleton [1999] and Duffee [1999], and references therein.) If there exist market-traded instruments whose values are linked to such events, then the otherwise-unobserved intensity processes can be “backed out.”

data of spot and option prices.

It should be noted that while our approach relies on one option (C_n) a day to back out V_n^ϑ , it does not preclude us from using multiple options. In fact, in Section 5.2, we introduce, in addition to the time series of spot and near-the-money short-dated option prices, a time series of in-the-money call options to help us identify jump-risk premium simultaneously with volatility-risk premium. But our approach does assume that this particular time series of option prices $\{C_n\}$ is measured precisely. This is partly the motivation for us to use near-the-money short-dated options, as they usually are the most liquid options.

In the rest of the section, we will first provide a detailed description of the IS-GMM estimators, then provide a discussion on how to choose the moment conditions optimally. The large sample properties of the IS-GMM estimators are established in Appendix A.

3.1 “Implied-State” GMM Estimators

Fixing¹³ some time interval Δ , we sample the continuous-time state process $\{S_t, V_t, r_t, q_t\}$ at discrete times $\{0, \Delta, 2\Delta, \dots, N\Delta\}$, and denote the sampled process $\{S_{n\Delta}, V_{n\Delta}, r_{n\Delta}, q_{n\Delta}\}$ by $\{S_n, V_n, r_n, q_n\}$. Letting¹⁴

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

denote the date- n “excess” return, it is easy to see that transition distribution of $\{y_n, V_n\}$ 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

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

where ϑ_0 is the true model parameter, $h : \mathbb{R}^{n_y} \times \mathbb{R}_+^{n_v} \times \Theta \rightarrow \mathbb{R}^{n_h}$ is some test function to be chosen,¹⁵ 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 ,

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

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

¹³In this section, we treat the parameters θ_r and θ_q associated with the dynamics of risk-free rate r and dividend yields q as given. In practice, we first obtain maximum-likelihood (ML) estimates of θ_r and θ_q using time series of interest rates and dividend yields, respectively. We then treat the ML estimates of θ_r and θ_q as true parameters, and adopt the “implied-state” GMM estimation strategy outlined here. Any loss of efficiency from this 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.

¹⁴In 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 $\tilde{y}_n = \ln S_n - \ln S_{n-1} - (r_{n-1} - q_{n-1}) \Delta$ 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. Alternative proxies for $\int_{(n-1)\Delta}^{n\Delta} (r_t - q_t) dt$, such as $(r_n - q_n) \Delta$ and $[(r_n + r_{n-1})/2 - (q_n + q_{n-1})/2] \Delta$, are also considered. The empirical results reported in this paper are robust with respect to all three proxies.

¹⁵We assume that h is continuously differentiable and integrable in the sense of (3.2).

Setting our situation apart from that of a typical GMM, we do not directly observe V_n . Our approach is to take advantage of market-observed spot price S_n and option price C_n , and explore the option-pricing relation. To be more specific, 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 (2.10),

$$c_n = f(V_n, \vartheta_0, r_n, q_n, \tau_n, k_n), \quad (3.3)$$

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 (2.10), in that Ξ is the maximal set for which a mapping $g : \Xi \rightarrow \mathbb{R}_+$ is uniquely defined by

$$f(g(c, \vartheta, r, q, \tau, k), \vartheta, r, q, \tau, k) = c, \quad (3.4)$$

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

$$V_n^\vartheta = g(c_n, \vartheta, r_n, q_n, \tau_n, k_n). \quad (3.5)$$

One important property of V_n^ϑ is that that the true date- n stochastic volatility V_n is retrieved when V_n^ϑ is evaluated at the true model parameter ϑ_0 . The concept of option-implied volatility is not new. A prominent example is the Black-Scholes implied volatility.

Given the option-implied volatility V_n^ϑ , we can now construct the sample analogue of the moment condition (3.2) by

$$G_N(\vartheta) = \frac{1}{N} \sum_{n \leq N} h(y_{(n, n_y)}, V_{(n, n_v)}^\vartheta, \vartheta), \quad (3.6)$$

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

$$\hat{\vartheta}_N = \arg \min_{\vartheta \in \Theta} G_N(\vartheta)^\top \mathcal{W}_N G_N(\vartheta), \quad (3.7)$$

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

Consistency and asymptotic normality of the IS-GMM estimators are established in Appendix A, which addresses the econometric issue raised by the use of options with time-varying contract variables such as maturity $\{\tau_n\}$ and strike-to-spot ratio $\{k_n\}$.

Our 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 [2000], 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 “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}) . Under certain integrability conditions (Duffie, Pan, and Singleton [1999]), one can show that, for any u_y and u_v in \mathbb{R} ,

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

where $\phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is defined by¹⁶

$$\phi(u_y, u_v, v) = \exp (A(u_y, u_v) + B(u_y, u_v) v), \quad (3.8)$$

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 [2000] 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 [2000]. 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

$$E_n (y_{n+1}^i V_{n+1}^j) = \frac{\partial^{(i+j)} \phi(u_y, u_v, V_n)}{\partial^{i} u_y \partial^{j} u_v} \Big|_{u_y=0, u_v=0}, \quad i, j \in \{0, 1, \dots\}. \quad (3.9)$$

Direct computation of the derivatives in (3.9), although straightforward, can be cumbersome for higher orders of i and j .¹⁷ 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.

We let $M : \mathbb{R}_+ \times \Theta \rightarrow \mathbb{R}^7$ denote the list of conditional moments of (y, V) defined by $M_1(V_n, \vartheta) = E_n^\vartheta(y_{n+1})$, $M_2(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^2)$, $M_3(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^3)$, $M_4(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^4)$, $M_5(V_n, \vartheta) = E_n^\vartheta(V_{n+1})$, $M_6(V_n, \vartheta) = E_n^\vartheta(V_{n+1}^2)$, 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

$$E_{n-1}^\vartheta(\epsilon_n) = 0, \quad \epsilon_n = [\epsilon_n^{y1}, \epsilon_n^{y2}, \epsilon_n^{y3}, \epsilon_n^{y4}, \epsilon_n^{v1}, \epsilon_n^{v2}, \epsilon_n^{yv}]^\top, \quad (3.10)$$

where

$$\begin{aligned} \epsilon_n^{y1} &= y_n - M_1(V_{n-1}, \vartheta), & \epsilon_n^{y2} &= y_n^2 - M_2(V_{n-1}, \vartheta) \\ \epsilon_n^{y3} &= y_n^3 - M_3(V_{n-1}, \vartheta), & \epsilon_n^{y4} &= y_n^4 - M_4(V_{n-1}, \vartheta) \\ \epsilon_n^{v1} &= V_n - M_5(V_{n-1}, \vartheta), & \epsilon_n^{v2} &= V_n^2 - M_6(V_{n-1}, \vartheta) \\ \epsilon_n^{yv} &= y_n V_n - M_7(V_{n-1}, \vartheta). \end{aligned} \quad (3.11)$$

¹⁶See also Heston [1993], Bates [2000], and Das and Sundaram [1999].

¹⁷Das 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.

This choice (3.10) 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 (3.10). In the spirit of Hansen [1985], we define the “optimal” moment conditions by

$$\mathcal{H}_{n+1} = \mathcal{Z}_n \epsilon_{n+1}, \quad \text{with} \quad \mathcal{Z}_n = \mathcal{D}_n^\top \times (\text{Cov}_n^\vartheta(\epsilon_{n+1}))^{-1}, \quad (3.12)$$

where $\text{Cov}_n^\vartheta(\epsilon_{n+1})$ 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

$$\begin{aligned} \mathcal{D}_n^i &= -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta} - g_\vartheta(c_n, \vartheta) \frac{\partial M_i(v, \vartheta)}{\partial v} \Big|_{v=V_n}, & i = 1, 2, 3, 4, \\ \mathcal{D}_n^i &= -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta}, & i = 5, 6, 7, \end{aligned} \quad (3.13)$$

where $c_n = f(V_n, \vartheta_0)$ with f given by the option-pricing formula (2.10), and where $g_\vartheta(c, \vartheta) = \partial g(c, \vartheta) / \partial \vartheta$ with g defined by (3.4). (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

$$g_\vartheta(c_n, \vartheta) = -f_v^{-1}(V_n, \vartheta) f_\vartheta(V_n, \vartheta), \quad (3.14)$$

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 (2.10). The partial derivatives f_v and f_ϑ can be calculated explicitly, up to numerical integration, by differentiating through the integrals in (2.11).

Each element \mathcal{H}_{n+1}^j of the “optimal” observations $\mathcal{H}_{n+1} = (\mathcal{H}_{n+1}^1, \dots, \mathcal{H}_{n+1}^{n_\vartheta})$ 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 3.7 by replacing the unobserved stochastic volatility V_n with the option-implied stochastic volatility V_n^ϑ .

Finally, it should be noted¹⁸ that the efficiency of this “optimal-instrument” scheme is limited in that the Jacobian \mathcal{D} constructed in (3.13) for $i = 5, 6, 7$ differs from that of Hansen [1985]. Specifically, 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.

¹⁸When 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 [1999] for details.

4 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).

4.1 S&P 500 Index and Near-the-Money Short-Dated Options

We construct a time-series $\{S_n, C_n\}$ of the S&P 500 index and near-the-money short-dated option prices, from January 1989 to December 1996, with “weekly” frequency (every 5 trading days). Details are 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.¹⁹ 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.²⁰ 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 6.

4.2 Time Series of In-the-Money Short-Dated Calls

On each date n , we select an in-the-money call C_n^{ITM} with the same maturity as the near-the-money option C_n described above, but a different strike price. Among all of the possible ITM call options (with strike price less than that of the near-the-money option C_n), we select the one C_n^{ITM} with strike-to-spot ratio closest to 0.95. If there is no such ITM calls available, we choose an OTM put, again with strike-to-spot ratio closest to 0.95, and convert the price to that of an ITM call using put/call parity. The sample mean of the strike-to-spot ratio $\{k_n^{ITM}\}$ is 0.952 with a sample standard deviation of 0.007.

¹⁹Both time to expiration τ_n and sampling interval Δ are annualized, using a 365-calendar-day year and a 252-business-day year, respectively.

²⁰One 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.

4.3 Cross-Sectional Data on Calls and Puts

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 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.²¹ There are in total 11,434 observations for the group of high-volatility days, 33,919 observations for the medium-volatility days, and 19,589 observations for the low-volatility days.

5 Empirical Results

Estimation results are organized as following: the role of risk premia in reconciling spot and option dynamics, estimating jump-risk premia and volatility-risk premia simultaneously using an additional time-series of in-the-money calls, diagnostic tests on volatility dynamics, and implications of the cross-sectional behavior of option prices. The estimation results associated with the risk-free rates r and dividend yields q are given in Appendix H.

5.1 Risk Premia: Reconciling Spot and Option Dynamics

In this section, we use the time series $\{S_n, C_n\}$ of S&P 500 index and near-the-money short-dated option prices to estimate the parameters of the SVJ and nested models, and examine how well the models accommodate the joint time-series data. In particular, we focus on the role of risk premia in reconciling the spot and option dynamics. We find jump-risk premia to play a crucial role.

In order to understand the importance of risk premia in options, we start with a nested model with zero volatility-risk premia ($\eta^v = 0$), and no jumps ($\lambda_0 = \lambda_1 = 0$), therefore zero jump-risk premia. We call this model the SV0 model. Using the results of the goodness-of-fit tests²² summarized in Table 2, we see that the SV0 model is strongly rejected by the joint

²¹For example, this happens when the quoted option price is less than its intrinsic value.

²²These tests focus on how well the respective models satisfy the fundamental moment conditions (3.10), and are constructed directly from the heteroskedasticity-corrected version $\tilde{\epsilon}$ of ϵ , defined by $\tilde{\epsilon}_n^i = \epsilon_n^i / \sqrt{E_{(n-1)}(\epsilon_n^i)^2}$, $i \in \{1, \dots, 7\}$. 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. 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.

time-series data. In particular, 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 $E(\tilde{\epsilon}_n^{y^2}) = 0$, 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 $\epsilon_n^{y^2} \approx y_n^2 - V_{n-1}^\vartheta \Delta$. The significantly negative test statistics associated with $E(\epsilon_n^{y^2})$ therefore indicates that the volatility realized in the spot market is significantly less than that observed from the options market (through the SV0 model).

Table 1: 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 (1.9)	0.0242 (0.0044)	0.38 (0.04)	-0.57 (0.05)	4.4 (1.8)	$\equiv 0$	$\equiv 0$			
SV	7.1 (2.1)	0.0137 (0.0023)	0.32 (0.03)	-0.53 (0.06)	8.6 (2.3)	7.6 (2.0)	$\equiv 0$			
SVJ0	7.1 (1.9)	0.0134 (0.0029)	0.28 (0.04)	-0.52 (0.07)	3.1 (2.9)	$\equiv 0$	27.1 (11.8)	-0.3% (1.7%)	3.25% (0.64%)	-18.0% (1.6%)
SVJ	7.5 (2.0)	0.0135 (0.0026)	0.28 (0.04)	-0.49 (0.09)	-1.5 (6.9)	-7.5 (11.6)	55.8 (54.3)	-0.5% (1.3%)	2.70% (0.73%)	-17.1% (1.6%)

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

One possible explanation is that the SV0 model does not incorporate investors' aversion toward the risk of changes in volatility. To examine the role of volatility-risk premia, we relax the assumption that volatility risk is not priced, and treat the volatility-risk premia coefficient η^v as a free parameter. We call the relaxed model the SV model. Evidence in support of volatility-risk premia can be summarized as following.²³ First, Table 1 shows that 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.") Second, letting $\mathcal{H}_{n+1}(\eta^v)$ denote the "optimal" moment associated with η^v , as described in Section 3.2, we further perform a Lagrange-multiplier test²⁴ of the SV0 model against the SV model using the moment condition $E_n[\mathcal{H}_{n+1}(\eta^v)] = 0$. The SV0 model (that with $\eta^v = 0$) is rejected against the SV model, with a p -value of 0.0002. Third, Table 2 shows that the moment condition $E(\epsilon_n^{y^2}) = 0$ is no longer strongly violated for the SV model.

The above findings suggest that allowing volatility-risk premia partially reconcile the tension between the spot and option prices that arises in the SV0 model. Parallel to such

²³The significance of volatility-risk premia is also reported in Guo [1998], Benzoni [1998], Poteshman [1998], Kapadia [1998], and Chernov and Ghysels [2000]. As volatility-risk premia are 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.

²⁴This 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).

Table 2: Goodness-of-Fit Tests

	$\tilde{\epsilon}$	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 $\chi^2(4)$	28.2** (10^{-5})	8.1 (0.09)	1.8 (0.77)	1.4 (0.84)
	all v $\chi^2(2)$	9.1* (0.01)	11.4** (0.003)	3.2 (0.20)	3.1 (0.21)
	all $\chi^2(7)$	59.9** (10^{-10})	31.6** (10^{-5})	7.6 (0.37)	7.1 (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.

supporting evidence, there is also strong evidence suggesting the inability of the volatility-risk premia in reconciling spot and option dynamics. In particular, Table 2 shows that the joint test on $E(\tilde{\epsilon}_n) = 0$ is still strongly rejected (with a p -value of 10^{-5}) in the SV model. Moreover, the SV model estimates imply a stochastic-volatility process that is explosive under the risk-neutral measure, 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 illustrated in Figure 1.

We next examine the role of jump-risk premia. Taking the SV0 model as the starting point, we introduce jump and jump-risk premium by relaxing the constraint that $\lambda_1 = 0$ and treating λ_1 as a free parameter. We call this relaxed jump model the SVJ0 model. In this model, the premia for the jump-size risk are captured by the difference between the mean jump size μ and its risk-neutral counterpart μ^* , while the volatility-risk premia are set to be zero. Our first observation is that, in contrast to the SV model, the SVJ0 model is not rejected by the joint time-series data. (The p -value of the joint test of the hypothesis that $E(\tilde{\epsilon}_n) = 0$ is 0.34 for the SVJ0.)

One important difference between the two types of risk premia considered in this paper is that, for short-dated options, the jump-risk premia respond quickly to market volatility, while the volatility-risk premia do not. Figure 2 shows that when the market volatility doubles from 10% to 20%, the percentage premium paid for jump risk nearly doubles, while that for volatility risk increases only by a small amount. This different responsiveness to the market volatility implies that volatility-risk premia under-estimate the risk premia implicit in

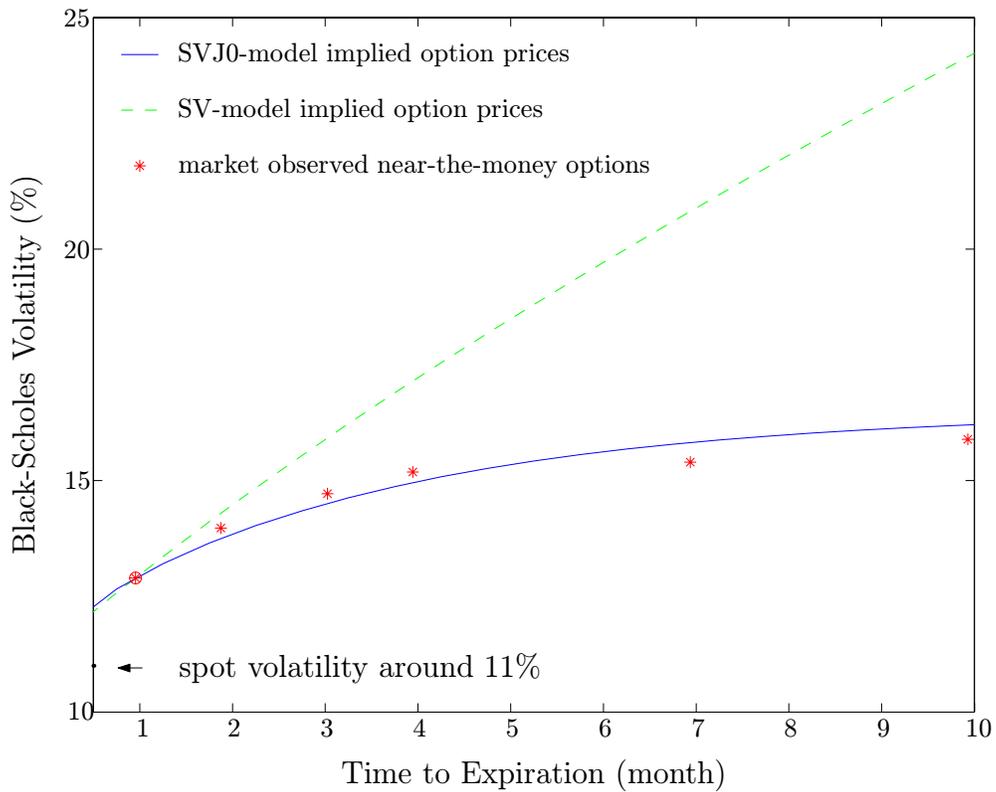


Figure 1: On a medium-volatility day (Nov. 22, 96), the SVJ0 model (with jump-risk premia) is capable of pricing both short-dated and long-dated options, while the SV model (with volatility-risk premia) severely over-prices long-dated options.

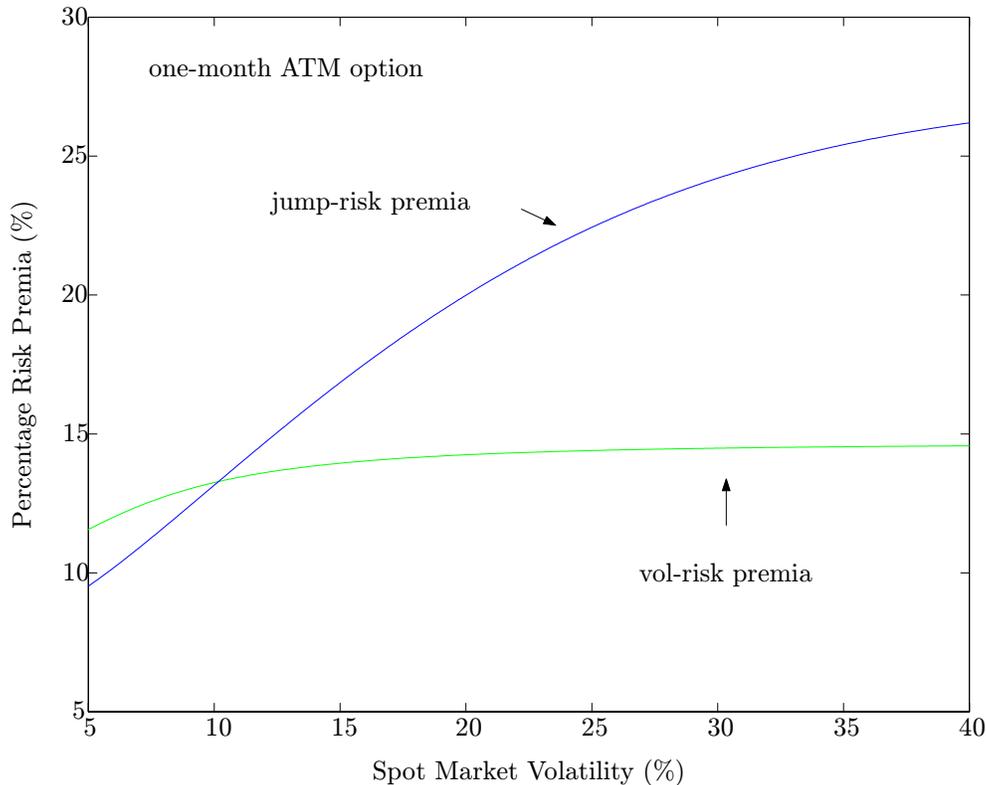


Figure 2: Different responsiveness to market volatility. The percentage risk premia measure the amount of cents one has to pay for risk premia for every dollar invested in the one-month at-the-money options. The market volatility is in terms of the Black-Scholes volatility.

short-dated options during high-volatility periods, while over-estimating during low-volatility periods. This partially explains why the SV model is strongly rejected by the joint time-series data.

Another important difference between the two types of risk premia is that premia for volatility risk take time to build up, while those for jump risk pick up the action right away. Consequently, in order to explain the risk premia implicit in the near-the-money short-dated options, the volatility-risk premia have to work really hard within the short time horizon, resulting in an explosive risk-neutral volatility process that severely over-prices long-dated options (Figure 1). The jump-risk premia, on the hand, provide a better characterization of the risk premia implicit in the long- and short-dated options, as illustrated in Figure 1.²⁵

Finally, we examine the impact of the constraints, $\eta^v = 0$ and $\lambda_0 = 0$, imposed by the SVJ0 model. Letting $\mathcal{H}_n(\eta^v)$ denote the “optimal” moment associated with η^v , we test the SVJ0 model against the alternative that $\eta^v \neq 0$. Using the moment condition $E[\mathcal{H}_n(\eta^v)] = 0$, which is true under the alternative, we perform a Lagrange-multiplier test of SVJ0. The SVJ0 model (that with $\eta^v = 0$) is not rejected against the alternative that $\eta^v \neq 0$ at traditional confidence levels. (The p -value is 0.55.) Similarly, letting $\mathcal{H}_n(\lambda_0)$ denote the “optimal”

²⁵See Section 5.4 for a detailed description of the SV and SVJ0 models in capturing options across maturity.

moment associated with λ_0 , we test the SVJ0 model against the alternative that $\lambda_0 \neq 0$. The SVJ0 model is not rejected against the alternative that $\lambda_0 \neq 0$ at traditional confidence levels.²⁶ (The p -value is 0.12.)

5.2 Risk Premia: Estimated Using Additional Options

In this section, we focus on the SVJ model, which relaxes from the SVJ0 model by allowing the volatility-risk premia coefficient η^v to be a free parameter, therefore incorporating both volatility-risk premia and jump-risk premia.

We first use the joint time series $\{S_n, C_n\}$ to estimate the SVJ model. From Table 1, we see that the relaxed parameter η^v is estimated to be negative, while the jump-arrival coefficient λ_1 is estimated to be twice as much as its SVJ0-model counterpart. In other words, the estimated SVJ model “overstates” (relative to the SVJ0 model) jump-risk premium, compensating it with negative volatility-risk premia. Moreover, neither of the estimated parameters η^v and λ_1 is statistically significant. In other words, as we simultaneously incorporate the premia for volatility and jump risks, the joint time series of spot and near-the-money short-dated option prices becomes inadequate to help us pin down the relative magnitude of volatility-risk and jump-risk premia.

As we extend our analysis to cross-sectional options data in Section 5.4, this “overstated” jump-risk premium shows up in the form of exaggerated volatility “smirks” implied by the estimated SVJ model. Using this intuition, we introduce an additional time series C_n^{ITM} of in-the-money calls (data collection details given in Section 4.2). We assume that such ITM calls are priced with errors. The pricing errors are assumed to be independent across time, and independent of the three sources of uncertainty in the SVJ model. Moreover, we assume that, on each date n , the standard deviation of the date- n pricing error is proportional the bid ask spread.

Letting c_n^{ITM} and k_n^{ITM} be the date- t price-to-spot ratio and strike-to-spot ratio of the in-the-money call option, we form the moment condition

$$E(\mathcal{H}_n^{ITM}) = 0, \quad \mathcal{H}_n^{ITM} = \frac{c_n^{ITM} - f(V_n, \vartheta, r_n, q_n, \tau_n, k_n^{ITM})}{\delta_n^{ITM}}, \quad (5.1)$$

where δ_n^{ITM} is the date- n bid ask spread (per unit spot price), and f is the SVJ option pricing formula defined in (2.10). In addition to the 10 “optimal” moment conditions associated with the 10 SVJ-model parameters, we introduce the moment condition (5.1) to take advantage of the risk-premia information embedded in the ITM call (or OTM put) option prices.

²⁶It should be noted that this test result does not imply that the constant component λ_0 of the jump arrival intensity is not important in capturing the jump behavior in the index dynamics. In fact, using longer time-series data with daily frequency, Chernov, Gallant, Ghysels, and Tauchen [1999] report evidence in support of such a constant component. Given the relatively short sample and weekly frequency used in this paper, our ability to pin down the jump component in the index dynamics is very limited. (For example, the estimate for μ is not significant at all.) The key point one would want to take away from this result is that in terms of reconciling spot and option dynamics, the state-dependent component λ_1 plays a far more important role than the constant component λ_0 . In other words, the state-dependent aspect of the jump-risk premium dominates.

Table 3: Estimating the SVJ Model using an Additional Time Series of ITM Calls

κ_v	\bar{v}	σ_v	ρ	η^s	η^v	λ_1	μ	σ_J	μ^*
6.4	0.0153	0.30	-0.53	3.6	3.1	12.3	-0.8%	3.87%	-19.2%
(1.8)	(0.0029)	(0.04)	(0.07)	(2.4)	(2.2)	(1.9)	(2.4%)	(0.72%)	(1.8%)

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

From the estimation results summarized in Table 3, we see that, by using the additional time series of ITM call options as a disciplinary measure, we obtain parameter estimates for η^v and λ_1 that are very different from those reported in Table 1 for the SVJ model. In particular, the volatility-risk premia coefficient η^v is estimated to be positive, indicating a positive volatility-risk premia. The accuracy of the estimator for λ_1 is greatly improved, but η^v is still statistically insignificant.²⁷ Looking for possible tensions in the system, we repeat the goodness-of-fit tests (as those reported in Table 2) for this model. The p-value of the overall goodness-of-fit test is 0.40. One can also examine the goodness of fit by looking at the over-identifying restriction introduced by the additional moment condition (5.1). The p-value for that test is 0.58.

To understand the nature of the jump-risk premia estimated in the SVJ model, we take a close examination of the parameter estimates in Table 1. The risk-neutral mean μ^* of the relative jump size is estimated to be -19%, while its counterpart μ for the data-generating process is estimated to be -0.8%. This implies that, when weighted by aversion to large price movements, negative jumps are perceived to be more negative ($\mu^* - \mu$ is estimated at -18%, with a standard error of 4%). 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.

Under the setting of the SVJ model, the price of an option now has three components: the first is associated with volatility-risk premia, the second associated with jump-risk premia, while the third associated with the actual dynamics. For intuition purpose, we offer some

²⁷Given the intuition that volatility-risk premia take time to build up, therefore playing a more important role in long-dated options, one can expect to obtain a more accurate estimator for the volatility-risk premium coefficient η^v using an additional time series of long-dated options. On the other hand, this attempt to identify volatility-risk premia through the help of long-dated options could easily be overshadowed by a more seriously limitation of the SVJ model. Specifically, the inability of the SVJ model to capture the "long memory" in volatility. See the next two sections for details.

rough estimates of the breakdown between the two types of risk premia. Consider a one-month option traded at 20% market volatility, if the option is at the money, then about 65% of the overall risk premia is paid for jump risk, and the rest for volatility risk; if the option is a 10% OTM put, then the jump-risk premium component increases to 85%; for a 10% OTM call option, on the other hand, this percentage decreases to 40%. In other words, investors are more worried about jump risk in OTM puts than OTM calls. Moreover, the premia for jump risk become a more important component of the overall risk premia during volatile markets. This is, of course, consistent with the intuition that jump-risk premia are far more responsive to market volatility than volatility-risk premia.

Under the setting of the SVJ model, the time- t instantaneous equity risk premium has two components: the first component $\eta^s V_t$ compensates for the usual diffusive return risk, and the second component $\lambda_1 V_t (\mu - \mu^*)$ compensates for the jump-size uncertainty. Measuring the average volatility level by its long run-mean \bar{v} , the SVJ-model estimates imply that the average mean excess rate of (cum-dividend) return demanded for usual return risk is 5.5% per year (with a standard deviation of 3.4%), while that demanded for jump risk is 3.5% per year (with a standard deviation of 0.7%). This shows that a significant portion of the equity risk premium is assigned as a premium to compensate for investors' aversion toward jumps. Given that less than 3% of the total return variance is due to jump risk, these numbers suggest that compensation for jump risk is very different from that for diffusive risk.

5.3 Volatility Dynamics: Diagnostic Tests

In this section, we look at possible model mis-specification, focusing in particular on the volatility dynamics.

Our first empirical finding is that the stochastic volatility model of Heston [1993] is not rich enough to capture the term structure of volatility implied by the data.²⁸ Letting $E_n(\epsilon_{n+1}^{v1}) = 0$ be the fundamental moment condition associated with the first moment of volatility, the SVJ model implies that ϵ^{v1} is serially uncorrelated. Table 4 shows that the hypothesis of $E(\epsilon_n^{v1} \epsilon_{n+1}^{v1}) = 0$ is strongly rejected by all five cases considered in this paper. The sample estimate of $\text{corr}(\epsilon_n^{v1}, \epsilon_{n+1}^{v1})$ 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 auto-correlation can be calculated as²⁹

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

Using the fact that the one-step auto-correlation $\text{corr}(V_n, V_{n-1})$ is $\exp(-\kappa_v \Delta)$, we have $\text{corr}(\epsilon_n^{v1}, \epsilon_{n-1}^{v1}) = e^{-\kappa_v \Delta t} [e^{-2\kappa_v \Delta t} - \text{corr}(V_{n+1}, V_{n-1})]$. A negative and significant sample estimate of $\text{corr}(\epsilon_n^{v1}, \epsilon_{n-1}^{v1})$ therefore indicates that the data call for $\text{corr}(V_{n+1}, V_{n-1}) > \exp(-2\kappa_v \Delta t)$.

²⁸Similar findings on the long memory of stochastic volatility implied by options data can be found in Stein [1989] and Bollerslev and Mikkelsen [1996].

²⁹Here, 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.

This is in contrast to the model-prescribed two-step auto-correlation $\text{corr}(V_{n+1}, V_{n-1}) = \exp(-2\kappa_v \Delta t)$. In other words, the stochastic volatility model is not capable of fitting one- and two-step auto-correlations simultaneously. The reason is quite simple: The model prescribes a term-structure of volatility, $\text{corr}(V_n, V_{n+m}) = \exp(-m\kappa_v \Delta t)$, which “dies” too quickly, relative to the data.³⁰

Table 4: Diagnostic Tests on Volatility Dynamics

	SV0	SV	SVJ0	SVJ	SVJ ^{ITM}
$E(\epsilon_n^v \epsilon_{n+1}^v) = 0$	-2.37*	-2.62**	-2.04*	-2.19*	-2.38*
$E(\epsilon_n^{v^3}) = 0$	-0.69	0.85	0.70	0.54	-0.59
$E(\epsilon_n^{v^4}) = 0$	-0.38	0.65	0.62	0.53	0.05
$E(\tilde{\epsilon}_n^{v^3}) = 0$	-0.58	1.46	2.17*	2.59**	0.13
$E(\tilde{\epsilon}_n^{v^4}) = 0$	0.95	1.68	2.67**	3.75**	1.03

* and ** indicate significance under a 5% and 1% test, respectively. SVJ^{ITM} indicates the case in Section 5.2 where one additional time series of ITM calls is used in model estimation.

We next focus on the third and fourth moments of the volatility process, looking for evidence of jumps in volatility, as conjectured by Bates [2000]. Letting $E_n(\epsilon_{n+1}^{v^3}) = 0$ and $E_n(\epsilon_{n+1}^{v^4}) = 0$ be the moment conditions associated with the third and fourth moments of volatility, we report in Table 4 the conditions $E(\epsilon_{n+1}^{v^3}) = 0$ and $E_n(\epsilon_{n+1}^{v^4}) = 0$, and their respective heteroskedasticity-corrected versions, $E(\tilde{\epsilon}_{n+1}^{v^3}) = 0$ and $E_n(\tilde{\epsilon}_{n+1}^{v^4}) = 0$. The heteroskedasticity-corrected ($\tilde{\epsilon}$) tests evidently have more asymptotic power than their respective uncorrected (ϵ) counterparts. For the SVJ0 and SVJ models, the sample estimates of the moment conditions $E_n(\tilde{\epsilon}_{n+1}^{v^3}) = 0$ and $E_n(\tilde{\epsilon}_{n+1}^{v^4}) = 0$ are found to be positive and significantly different from zero, indicating the possibility of jumps (with positive mean jump size) in the stochastic-volatility process,³¹ or at least fatter-tailed innovations in the volatility process. Our overall findings, however, are mixed. It could very well be explained by the limited power of our test statistics.

Finally, we offer a brief discussion on volatility of volatility. Both Bates [2000] and Bakshi, Cao, and Chen [1997] report that in order to explain the volatility “smiles” and “smirks” found in the cross-sectional options data, the 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 Bates [2000] and Bakshi, Cao, and Chen [1997] is also reflected in our goodness-of-fit test associated with ϵ^{v^2} (Table 2). In this paper, however, we uncover this inconsistency from a time-series investigation of spot and near-the-money short-dated option prices. Our estimation results also suggest that a higher volatility

³⁰Multiple 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 [2000] and a stochastic-volatility model with stochastic long-run mean suggested by Duffie, Pan, and Singleton [1999].

³¹Examples 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].

of volatility coefficient σ_v , is usually associated with a higher long-run mean of volatility \bar{v} . This is consistent with findings (Jones [1999]) 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. Alternatively, one can introduce a second volatility factor.

5.4 Implications of Cross-Sectional Behavior

In this section, we extend our analysis to cross-sectional option data. Clearly, cross-sectional options data contain a wealth of information. For example, Bates [2000] and Bakshi, Cao, and Chen [1997] estimate essentially the same class of models (the risk-neutral version) as ours using the cross-sectional options data. One disadvantage of their approach is that the dynamic aspect of the state variable, and the time-series aspect of the spot and options data are not taken care of.³² The time-series approach adopted in this paper, on the other hand, takes full advantage of the dynamic information of the state variable, as well as the information implicit in the joint time-series data of spot and option prices. One disadvantage of our approach is that the rich information embedded in the cross-sectional option prices is not taken full advantage of.

This section sets out to learn more from the cross-sectional aspect of the data. (Details on data collection are given in Section 4.3.) Equipped with the time-series estimation results reported in Sections 5.1 and 5.2, we ask: How well do these models price the cross-sectional options data observed in the market? The short answer is: The results are surprisingly encouraging for the case of SVJ0, given that only very limited amount of cross-sectional information is impounded in the SVJ0 model estimates — just one option a day!

Table 5 summarizes the cross-sectional pricing errors.³³ Focusing first on the SVJ0 model, we see that on days of medium volatility, the pricing errors are well below (except for one case) the respective average bid/ask spreads³⁴ across all maturities and moneyness. This result is quite encouraging, given the limited dimensionality of the model (2 state variables) and the high dimensionality of the data. On days of high volatility, the SVJ0 model consistently under-prices medium and long-dated options, while on days of low volatility, the SVJ0 model consistently over-prices such options. This can be easily explained by the fact the one-factor stochastic volatility model considered here is not rich enough to account for the “long memory” of volatility (Section 5.3). In particular, the SVJ0 model reverts to its long-run mean too fast (compared with what is implied by the data).

For a pictorial exposition, we plot in Figures 3 volatility smiles (across different maturi-

³²In particular, 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 [2000] and Bakshi, Cao, and Chen [1997] show that the answer is “no.”)

³³The 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. The positive and negative signs in the parentheses indicate whether, on average, the model over-prices or under-prices.

³⁴The bid/ask spreads are measured by the difference between offer and ask prices, each measured in terms of *BS vol*.

Table 5: The SVJ 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 (-)
SVJ ^{ITM}		2.5 (-)	1.7 (-)	5.2 (-)	0.6 (+)	0.4 (-)	0.7 (-)	1.0 (-)	0.8 (-)	4.7 (-)
bid/ask		3.2	1.3	3.9	2.7	1.0	2.3	1.1	0.5	2.8
SV0	$60 < \tau \leq 90$	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 (-)
SVJ ^{ITM}		2.2 (-)	1.8 (-)	1.9 (-)	0.9 (+)	1.1 (+)	1.4 (+)	2.0 (+)	1.7 (+)	2.7 (-)
bid/ask		1.9	1.0	1.8	0.9	0.6	0.9	1.0	0.4	1.4
SV0	$\tau \geq 180$	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 (-)
SVJ ^{ITM}		2.6 (-)	1.6 (+)	2.2 (+)	2.4 (+)	2.8 (+)	3.2 (+)	3.2 (+)	3.6 (+)	4.0 (+)
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. SVJ^{ITM} indicates the case in Section 5.2 where one additional time series of ITM calls is used in model estimation.

ties) on a medium-volatility day. Again, we see that the SVJ0 model prices quite well options across all maturities and moneyness. This is quite remarkable, as only one option (marked by circle) is used, while the rest of the options are purely out of sample. Volatility smiles on the most and least volatile days are plotted in Figures 4 and 5, respectively. Again, we see that while the SVJ0 model can capture the volatility “smirks” relatively well, it under-prices long-dated options on the most volatile day, and over-prices long-dated options on the least volatile day. This is consistent with our time-series diagnostic results that the one-factor volatility mean-reverts too fast.

On days of high and low volatilities, we also see large volume of quotes for deep-in-the-money puts, whose prices none of our models are capable of explaining. 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 models. 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.

Our cross-sectional investigation shows that neither the SV0 nor SV model is capable of explaining options across moneyness. This result is not new. Fitting this class of stochastic volatility models directly to cross-sectional option prices, Bates [2000] and Bakshi, Cao, and Chen [1997] reach the same conclusion.

Table 5 also shows that the SV model severely over-prices long-dated options. As discussed in Section 5.1, the SV model relies on volatility-risk premia to explain the risk premia implicit in short-dated options. In doing so, the volatility-risk premia are severely overstated, resulting in an explosive volatility process under the “risk-neutral” measure, which in turn over-prices long-dated options.

For short-dated options, the SVJ model tends to exaggerate volatility “smirks,” which can be explained by the fact that the SVJ model compensates an “over-stated” jump-risk premia by an negative volatility-risk premia (Section 5.2). Under the setting of SVJ^{ITM} , we introduce an additional time series of ITM calls to help pin down the volatility-risk and jump-risk premia simultaneously. Consequently, we see a much improved characterization of volatility “smiles.”

6 Concluding Remarks

In this paper, we examine how different risk factors are priced in the S&P 500 index options, and provide strong evidence in support of a jump-risk premium that is highly correlated with the market volatility. We find that this jump-risk premium plays an important role in explaining both the joint time-series behavior of spot and option prices, and the cross-sectional behavior of option prices.

We conclude with some remarks on the economic implications of the pricing kernel estimated from the joint time-series data. A formal treatment, however, is beyond the scope of this paper. Using the estimation results reported in Section 5.2, we find that the excess mean rate of return demanded for the usual “diffusive” return risk is 5.5% per year, while that for jump risk is about 3.5% per year. Given that less than 3% of the total return variance is due to jump risk, these numbers indicate that the compensation for jump risk is very different

from that for diffusive risk. To explain these empirical results within the framework of rational expectations, it may be fruitful to explore utility models showing potentially extreme aversion to big losses or negative skewness³⁵ (as in, for example, Gul [1991]). Alternatively, if the spot and options markets are not fully integrated, then such significant jump-risk premia could be partially proxying for market frictions that are specific only to the options market.

³⁵Harvey and Siddique [2000] document evidence of systematic skewness using cross-sectional equity returns. While simple utility functions such as the one with constant relative risk aversion coefficient does allow aversion to variance and preference for skewness, the magnitude of the two is tied down by one parameter (the risk aversion coefficient). Explicit modeling of skewness preference can be found, among others, in Rubinstein [1973] and Kraus and Litzenberger [1976].

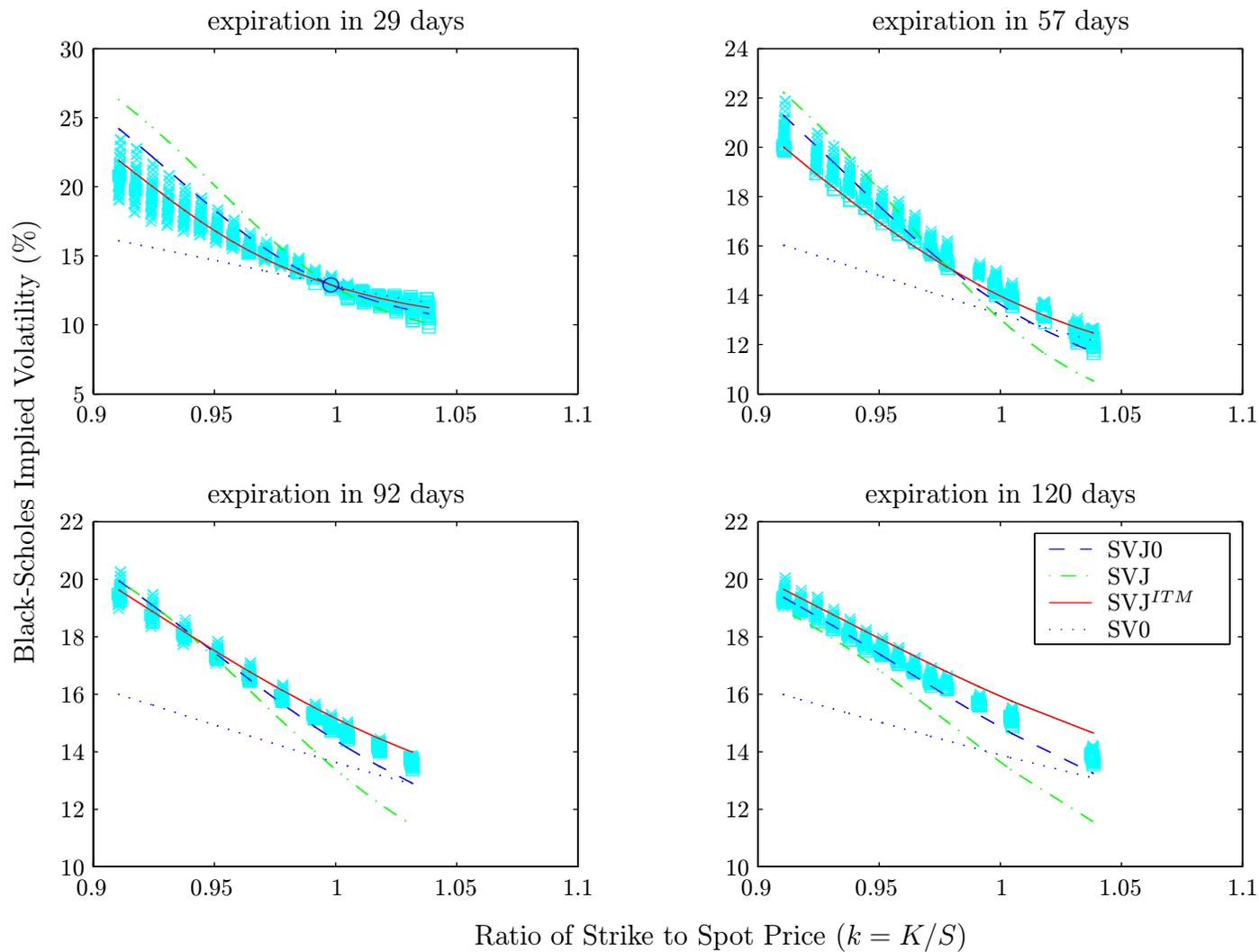


Figure 3: Smiles curves on a “medium-volatility” day. All observations are observed between 10:10am to 10:20am on November 22, 1996. The call options are marked by ‘×,’ and the put options by ‘□’.

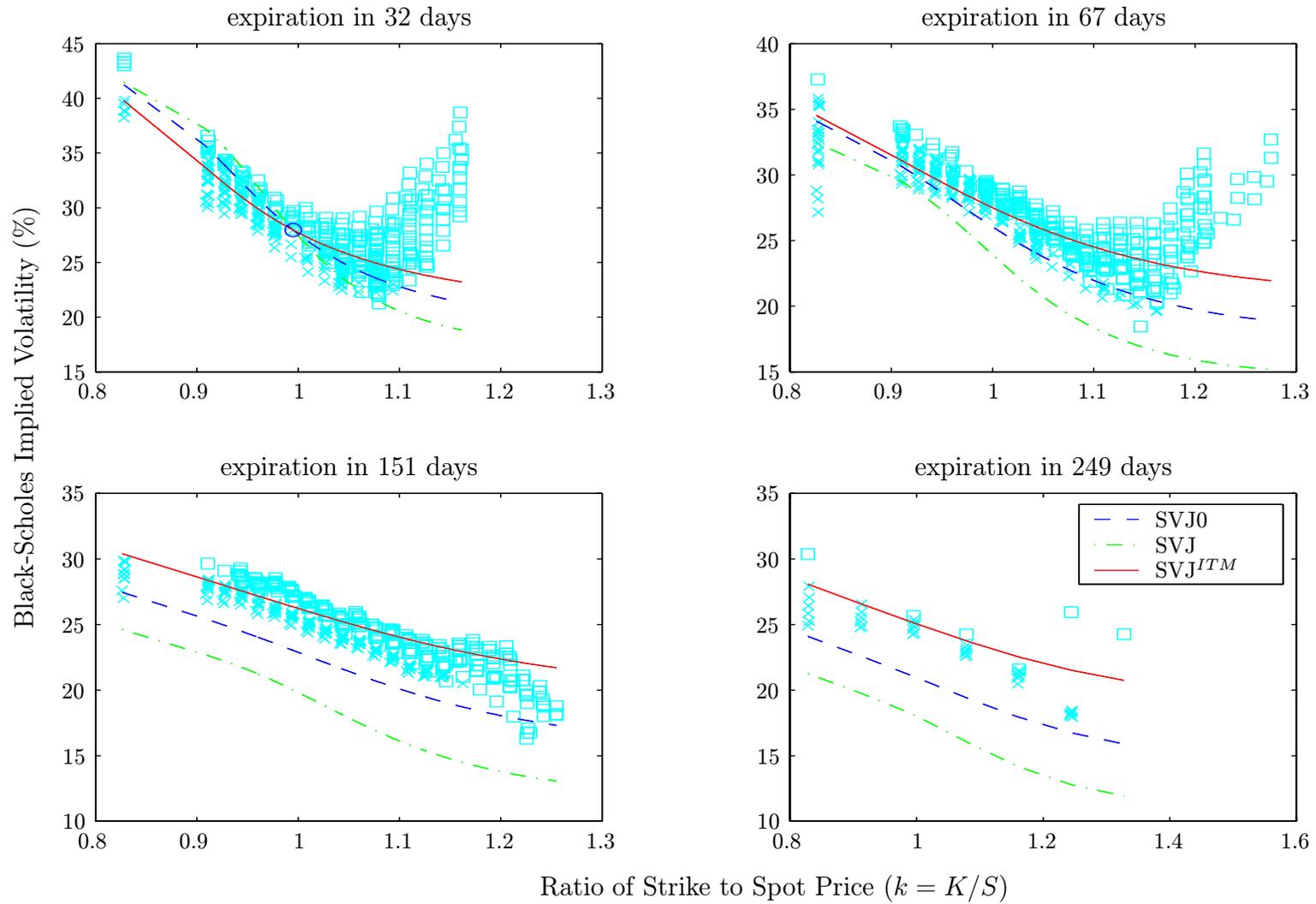


Figure 4: Smiles curves on a “high-volatility” day. All observations are observed between 10:00am to 11:00am on October 16, 1990. The call options are marked by ‘×,’ and the put options by ‘□’.

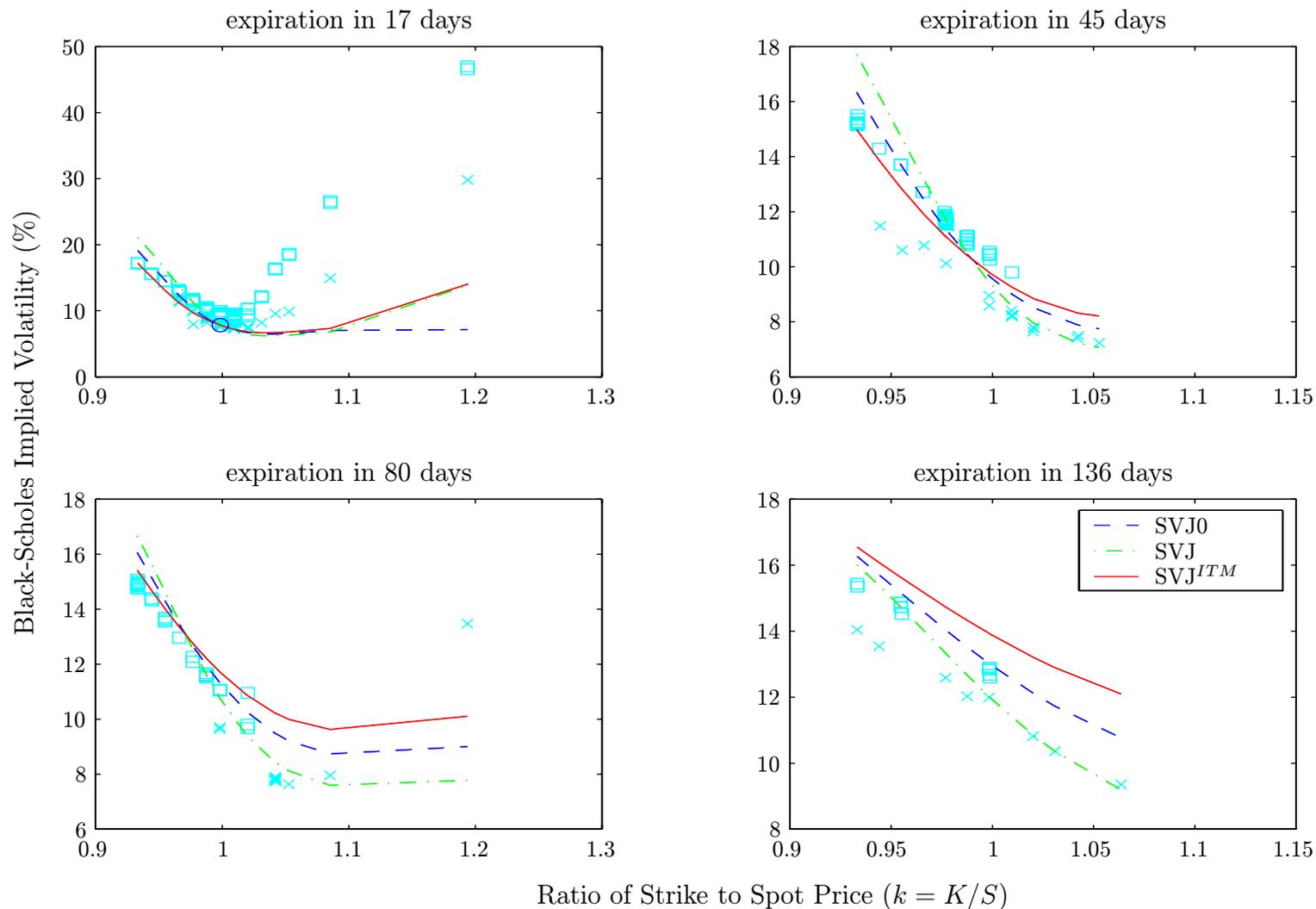


Figure 5: Smiles curves on the least volatile day of the sample. All observations are observed between 10:00am to 11:00am on August 3, 1994. The call options are marked by 'x,' and the put options by '□'.

Appendices

A 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.³⁶

A.1 Stationarity Assumption for Contract Variables

Figure 6 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.³⁷ 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.

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.³⁸ 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 A.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 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.

³⁶For 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.

³⁷To 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.

³⁸The “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 progresses. See Gallant and White [1988] and references therein.

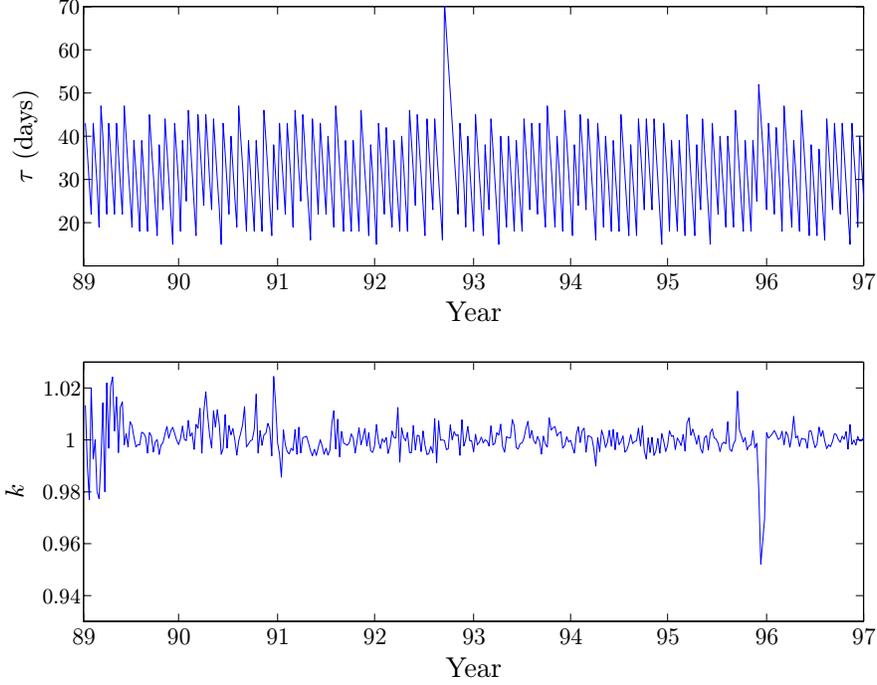


Figure 6: Time series of contract variables: time-to-expiration τ and strike-to-spot ratios k .

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.

A.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 $\nu : \mathbb{R}_+ \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is defined by

$$\nu(v, \vartheta, r, q, \tau, k) = g(f(v, \vartheta_0, r, q, \tau, k), \vartheta, r, q, \tau, k), \quad (\text{A.1})$$

where g is defined by (3.4), 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 $X_n = [y_{(n, n_y)}, V_{(n, n_v)}, r_{(n, n_v)}, q_{(n, n_v)}, k_{(n, n_v)}]$ 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

$$H(X_n, \vartheta, Y_n) = h(y_{(n, n_y)}, \nu(V_{(n, n_v)}, \vartheta, r_{(n, n_v)}, q_{(n, n_v)}, \tau_{(n, n_v)}, k_{(n, n_v)}), \vartheta), \quad (\text{A.2})$$

where $r_{(n,n_v)} = [r_n, r_{n-1}, \dots, r_{n-n_v+1}]$, 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 A.1 (Time Stationarity of Y) $\{Y_n\}$ has finitely many outcomes, denoted $\{1, 2, \dots, I\}$. For each outcome i and each positive integer N , let $A_N^{(i)} = \{n \leq N : Y_n = i\}$ 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

$$\lim_N \frac{\#A_N^{(i)}}{N} = w_i \quad a.s., \quad (\text{A.3})$$

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

For a proof of the geometric ergodicity of the state vector $\{y_n, V_n, r_n, q_n\}$, see Appendix B.9 of Pan [2000]. Assuming further that $\{k_n\}$ and $\{y_n, V_n, r_n, q_n\}$ are jointly geometrically ergodic, we know that $X_n = [y_{(n,n_y)}, V_{(n,n_v)}, r_{(n,n_v)}, q_{(n,n_v)}, k_{(n,n_v)}]$ 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 A.2 below then follows from Glynn [1999], under Assumption A.1 and the additional assumption of independence between X and Y . This independence assumption is trivially satisfied in our setting because Y is deterministic.³⁹

Assumption A.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

$$\sup_{\vartheta \in \Theta} |G_N^{(i)}(\vartheta) - G_\infty^{(i)}(\vartheta)| \rightarrow 0 \quad a.s.. \quad (\text{A.4})$$

Given the pointwise-SLLN portion of Assumption A.2, in order to establish the uniform SLLN of Assumption A.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 to establishing the strong consistency of $\{\hat{\vartheta}_N\}$. A proof can be found in Appendix B.8 in Pan [2000].

Proposition A.1 (USLLN of $H(X, \vartheta, Y)$) Under Assumptions A.1 and A.2, for each ϑ , $G_\infty(\vartheta) = \lim_N G_N(\vartheta)$ exists, and

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

where $G_N(\vartheta)$, defined by (3.6), is the sample moment of the observation function.

³⁹Independent-sampling strong laws for more general processes can be found in Glynn and Sigman [1998].

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

Assumption A.3 (Convergence of Weighting Matrices) $\mathcal{W}_N \rightarrow \mathcal{W}_0$ almost surely for some constant symmetric positive-definite matrix \mathcal{W}_0 .

Under Assumption A.3 and the conditions of Proposition A.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 $C : \Theta \rightarrow \mathbb{R}$ 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 A.1, and the fact that $V_n^{\vartheta_0} = V_n^\Delta$.

Assumption A.4 (Uniqueness of Minimizer) $C(\vartheta_0) \neq C(\vartheta)$, $\vartheta \in \Theta$, $\vartheta \neq \vartheta_0$.

Theorem A.1 (Strong Consistency) Under Assumptions A.1–A.4, the IS-GMM $\{\vartheta_N\}$ estimator converges to ϑ_0 almost surely as $N \rightarrow \infty$.

Given the Uniform SLLN (Proposition A.1), the proof is standard and omitted. (See, for example, the proof of Theorem 3.3 in Gallant and White [1988].)

A.2.1 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 A.5 (CLT) $\sqrt{N}G_N(\vartheta_0)$ converges in distribution as $N \rightarrow \infty$ 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 $\|h(y_{(n, n_y)}, V_{(n, n_v)})\|^{2+\delta}$, for some $\delta > 0$, over the stationary distribution of $(y_{(n, n_y)}, V_{(n, n_v)})$. (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

$$d(\vartheta, X_n, Y_n) = \frac{\partial}{\partial \vartheta} h(y_{(n, n_y)}, V_{(n, n_v)}^\vartheta, \vartheta) + \sum_{i=n-n_v+1}^n \frac{\partial}{\partial v_i} h(y_{(n, n_y)}, V_{(n, n_v)}^\vartheta, \vartheta) g_\vartheta(c_i, \vartheta, r_i^\Delta, q_i^\Delta, \tau_i, k_i), \quad (\text{A.5})$$

where $g_\vartheta(c, \vartheta, r, q, \tau, k) = \partial g(c, \vartheta, r, q, \tau, k) / \partial \vartheta$, with g defined by (3.4), and where $c_i = f(V_i^\Delta, \vartheta_0, r_i^\Delta, q_i^\Delta, \tau_i, k_i)$. The first term on the right-hand side of (A.5) 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 $V_i^\vartheta = g(c_i, \vartheta, r_i^\Delta, q_i^\Delta, \tau_i, k_i)$ on ϑ , for $i \in \{n - n_v + 1, \dots, n\}$. 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 A.6 (Convergence of “Jacobian Estimator”) *For some constant $(n_h \times n_\vartheta)$ matrix d_0 of rank n_ϑ : (i) $\frac{1}{N} \sum_{n \leq N} d(\vartheta_0, X_n, Y_n)$ converges in probability as $N \rightarrow \infty$ to d_0 . (ii) For any $\{\vartheta_n\}$ converging in probability as $n \rightarrow \infty$ to ϑ_0 , $\frac{1}{N} \sum_{n \leq N} d(\vartheta_n, X_n, Y_n)$ converges in probability as $N \rightarrow \infty$ to d_0 .*

Part (i) of Assumption A.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 A.2 (Asymptotic Normality) *Under Assumptions A.1–A.6, $\sqrt{N}(\vartheta_N - \vartheta_0)$ converges in distribution as $N \rightarrow \infty$ to a normal random vector with mean zero and covariance matrix*

$$\Lambda = (d_0^\top \mathcal{W}_0 d_0)^{-1} d_0^\top \mathcal{W}_0 \Sigma_0 \mathcal{W}_0 d_0 (d_0^\top \mathcal{W}_0 d_0)^{-1}. \quad (\text{A.6})$$

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 $\Lambda = d_0^{-1} \Sigma_0 (d_0^\top)^{-1}$.

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}^π and \mathcal{B}^π are local martingales. Now let $C = (C^{(1)}, \dots, C^{(n)})$ denote the process process of any n options, priced by the state-price density π , as in (2.8). 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 $X = (\mathcal{S}, \mathcal{B}, C^{(1)}, \dots, C^{(n)})$, 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.⁴⁰

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

⁴⁰A 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 $\theta_T \cdot X_T > 0$.

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 deflator. 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 $\theta_t \cdot X_t \geq 0$ 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 $\xi_t = \pi_t \exp\left(\int_0^t r_s ds\right)$ define in Section 2.3 to be a martingale.

Proposition C.1 (i) A sufficient condition for ξ to be a martingale is that

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

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

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

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 $E(\exp(U_i^\pi) - 1) = 0$, we have, for $0 \leq t \leq s \leq T$,

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

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,

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

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 that⁴¹

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

⁴¹See, for example, Duffie, Pan, and Singleton [1999] for details.

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

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

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 $\{\omega(t) : 0 \leq t \leq T\}$ is bounded.

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

The coefficients B and A of (3.8) are defined by

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

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

The coefficients α_v and β_v of (2.9) are defined by

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

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

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

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

Next, α_q and β_q are defined by

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

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 (2.11) 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 (2.11), assuming that the transform $\psi(c, v, r, q, \tau)$ defined by (2.9) 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,

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

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

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

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 $I_1(u) = \text{Im} \left(\tilde{\psi}_1(u) \exp(-iu\bar{x}) \right)$ denote the integrand, we approximate by

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

where $\lfloor x \rfloor$ is an integer such that $\lfloor x \rfloor - 1 < x \leq \lfloor x \rfloor$. 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

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

We can select a step size Δu_1 such that

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

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 , $|I_1(u)| \leq \exp(-uA_1 + A_0)$ where $A_1 = (v + \bar{v}^* \kappa^* \tau) \sqrt{1 - \rho^2} / \sigma_v$, and $A_0 = (v + \bar{v}^* \kappa^* \tau) (\kappa^* - \sigma_v \rho) / \sigma_v^2 + \ln(4(1 - \rho^2)) \kappa^* \bar{v}^* / \sigma_v^2$. For the desired accuracy δ , we can therefore choose U_1 such that

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

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_{X_1} and variance $\sigma_{X_1}^2$ of \tilde{X}_1 are finite, the tail probabilities can be controlled by Chebyshev's inequality:

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

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

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

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 $E \left[\exp(u\tilde{X}_1) \right] = \tilde{\psi}_1(-iu)$, 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 $E_t(y_{t+\Delta}^i V_{t+\Delta}^j)$, 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.⁴² The joint moments of y and V can be calculated recursively by⁴³

$$\begin{aligned} E_t(y_{t+\Delta}^0 V_{t+\Delta}^m) &= \sum_{j=0}^{m-1} C_{m-1}^j E_t(y_{t+\Delta}^0 V_{t+\Delta}^j) p_{y,v}^{(0,m-j)}(V_t), \quad m \geq 1, \\ E_t(y_{t+\Delta}^n V_{t+\Delta}^m) &= \sum_{i=0}^{n-1} \sum_{j=0}^m C_{n-1}^i C_m^j E_t(y_{t+\Delta}^i V_{t+\Delta}^j) p_{y,v}^{(n-i,m-j)}(V_t), \quad n \geq 1, \quad m \geq 0, \end{aligned} \quad (\text{F.1})$$

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

$$p_{y,v}^{(i,j)}(V_t) = A_{y,v}^{(i,j)} + B_{y,v}^{(i,j)} V_t, \quad (\text{F.2})$$

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

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

$$\begin{aligned} B_{y,v}^{(0,m)} &= \frac{m}{2} f_0 \sigma_v^2 B_{y,v}^{(0,m-1)}, \quad m \geq 2, \\ B_{y,v}^{(n,0)} &= \lambda_1 J_n f_0 - \frac{1}{2} \sum_{i=1}^{n-1} C_n^i B_{y,v}^{(i,0)} f_{n-i}, \quad n \geq 3, \\ B_{y,v}^{(n,1)} &= \frac{1}{2} \kappa f_0 f_n + \frac{1}{2} \sigma_v^2 f_0 B_{y,v}^{(n,0)} - \frac{1}{2} \sum_{i=1}^{n-1} C_n^i f_i B_{y,v}^{(n-i,1)}, \quad n \geq 2, \\ B_{y,v}^{(n,m)} &= \frac{m}{2} \sigma_v^2 f_0 B_{y,v}^{(n,m-1)} - \frac{1}{2} \sum_{i=1}^n C_n^i f_i B_{y,v}^{(n-i,m)}, \quad n \geq 1, \quad m \geq 2, \end{aligned}$$

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

$$\begin{aligned} g_0 &= 2, \quad g_n = 2\gamma_n + \frac{1}{f_0} \sum_{i=1}^{n-1} \Gamma_{n-i} g_i, \quad n \geq 1, \\ f_0 &= \frac{1 - \exp(-\kappa\Delta)}{\kappa}, \quad f_1 = g_1 - (\kappa\gamma_1 + \sigma_v \rho) f_0, \quad f_n = g_n - \kappa\gamma_n f_0, \quad n \geq 2, \end{aligned}$$

⁴²This 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 .

⁴³For 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 $\lambda_1 = 0$.

with

$$\begin{aligned}\gamma_1 &= -\left(\frac{\sigma_v}{\kappa}\right)^2 \left(\rho \frac{\kappa}{\sigma_v} + \eta^s - \frac{1}{2} + \lambda_1(J_1 - \mu^*)\right), \quad \gamma_2 = -\gamma_1^2 - \left(\frac{\sigma_v}{\kappa}\right)^2 (1 - \rho^2 + \lambda_1 J_2), \\ \gamma_n &= -\sum_{i=1}^{n-1} \gamma_i \gamma_{n-i} C_{n-1}^i - \left(\frac{\sigma_v}{\kappa}\right)^2 \lambda_1 J_n, \quad n \geq 3 \\ \Gamma_0 &= \frac{\exp(-\kappa \Delta)}{\kappa}, \quad \Gamma_n = -\kappa \Delta \sum_{i=0}^{n-1} C_{n-1}^i \gamma_{n-i} \Gamma_i, \quad n \geq 1.\end{aligned}$$

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

$$\begin{aligned}A_{y,v}^{(0,n)} &= \frac{n-1}{2} \sigma_v^2 f_0 A_{y,v}^{(0,n-1)}, \quad n \geq 2, \\ A_{y,v}^{(n,0)} &= \lambda_0 J_n^0 \Delta - \frac{\kappa \bar{v}}{\sigma_v^2} \left(\kappa \gamma_n \Delta + \hat{f}_n - \hat{g}_n\right), \quad n \geq 2, \\ A_{y,v}^{(n,1)} &= -\frac{\kappa \bar{v}}{2} f_0 f_n - \frac{1}{2} \sum_{i=1}^{n-1} C_n^i f_i A_{y,v}^{(n-i,1)}, \quad n \geq 1, \\ A_{y,v}^{(n,m)} &= -\kappa \bar{v} \sigma_v^{2(m-1)} m! f_n \left(\frac{f_0}{2}\right)^m - \frac{1}{2} \sum_{i=1}^{n-1} \hat{C}_n^i(m) f_i A_{y,v}^{(n-i,m)}, \quad n \geq 1, m \geq 2,\end{aligned}$$

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 A.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,

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

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

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

where $G_N = (N)^{-1} \sum_n \mathcal{H}_n$ is the sample analogue of the “optimal” moments. We know that $\partial \mathcal{G}_N(\bar{\vartheta}_N)/\partial \vartheta$ converges to a constant full-rank matrix d_0 in probability, under Assumption A.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

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

where $\stackrel{a}{\approx}$ 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 Estimation of Interest Rates and Dividend Yields

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 given q_{n-1} 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 $\{q_n\}$. The observed T -year LIBOR rates $\{R_n\}$ (converted to continuous compounding rates) can be expressed in terms of r_n by (Cox, Ingersoll, and Ross

[1985])

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

where θ_r^0 denotes the true parameter vector, and where α_r and β_r are as defined in (D.3). The one-period conditional density $p^R(\cdot | R_{n-1}; \theta_r)$ of R_n given R_{n-1} is therefore given by

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

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

Table 6: 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 6. 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.

References

- Ait-Sahalia, Y. and A. Lo (2000). Nonparametric Risk Management and Implied Risk Aversion. *Journal of Econometrics* **94**, 9–51.
- Ait-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 (2000). Spanning and Derivative-Security Valuation. *Journal of Financial Economics* **55**, 205–238.
- Bates, D. (2000). Post-'87 Crash fears in S&P 500 Futures Options. *Journal of Econometrics* **94**, 181–238.
- Benzoni, L. (1998). Pricing Options under Stochastic Volatility: An Econometric Analysis. Working Paper, J.L. Kellogg Graduate School of Management, Northwestern University.
- Bernardo, A. and O. Ledoit (2000). Gain, Loss, and Asset Pricing. *Journal of Political Economy* **108**, 144–172.
- 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.
- Canina, L. and S. Figlewski (1993). The Informational Content of Implied Volatility. *The Review of Financial Studies* **6**, 659–681.
- Chacko, G. and L. Viceira (1999). Spectral GMM Estimation of Continuous-Time Process. Working Paper, Harvard Business School.
- Chernov, M., R. Gallant, E. Ghysels, and G. Tauchen (1999). Estimating Jump-Diffusions with a Panel of Options and Returns. Working Paper, Department of Finance, Pennsylvania State University.

- Chernov, M. and E. Ghysels (2000). 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* **56**.
- Christensen, B. and N. Prabhala (1998). The Relation between Implied and Realized Volatility. *Journal of Financial Economics* **50**, 125–150.
- Cochrane, J. and J. Saá-Requejo (2000). Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets.
- 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.
- Cox, J. C. and S. Ross (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics* **3**, 145–166.
- Danielsson, J. (1994). Stochastic Volatility in Asset Prices Estimation with Simulated Maximum Likelihood. *Journal of Econometrics* **64**, 375–400.
- 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.
- Davies, R. (1973). Numerical Inversion of a Characteristic Function. *Biometrika* **60**, 415–417.
- Day, T. E. and C. M. Lewis (1992). Stock Market Volatility and the Information Content of Stock Index Options. *Journal of Econometrics* **52**, 267–287.
- 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 Uncertainty. *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.

- Figlewski, S. (1997). Forecasting Volatility. *Financial Markets, Institution & Instruments* **6**, 1–88.
- Fouque, J.-P., G. Papanicolaou, and K. R. Sircar (2000). Mean-Reversion of S&P 500 Volatility. Working Paper, Department of Mathematics, Stanford University.
- 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.
- Gallant, R., D. Hsieh, and G. Tauchen (1997). Estimation of Stochastic Volatility Models with Diagnostics. *Journal of Econometrics* **81**, 159–192.
- 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 Currency 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.
- Harvey, A. C., E. Ruiz, and N. G. Shephard (1994). Multivariate Stochastic Variance Models. *The Review of Economic Studies* **61**, 247–264.
- Harvey, C. R. and A. Siddique (2000). Conditional Skewness in Asset Pricing Tests. *Journal of Finance*, forthcoming.
- 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.

- 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. (1999). Recovering Risk Aversion from Options Prices and Realized Returns. *Review of Financial Studies* **13**, 1611–1632.
- Jackwerth, J. C. and M. Rubinstein (1996). Recovering Probability Distributions from Option Prices. *Journal of Finance* **51**, 1611–1631.
- Jacquier, E., N. Polson, and P. Rossi (1994). Bayesian Analysis of Stochastic Volatility Models (with Discussion). *Journal of Business & Economic Statistics* **12**, 371–417.
- 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.
- Jorion, P. (1989). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies* **4**, 427–445.
- 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.
- Kraus, A. and R. Litzenberger (1976). Skewness Preference and the Valuation of Risk Assets.
- Lamoureux, C. and W. Lastrapes (1993). Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatility. *The Review of Financial Studies* **6**, 293–326.
- Liu, J. (1997). Generalized Method of Moments Estimation of Affine Diffusion Processes. Working Paper, Graduate School of Business, Stanford University.
- Longstaff, F. (1995). Option Pricing and the Martingale Restriction. *Review of Financial Studies* **8**, 1091–1124.
- Lucas, R. E. (1978). Asset Prices in an Exchange Economy. *Econometrica* **46**, 1429–1445.
- Melino, A. (1994). Estimation of Continuous-Time Models in Finance. In C. Sims (Ed.), *Advances in Econometrics*. Cambridge University Press.
- Melino, A. and S. Turnbull (1990). Pricing Foreign Currency Options with Stochastic Volatility. *Journal of Econometrics* **45**, 239–265.
- Merton, R. (1976). Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* **3**, 125–44.
- 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.
- Pan, J. (2000). *Jump-Diffusion Models of Asset Prices: Theory and Empirical Evidence*. Ph. D. thesis, Graduate School of Business, Stanford University.
- Pastorello, S., E. Renault, and N. Touzi (1996). Statistical Inference for Random Variance Option Pricing. Working Paper, Dipartimento di Scienze Economiche, Università di Bologna, Italy.
- Patilea, V. and E. Renault (1997). Continuously Updated Extremum Estimators. Working Paper, Institut de Statistique, Université Catholique de Louvain.
- Pham, H. and N. Touzi (1996). Equilibrium State Prices in a Stochastic Volatility Model. *Mathematical Finance* **6**, 215–236.
- Poteshman, A. M. (1998). Estimating a Genral Stochastic Variance Model from Options Prices. Working Paper, Department of Finance, University of Illinois at Urbana-Cahmpaign.
- Poteshman, A. M. (2000). Forecasting Future Variance from Option Prices. Working Paper, Department of Finance, University of Illinois at Urbana-Cahmpaign.
- 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. (1997). Econometric Models of Option Pricing Errors. In D. Kreps and K. Wallis (Eds.), *Advances in Economics and Econometrics: Theory and Applications*, Volume 3. Cambridge University Press.
- 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.
- Rosenberg, J. and R. Engle (1999). Empirical Pricing Kernels. Working Paper, Stern School of Business, New York University.
- Rubinstein, M. (1973). The Fundamental Theorem of Parameter-Preference Security Valuation. *Journal of Financial and Quantitative Analysis* **8**, 61–69.
- Scott, L. (1997). Pricing Stock Options in a Jump-diffusion Model with Stochastic Volatility and Interest Rates: Application of Fourier Inversion Methods. *Mathematical Finance* **7**, 345–358.
- Singleton, K. (1999). *Lecture Notes on Econometric Analysis of Dynamic Asset Pricing Models*. Graduate School of Business, Stanford University.
- Singleton, K. (2000). Estimation of Affine Diffusion Models Based on the Empirical Characteristic Function. *Journal of Econometrics*, *Forthcoming*.

- 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.
- Taylor, S. J. (1986). *Modelling Financial Time Series*. New York: John Wiley.