# The jump-risk premia implicit in options: evidence from an integrated time-series study[☆]

## Jun Pan[*]

*Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, USA*

## 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. © 2002 Elsevier Science S.A. All rights reserved.

*JEL classification:* G12

*Keywords:* Option pricing; Stochastic volatility; Jump-risk premium; Implied-state generalized method of moments; Volatility "smiles" and "smirks"

## 1. Introduction

It has been widely documented that stock returns exhibit both stochastic volatility and jumps. The importance of such risk factors arises not only from time-series studies of stock prices,[1] but also from cross-sectional studies of stock options (Bakshi et al., 1997; Bates, 2000). This brings us to ask an important question that was left unanswered by past studies: To what extent are such risk factors, namely volatility and jump risks, priced in the financial market? In particular, what is the market price of the jump risk? Is the jump risk priced differently from the diffusive risk? Answers to these questions have a direct impact on investors' decision-making, and could also shed some light on how investors react to various types of uncertainty.

In this paper, we address these issues by focusing on the risk premia implicit in the S&P 500 index options. Although there is a rich body of empirical studies on options, our understanding of the risk premia in options is still limited. In particular, the role of jump-risk premia in option pricing has not been examined to date.[2] In light of this, we adopt the Bates (2000) model, which extends the stochastic volatility model of Heston (1993) by incorporating state-dependent price jumps. Under such a setting, the S&P 500 index returns are affected by three distinctively different risk factors: (1) the diffusive price shocks, (2) the price jumps, and (3) the diffusive volatility shocks. The dynamic properties of such risk factors and, more important for our purpose, the market prices of such risk factors determine how options on the S&P 500 index are priced. To facilitate our analysis of how various risk factors are priced, we introduce a parametric pricing kernel to price all three risk factors, including the volatility risk and the jump risk. An important feature of the jump-risk premium considered here is that it is allowed to depend on the market volatility.

To estimate the price dynamics simultaneously with the parametric pricing kernel, we adopt an integrated approach to the time-series data on the S&P 500 index and options. Such an integrated approach has long been advocated in the literature, but its implementation has only been recent. For example, in a non-parametric setting, Aït-Sahalia et al. (2001) compare the risk-neutral densities estimated separately from spot prices and option prices. In a parametric setting, Chernov and Ghysels (2000) use joint time-series data to estimate the Heston (1993) model.

Our integrated approach differs from those adopted by earlier studies in that we device an "implied-state" generalized method of moments (IS-GMM) estimation strategy to take full advantage of the analytical tractability of the Bates (2000) model. For a given set $\vartheta$ of model parameters, we proxy for the unobserved volatility, $V_t$, with an option-implied volatility, $V_t^{\vartheta}$, inverted from the time-$t$ spot price, $S_t$, and a near-the-money short-dated option price, $C_t$, using the

---

[1] See, for example, Jorion (1989), Andersen et al. (1998), Eraker et al. (2000), Chernov et al. (1999, 2000), and the references therein.

[2] In pure-diffusion settings, the role of volatility-risk premia has been examined by Guo (1998), Benzoni (1998), Chernov and Ghysels (2000), Poteshman (1998), and Bakshi and Kapadia (2001). These studies, however, do not address the issue of jump-risk premia nor do they provide insight on the relative importance of the premia for jump and volatility risks.

model-implied option-pricing formula. Access to the option-implied stochastic volatility, $V^{\vartheta}$, allows us to explore the joint distribution of spot and option prices by focusing directly on the dynamic structure of the state variables: the stock price $S$ and the stochastic volatility $V$. This approach is particularly attractive in our setting, as the conditional moment-generating function of $(S, V)$ is explicitly known, while the conditional distribution of $(S, C)$ could be complicated because of the nonlinearity of option pricing.

Under such an integrated approach, the role of risk premia arises naturally as we attempt to reconcile the dynamics implied by the joint time series $\{S_t, C_t\}$ of the S&P 500 index and the near-the-money short-dated option prices. For example, for a model in which neither jump risk nor volatility risk is priced, strong inconsistency arises between the level of volatility observed in the spot market and that implied, through the model, by the options market. In particular, the option-implied volatility is too high to be rationalized by the ex-post realized volatility observed in the spot market. Similar findings are reported by Jackwerth and Rubinstein (1996) in the setting of the Black and Scholes (1973) model. One natural explanation for this bias is that option prices contain a risk-premium component associated with additional risk factors such as stochastic volatility and jumps. Allowing for a volatility-risk premium, and fitting the stochastic-volatility model of Heston (1993) to the joint time-series data $\{S_t, C_t\}$, we find a significant volatility-risk premium, as well as an improvement in goodness-of-fit. Overall, however, 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 overpriced long-dated options.

Given the poor performance of the model with volatility-risk premia, an alternative approach is to introduce jump and jump-risk premium to the Heston (1993) model. Indeed, fitting the Bates (2000) model (with a state-dependent jump-risk premium) to the joint time-series data $\{S_t, C_t\}$, we find a significant premium for jump risk. In contrast to the case with volatility-risk premium, the model is not rejected by the joint time-series data, and the estimated level of jump-risk premium does not result in any distortion of the long-dated option prices. Finally, when allowing both types of risk premia to reconcile simultaneously the spot and option dynamics, we find that the state-dependent jump-risk premium dominates by far the volatility-risk premium.

In addition to the joint time-series data $\{S_t, C_t\}$ that have so far been the focus of our model estimation and empirical analyses, we also observe on each date $t$ a cross-section of options with varying degrees of moneyness and maturities. This rich set of options data provides a challenging test of our empirical results. Using the model parameters estimated exclusively from the joint time series $\{S_t, C_t\}$ of one spot and one option, we find that the Bates (2000) model explains the cross-sectional options data surprisingly well, capturing, in particular, the changes over time of the "smirk" patterns that are exhibited in options with different moneyness. Again, the key component here is the state-dependent jump-risk premium, under which small negative jumps in the actual dynamics are perceived to be significantly more negative in the "risk-neutral" measure. The importance of such negative "risk-neutral" jumps

in capturing the "smirk" patterns has also been documented by Bakshi et al. (1997) and Bates (2000). We reach to the same conclusion, however, in different ways. While their model estimates (the "risk-neutral" part) are obtained by fitting directly to the entire cross-sectional options, ours are estimated using the time series $\{S_t, C_t\}$ of only one spot and one option. While their empirical implication stays at the level of the "risk-neutral" dynamics, ours goes one step further by noting that volatility "smirks" are primarily due to investors' fear of large adverse price jumps. Moreover, given that the jump-risk premium is, in fact, estimated using near-the-money option prices, our results show that such fear of jump risk is reflected not only in deep out-of-the-money (OTM) puts, but in near-the-money options as well.

To understand the pricing kernel that links the two markets, the spot and options data can be exploited in alternative ways. Under a consumption-based asset pricing setting, Jackwerth (2000) and Aït-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 impacts 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 distinctly different risk attitudes toward jump risk and diffusive risk.

Finally, it should be noted that our empirical results build on the premise that the options market is fully integrated with the spot market, sharing the same price dynamics and the same market prices of risks. If the options market is somehow segmented from the spot market because of some option-specific factors such as liquidity, it then becomes impossible to identify the parametric pricing kernel in our current setting. In a related issue, Jackwerth and Rubinstein (1996) raise the possibility of a "peso" component (extreme and rare events) in option prices, which is found to be important by Aït-Sahalia et al. (2001). Given that jumps are inherently infrequent, and our sample period is relatively short, it is empirically hard to separate the "peso" hypothesis from that of jump-risk premia. In principle, however, these two hypotheses are fundamentally different. The "peso" explanation implies that OTM put options are priced with premia because of the potential occurrence of extreme events, which have not yet been materialized. By contrast, the risk-premia explanation emphasizes investors' aversion to such extreme events – options (especially OTM put options) are priced with premia not only because of the likelihood and magnitude of such rare events, but also because of investors' aversion to such events.

The remainder of this paper is organized as follows. Section 2 specifies the Bates (2000) model, the parametric pricing kernel, and the option-pricing formula. Section 3 outlines the integrated approach adopted in this paper and provides details

on the estimation strategy. Section 4 describes the data. Section 5 summarizes the empirical findings. Section 6 concludes the paper. Technical details are provided in appendices.

## 2. The model

We adopt the Bates (2000) model to characterize the stock return dynamics. As summarized in Section 2.1, this model introduces three sources of uncertainty to the underlying price dynamics: (1) diffusive return shocks, (2) volatility shocks, and (3) jump risks.[3] The market prices of these risk factors are characterized in Section 2.2. A brief description of option pricing under this dynamic setting is presented in Section 2.3. Details about the state-price density that gives rise to the market prices of risks are given in Appendix A, and additional information about option pricing is provided in Appendix B.

### 2.1. The data-generating process

We fix a probability space $(\Omega, \mathscr{F}, P)$ and an information filtration $(\mathscr{F}_t)$ satisfying the usual conditions (see, for example, Protter, 1990), and let $S$ be the ex-dividend price process of a stock that pays dividends at a stochastic proportional rate $q$. Adopting the model of Bates (2000), we assume the following data-generating process for the stock price $S$

$$\mathrm{d}S_t = \left[ r_t - q_t + \eta^s V_t + \lambda V_t \left( \mu - \mu^* \right) \right] S_t \, \mathrm{d}t + \sqrt{V_t} S_t \, \mathrm{d}W_t^{(1)}$$
$$+ \, \mathrm{d}Z_t - \mu S_t \lambda V_t \, \mathrm{d}t, \tag{2.1}$$

$$\mathrm{d}V_t = \kappa_v(\bar{v} - V_t) \, \mathrm{d}t + \sigma_v \sqrt{V_t} \left( \rho \, \mathrm{d}W_t^{(1)} + \sqrt{1 - \rho^2} \, \mathrm{d}W_t^{(2)} \right), \tag{2.2}$$

where $r$ is a stochastic interest-rate process, $W = [W^{(1)}, W^{(2)}]^\top$ is an adapted standard Brownian motion in $\mathbb{R}^2$, and $Z$ is a pure-jump process.

This model captures two important features of the stock return dynamics, namely stochastic volatility and price jumps, and still provides analytical tractability for option pricing and model estimation. First, stochastic volatility is modeled by the autonomous process $V$ defined by Eq. (2.2), which is a one-factor "square-root" process with constant long-run mean $\bar{v}$, mean-reversion rate $\kappa_v$, and volatility coefficient $\sigma_v$.[4] This volatility specification, introduced by Heston (1993), allows the "Brownian" shocks to price $S$ and volatility $V$ to be correlated with constant coefficient $\rho$, capturing an important stylized fact that stock returns are typically negatively correlated with changes in volatility (Black, 1976).

---

[3] To be more precise, this model involves five sources of uncertainty. As it becomes clear in Section 2.1, the two additional shocks are associated with the riskfree rate and the dividend yields.

[4] It should be noted that we call the variance $V$ the volatility, which is typically referred to as the standard deviation of returns. This change of terminology should not cause any confusion.

Second, this model captures price jumps via the pure-jump process $Z$, which contains two components: random jump-event times and random jump sizes. The jump-event times $\{\mathcal{T}_i : i \geqslant 1\}$ arrive with a state-dependent stochastic intensity process $\{\lambda V_t : t \geqslant 0\}$ for some non-negative constant $\lambda$.[5] Given the arrival of the $i$th jump event, the stock 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 $\mu_J$ and variance $\sigma_J^2$, independent of $W$, of inter-jump times, and of $U_j^s$ for $j \neq i$. Intuitively, the conditional probability at time $t$ of another jump before $t + \Delta t$ is, for some small $\Delta t$, approximately $\lambda V_t \Delta t$ and, conditional on a jump event, the mean relative jump size is $\mu = \mathrm{E}(\exp(U^s) - 1) = \exp(\mu_J + \sigma_J^2/2) - 1$. Combining the effects of random jump timing and sizes, the last term $\mu S_t \lambda V_t \, \mathrm{d}t$ in Eq. (2.1) compensates for the instantaneous change in expected stock returns introduced by the pure-jump process $Z$.

Given that the pure-diffusion model of Heston (1993) cannot explain the tail-fatness of the stock return distribution (Andersen et al., 1998), nor can it explain the "smirkiness" exhibited in the cross-sectional options data (Bakshi et al., 1997; Bates, 2000), the extension to include jumps is well motivated. It should be emphasized, however, that our main motivation is to study how such jump risks are priced and, in particular, their role in reconciling the spot and option dynamics. As will become evident in the next section, we choose the linear specification $\lambda V$ of jump-arrival intensity to allow for a state-dependent jump-risk premium, i.e., to allow for the possibility that when the market is more volatile, the jump-risk premium implicit in option prices becomes higher.[6]

Focusing on the drift component of the stock price dynamics, we see that the stock price appreciates with interest rate $r_t$, pays out dividend rate $q_t$, and appreciates with two risk-premium components: $\eta^s V_t$ and $\lambda V_t(\mu - \mu^*)$, which are associated with the premia for "Brownian" return risks and jump risks, respectively. We postpone a formal discussion on the risk-premium components to Section 2.2.

The short interest-rate process $r$ is of the type modeled by Cox et al. (1985). Specifically, $r$ and the dividend-rate process $q$ are defined by

$$\mathrm{d}r_t = \kappa_r(\bar{r} - r_t)\,\mathrm{d}t + \sigma_r\sqrt{r_t}\,\mathrm{d}W_t^{(r)},$$

$$\mathrm{d}q_t = \kappa_q(\bar{q} - q_t)\,\mathrm{d}t + \sigma_q\sqrt{q_t}\,\mathrm{d}W_t^{(q)}, \tag{2.3}$$

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 $r$

---

[5] This jump model is of the Cox-process type. Conditional on the path of $V$, jump arrivals are Poisson with time-varying intensity $\{\lambda V_t : t \geqslant 0\}$. See, for example, Brémaud (1981).

[6] It should be noted that to maintain a parsimonious model, we leave out the constant component in jump arrival intensity. Using stock return data alone, evidence in support of state-dependent jump arrival intensity is documented by Johannes et al. (1998), while the studies of Chernov et al. (1999) and Andersen et al. (1998) emphasize the importance of the constant component. For the purpose of reconciling spot and option dynamics, we find the state-dependent component to play a dominating role. The constraint of zero constant component will be formally tested in Section 5.1. In addition to the linear specification, the case of nonlinear dependency of the arrival intensity on $V$ seems interesting, so is the case of state-dependent jump sizes. For analytical tractability, however, these specifications are not considered here.

and $q$ are autonomous one-factor square-root processes with constant long-run means ($\bar{r}$ and $\bar{q}$), mean-reversion rates ($\kappa_r$ and $\kappa_q$), and volatility coefficients ($\sigma_r$ and $\sigma_q$).

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. Our formulation of $r$ and $q$, however, precludes possible correlation between the two, as well as more plausible and richer dynamics for the short-rate process. But for the short-dated options used to fit our model, the particular stochastic nature of interest rates $r$ and dividend yields $q$ plays a relatively minor role.

Finally, one limitation of our volatility specification is that it does not allow volatility to jump, a feature that is found to be important in stock returns (Eraker et al., 2000). In an issue that is related, Jones (1999) points out that, compared with the constant elastic variance (CEV) model, the square-root specification does not allow volatility to increase fast enough. The severity of these limitations will be investigated in Section 5.4 by diagnostic tests on the volatility dynamics. Their implications on our understanding of the jump-risk premia will also be discussed. Overall, to maintain a parsimonious model is the main reason why volatility jumps are not considered here, and the CEV specification is not incorporated for analytical tractability.[7]

## 2.2. The market prices of risks

In contrast to the complete market setting of Black and Scholes (1973), the additional sources of uncertainty, in particular the random jump sizes, introduced in our setting make the market incomplete with respect to the riskfree bank account, the underlying stock, and the finite number of options contracts. Consequently, the state-price density (or pricing kernel) is not unique. Our approach is to focus on a candidate pricing kernel that prices the three important sources of risks: diffusive price shocks, jump risks, and volatility shocks.[8] Given their mild effects on short-dated option prices, the interest-rate and dividend-rate risks are not priced in this paper.

For exposition purpose, we present in this section the "risk-neutral" price dynamics defined by our candidate pricing kernel, leaving a detailed description of the pricing kernel in Appendix A. Letting $Q$ be the equivalent martingale measure

---

[7] Alternative models also include the log-normal model of Hull and White (1987). In fact, using high-frequency stock return data, Andersen et al. (2001) suggest that volatility is best described as a log-normal process. One disadvantage of such a model, however, is that once the negative correlation between stock returns and volatility is incorporated, option pricing becomes intractable. Moreover, Benzoni (1998) shows that for the purpose of option pricing, there is no qualitative difference between the log-normal model and the square-root model.

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

associated with our candidate pricing kernel, we assign the market prices of risks so that both $r$ and $q$ have the same joint distribution under $Q$ as under the data-generating measure $P$, and the dynamics of $(S, V)$ under $Q$ are of the form

$$dS_t = [r_t - q_t]S_t \, dt + \sqrt{V_t}S_t \, dW_t^{(1)}(Q) + dZ_t^Q - \mu^* S_t \lambda V_t \, dt, \tag{2.4}$$

$$dV_t = [\kappa_v(\bar{v} - V_t) + \eta^v V_t] \, dt + \sigma_v \sqrt{V_t}\Big(\rho \, dW_t^{(1)}(Q) + \sqrt{1 - \rho^2} \, dW_t^{(2)}(Q)\Big), \tag{2.5}$$

where $W(Q) = [W^{(1)}(Q), W^{(2)}(Q)]$ is a standard Brownian motion under $Q$. [See Appendix A for a formal definition of $W(Q)$.] The pure-jump process $Z^Q$ has a distribution under $Q$ that is identical to the distribution of $Z$ under $P$ defined in Eq. (2.1), except that under $Q$, the jump arrival intensity is $\{\lambda V_t : t \geqslant 0\}$ for some non-negative constant $\lambda$, and the jump amplitudes $U_i^s$ is normally distributed with $Q$-mean $\mu_J^*$ and $Q$-variance $\sigma_J^2$. In other words, under the risk-neutral measure, the conditional probability at time $t$ of another jump before $t + \Delta t$ is approximately $\lambda V_t \Delta t$ and, conditional on a jump event, the risk-neutral mean relative jump size is $\mu^* = E^Q(\exp(U^s) - 1) = \exp(\mu_J^* + \sigma_J^2/2) - 1$. Following the same discussion for the data-generating process, we see that the last term $\mu^* S_t \lambda V_t \, dt$ in Eq. (2.4) is the compensator for the pure-jump process $Z^Q$ under the risk-neutral measure. Consequently, the instantaneous risk-neutral expected rate of stock return is the short interest rate $r$ minus the dividend payout rate $q$.

Comparing the specification of the risk-neutral dynamics of $(S, V)$ with that of the data-generating process, one can obtain an intuitive understanding of how different risk factors are priced. Focusing first on the market prices of jump risks, we see that by allowing the risk-neutral mean relative jump size $\mu^*$ to be different from its data-generating counterpart $\mu$, we accommodate a premium for jump-size uncertainty. Similarly, a premium for jump-timing risk can be incorporated if we allow the coefficient $\lambda^*$ for the risk-neutral jump-arrival intensity to be different from its data-generating counterpart $\lambda$. In this paper, however, we concentrate mainly on the risk premium for jump-size uncertainty, while ignoring the risk premium for jump-timing uncertainty by supposing $\lambda^* = \lambda$. With this assumption, all jump risk premia will be artificially absorbed by the jump-size risk premium coefficient $\mu - \mu^*$. The time-$t$ expected excess stock return compensating for the jump-size uncertainty is $\lambda V_t(\mu - \mu^*)$.

We adopt this approach mainly out of empirical concern over our ability to separately identify the risk premia for jump timing and jump size uncertainties. For example, the arrival intensity of price jumps, as well as the mean relative jump size $\mu$, could be difficult to pin down using the S&P 500 index data under a GMM estimation approach. In Section 5.2, this constraint of $\lambda^* = \lambda$ will be relaxed to gauge the relative importance of premia for jump-timing and jump-size risks.

Premia for the "conventional" return risks ("Brownian" shocks) are parameterized by $\eta^s V_t$ for a constant coefficient $\eta^s$. 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 Eq. (2.5). For a positive coefficient $\eta^v$, 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 $\eta^v$.

Our specification of risk premia can also be relaxed. For example, the linear form of the volatility-risk premia $\eta^v V_t$ could be relaxed by introducing the polynomial form $\eta_0 + \eta_1 V_t + \eta_2 V_t^2 + \cdots + \eta_l V_t^l$, for some constant coefficients $\eta_0, \eta_1, \eta_2, \ldots, \eta_l$. Our specification rules out the possibility that $\eta_0 \neq 0$, because it could imply non-diminishing risk premia as the volatility approaches to zero. The quadratic term $\eta_2 V_t^2$ seems to be an interesting case, but is not examined in this paper for analytical tractability.

## 2.3. Option pricing

Let $\theta_r = [\kappa_r, \bar{r}, \sigma_r]^\top$, and $\theta_q = [\kappa_q, \bar{q}, \sigma_q]^\top$ denote the model parameters for the interest-rate process $r$ and the dividend-rate process $q$, respectively, and let $\vartheta$ denote the rest of the model parameters:

$$\vartheta = \left(\kappa_v, \bar{v}, \sigma_v, \rho, \lambda, \mu, \sigma_J, \eta^s, \eta^v, \mu^*\right). \tag{2.6}$$

Let $C_t$ denote the time-$t$ price of a European-style call option on $S$, struck at $K$ and expiring at $T = t + \tau$. Taking advantage of the affine structure of $(\ln S, V, r, q)$ and using the transform-based approach (see, for example, Stein and Stein, 1991; Heston, 1993; Scott, 1997; Bates, 2000; Bakshi et al., 1997; Bakshi and Madan, 2000; Duffie et al., 2000), we have

$$C_t = \mathrm{E}_t^Q\left[\exp\left(-\int_t^T r_u \, \mathrm{d}u\right)(S_T - K)^+\right] = S_t f\left(V_t, \vartheta, r_t, q_t, \tau, \frac{K}{S_t}\right), \tag{2.7}$$

where an explicit formulation for $f$ is given in Appendix B, and where for notational simplicity, we omit the explicit dependency of $f$ on $\theta_r$ and $\theta_q$.

## 3. Estimation

In this section, we focus on how to estimate the parametric model specified in Section 2 using the joint time-series data $\{S_n, C_n\}$ on spot and options. For notational convenience, we summarize the model parameters by $\vartheta$, as defined in Eq. (2.6). Treating the parameters $\theta_r$ and $\theta_q$ associated with the interest-rate process $r$ and the dividend-rate process $q$ as given,[9] our focus in this section is on the

---

[9] In practice, we first obtain maximum-likelihood (ML) estimates of $\theta_r$ and $\theta_q$ using time series of interest rates and dividend yields, respectively. We then treat the ML estimates of $\theta_r$ and $\theta_q$ as true parameters and adopt the "implied-state" GMM estimation strategy outlined here. Any loss of efficiency as a result of this approach is expected to be small, because the particular stochastic natures of $r$ and $q$ play a relatively minor role in pricing the short-dated options.

estimation of the true model parameter $\vartheta_0$, which is assumed to live in a compact parameter space $\Theta$.

Given that options are non-linear functions of the state variables, the joint dynamics of the market observables $S_n$ and $C_n$ could be complicated, irrespective of the analytical tractability of the state variables $(S, V)$. In order to take advantage of the analytical tractability of the state variables, we propose an "implied-state" generalized method of moments (IS-GMM) approach. The basic idea of our approach is to take advantage of the option-pricing relation $C_n = S_n f(V_n, \vartheta_0)$ by inverting, for a given set of model parameters $\vartheta$, a proxy $V_n^\vartheta$ for the unobserved volatility $V_n$ through $C_n = S_n f(V_n^\vartheta, \vartheta)$. Using $V_n^\vartheta$, we can focus directly on the dynamic structure of the state variables $(S, V)$. In our setting, the affine structure of $(\ln S, V)$ allows us to calculate the joint conditional moment-generating function of stock returns and volatility in closed-form, which, in turn, yields a rich set of moment conditions. Replacing $V_n$ by $V_n^\vartheta$ in the moment conditions, we can perform the usual GMM estimation, the only difference being that one of the state variables $V^\vartheta$ is parameter-dependent, hence the term "implied-state" GMM.

In the remainder of this section, we will first provide a detailed description of the IS-GMM estimators and then discuss the selection of the optimal moment conditions. The large sample properties of the IS-GMM estimators are established in Appendix C. A recursive formula for calculating the joint conditional moments of return and volatility is given in Appendix D.

## 3.1. "Implied-state" GMM estimators

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

$$y_n = \ln S_n - \ln S_{n-1} - \int_{(n-1)\Delta}^{n\Delta} (r_u - q_u)\, \mathrm{d}u \qquad (3.1)$$

denote the date-$n$ "excess" return,[10] it is easy to see that the transition distribution of $\{y_n, V_n\}$ depends only on the parameter vector $\vartheta$, and not on $\theta_r$ or $\theta_q$. Suppose, for the moment, that both the stock return $y_n$ and volatility $V_n$ can be observed. Our estimation problem then falls into a standard GMM setting. Specifically, we can select $n_h$ moment conditions such that

$$\mathrm{E}_{n-1}^{\vartheta_0}\left[ h\Big(y_{(n, n_y)},\ V_{(n, n_v)},\ \vartheta_0\Big)\right] = 0, \qquad (3.2)$$

---

[10] In order to construct the excess-return process $y$ defined by Eq. (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 $\Delta$. 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 $\Delta$ (our data are weekly), the effect of this approximation error on our results is assumed to be small. Alternative proxies for $\int_{(n-1)\Delta}^{n\Delta}(r_t - q_t)\,\mathrm{d}t$, 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.

where $\vartheta_0$ is the true model parameter, $h : \mathbb{R}^{n_y} \times \mathbb{R}^{n_v}_+ \times \Theta \to \mathbb{R}^{n_h}$ is some test function to be chosen,[11] $E^{\vartheta}_{n-1}$ denotes $\mathscr{F}_{(n-1)\varDelta}$-conditional expectation under the transition distribution of $(y, V)$ associated with parameter $\vartheta$, and, for some positive integers $n_y$ and $n_v$,

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

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

What distinguishes our situation from that of a typical GMM is that we do not observe the stock volatility $V_n$ directly. We can, nevertheless, take advantage of the market-observed spot price $S_n$ and option price $C_n$, and exploit the option-pricing relation

$$C_n = S_n f(V_n, \vartheta, r_n, q_n, \tau, k). \tag{3.3}$$

for an option with maturity $\tau$ and strike-to-spot ratio $k = K/S_n$. If the true model parameter $\vartheta_0$ is known, we can, in fact, back out the true volatility $V_n$ from this pricing relation using the market observables. For any other set of model parameters $\vartheta \in \Theta$, however, we can still back out a proxy $V_n^{\vartheta}$ for the unobserved volatility $V_n$ by solving $V_n^{\vartheta}$ from

$$C_n = S_n f\left(V_n^{\vartheta}, \vartheta, r_n, q_n, \tau, k\right). \tag{3.4}$$

The concept of backing out volatility from option prices is not novel, a prominent example being the Black–Scholes implied volatility. Our version of option-implied volatility $V_n^{\vartheta}$ differs from that of Black–Scholes in that ours is parameter dependent. Suppose that both volatility risk and jump risk are priced in the true model $\vartheta_0 : \eta^v \neq 0$ and $\mu^* \neq \mu$. If we start from a parameter set $\vartheta'$ that excludes jump risk and volatility risk from being priced, then the option-implied volatility $V_n^{\vartheta'}$ will be very different from the true volatility $V_n$. Alternatively, the closer $\vartheta$ is to the true model parameter $\vartheta_0$, the more accurate is the corresponding option-implied volatility $V_n^{\vartheta}$. When $V_n^{\vartheta}$ is evaluated at the true model parameter $\vartheta_0$, we retrieve the true volatility $V_n$.

Given the option-implied volatility $V_n^{\vartheta}$, we can now construct the sample analogue of the moment condition Eq. (3.2) by

$$G_N(\vartheta) = \frac{1}{N} \sum_{n \leqslant N} h\left(y_{(n,\, n_y)}, V^{\vartheta}_{(n,\, n_v)}, \vartheta\right), \tag{3.5}$$

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

$$\hat{\vartheta}_N = \arg \min_{\vartheta \in \Theta} G_N(\vartheta)^{\top} \mathscr{W}_N G_N(\vartheta), \tag{3.6}$$

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

This "implied-state" GMM approach raises several econometric issues. For example, one inherent feature of the exchange-traded options is that certain contract variables, such as time $\tau$ to expiration and strike-to-spot ratio $k$, vary from one observation to the next. This time dependency in contract variables could potentially

---

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

introduce a time dependency in the option-implied volatility, inadvertently resulting in an additional layer of complexity. In Appendix C, we show that, under mild technical conditions, the IS-GMM estimators are, indeed, consistent and asymptotically normal.

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] For example, Renault and Touzi (1996) propose an MLE-based two-step iterative procedure; Pastrorello et al. (1996) apply simulated method of moments (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; and, more recently, Eraker (2000) applies a Markov chain Monte Carlo based approach to joint time-series data on spot and options.[13]

Compared with these alternative approaches, the main motivation for us to adopt the IS-GMM approach is to take advantage of the affine structure of our dynamic model. Specifically, it allows us to focus directly on the joint dynamics of the state variables $(S, V)$, rather than the market observables $(S_n, C_n)$, which could be highly non-linear functions of the state variables. This is particularly attractive in our specification, because the affine structure of $(\ln S, V)$ provides a closed-form solution for the joint conditional moment-generating function of the stock return and volatility $(y, V)$ (Duffie et al., 2000), from which the joint conditional moments of $(y, V)$ can be calculated up to any desired order. As will be demonstrated in the subsequent section, such conditional moments can be used directly or indirectly (as optimal instruments) to build moment conditions. Moreover, they provide a rich set of diagnostic tests, allowing an explicit examination of how well various model constraints, e.g., constraints on the risk premium, fit with the joint time-series data of spot and option prices.

Finally, although our approach relies on one option $C_n$ per day to back out $V_n^\vartheta$, it does not preclude the use of multiple options. In fact, in Section 5.2, we introduce, in addition to the time series $\{C_n\}$ of near-the-money short-dated option prices, a time series $\{C_n^{\text{ITM}}\}$ of in-the-money call options to help 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 partially the motivation for us to use near-the-money short-dated options, since they typically are the most liquid options.

---

[12] This 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 employed 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 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".

[13] One advantage of the MCMC approach is that jumps are treated as an additional state variable, making it convenient to draw inferences about jump occurrences.

### 3.2. "Optimal" moment selection

This section provides a set of "optimal" moment conditions that takes advantage of the explicitly known moment-generating function of return and volatility ($y$, $V$), which is defined, for any $u_y$ and $u_v$ in $\mathbb{R}$, by

$$E_n\big[\exp(u_y y_{n+1} + u_v V_{v+1})\big] = \phi(u_y, u_v, V_n). \tag{3.7}$$

Given the explicitly known moment-generating function $\phi$ (see Appendix D), the joint conditional moments of returns and volatility can be derived by

$$E_n\big(y_{n+1}^i V_{n+1}^j\big) = \left.\frac{\partial^{(i+j)}\phi(u_y, u_v, V_n)}{\partial^i u_y \partial^j u_v}\right|_{u_y=0,\, u_v=0}, \quad i, j \in \{0, 1, \ldots\}. \tag{3.8}$$

Direct computation of the derivatives in Eq. (3.8), albeit straightforward, can nonetheless be cumbersome for higher orders of $i$ and $j$. Appendix D 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 $M_1(V_n, \vartheta) = E_n^\vartheta(y_{n+1})$, $M_2(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^2)$, $M_3(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^3)$, and $M_4(V_n, \vartheta) = E_n^\vartheta(y_{n+1}^4)$ denote the first four conditional moments of return, let $M_5(V_n, \vartheta) = E_n^\vartheta(V_{n+1})$ and $M_6(V_n, \vartheta) = E_n^\vartheta(V_{n+1}^2)$ denote the first two conditional moments of volatility, and let $M_7(V_n, \vartheta) = E_n^\vartheta(y_{n+1} V_{n+1})$ denote the first cross moment of return and volatility. We start with the following moment conditions:

$$E_{n-1}^\vartheta(\varepsilon_n) = 0, \quad \varepsilon_n = \big[\varepsilon_n^{y1}, \varepsilon_n^{y2}, \varepsilon_n^{y3}, \varepsilon_n^{y4}, \varepsilon_n^{v1}, \varepsilon_n^{v2}, \varepsilon_n^{yv}\big]^\top, \tag{3.9}$$

where

$$\varepsilon_n^{y1} = y_n - M_1(V_{n-1}, \vartheta), \quad \varepsilon_n^{y2} = y_n^2 - M_2(V_{n-1}, \vartheta),$$

$$\varepsilon_n^{y3} = y_n^3 - M_3(V_{n-1}, \vartheta), \quad \varepsilon_n^{y4} = y_n^4 - M_4(V_{n-1}, \vartheta),$$

$$\varepsilon_n^{v1} = V_n - M_5(V_{n-1}, \vartheta), \quad \varepsilon_n^{v2} = V_n^2 - M_6(V_{n-1}, \vartheta),$$

$$\varepsilon_n^{yv} = y_n V_n - M_7(V_{n-1}, \vartheta). \tag{3.10}$$

This choice of moment conditions is intuitive and provides some natural and testable conditions on certain lower moments and cross moments of $y$ and $V$. But these are not the most efficient moment conditions. In order to make them more efficient, we follow Hansen (1985) and introduce the following conditional instruments:

$$\mathscr{Z}_n = \mathscr{D}_n^\top \times \big(\mathrm{Cov}_n^\vartheta(\varepsilon_{n+1})\big)^{-1},$$

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

$$\mathscr{D}_n^i = -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta} - g_\vartheta(c_n, \vartheta)\left.\frac{\partial M_i(v, \vartheta)}{\partial v}\right|_{v=V_n}, \quad i = 1, 2, 3, 4,$$

$$\mathscr{D}_n^i = -\frac{\partial M_i(V_n, \vartheta)}{\partial \vartheta} \quad i = 5, 6, 7, \tag{3.11}$$

where $c_n$ is the date-$n$ option-to-spot ratio and $g_\vartheta(c_n, \vartheta)$ measures the sensitivity of the date-$n$ option-implied volatility $V_n^\vartheta = g(c_n, \vartheta)$ to $\vartheta$. [A formal definition of $g$ is given in Eq. (C.1).] Intuitively, the "optimal" instrument $\mathcal{L}_n$ has two components: the conditional covariance matrix $\mathrm{Cov}_n^\vartheta(\varepsilon_{n+1})$ corrects for the conditional heteroskedasticity in the original moment conditions $\varepsilon_{n+1}$, and the Jacobian $\mathcal{D}_n$ picks up the "conditional sensitivity" of $\varepsilon_{n+1}$ to the model parameters $\vartheta$. Taking advantage of the analytical tractability of our model, both components of $\mathcal{L}$ can be explicitly derived.

The "optimal" moment conditions can, therefore, be calculated by[14]

$$\mathcal{H}_{n+1} = \mathcal{L}_n \varepsilon_{n+1}. \tag{3.12}$$

Each element $\mathcal{H}_{n+1}^j$ of the "optimal" observations $\mathcal{H}_{n+1} = (\mathcal{H}_{n+1}^1, \ldots, \mathcal{H}_{n+1}^{n_\vartheta})$ is associated with an element $\vartheta_j$ of the parameter vector $\vartheta$. Intuitively, $\mathcal{H}_{n+1}^j$ is the weighted sum of the seven observations $\varepsilon_{n+1}$, normalized by the covariance matrix $\mathrm{Cov}_n^\vartheta(\varepsilon_{n+1})$, with weights proportional to the date-$n$ "conditional sensitivity" of $\varepsilon_{n+1}$ to $\vartheta_j$. Given this set $\mathcal{H}$ of "optimal" observations, we can apply our implied-state-variable approach outlined in Section 6 by replacing the unobserved stochastic volatility $V_n$ with the option-implied stochastic volatility $V_n^\vartheta$.

Finally, it should be noted that the efficiency of this "optimal-instrument" scheme is limited in that, in constructing $\mathcal{D}^5$, $\mathcal{D}^6$, and $\mathcal{D}^7$, we sacrifice efficiency by ignoring the dependence of $V^\vartheta$ on $\vartheta$. We do, however, gain analytic tractability, because calculations of the form $\mathrm{E}_n^\vartheta[g_\vartheta(c_{n+1}, \vartheta)]$, $\mathrm{E}_n^\vartheta[V_{n+1}g_\vartheta(c_{n+1}, \vartheta)]$, and $\mathrm{E}_n^\vartheta[y_{n+1}g_\vartheta(c_{n+1}, \vartheta)]$ would, indeed, be challenging.

# 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). This joint time-series is plotted in Fig. 1. The details of data collection 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, always between 10:00 and 10:30 a.m., varies from year to year. For example, it is set at 10:07–10:23 a.m. for all trading days in 1989; and at 10:14–10:16 a.m. for 1996. We adjust the length of the sampling window to accommodate

---

[14] Relative to full-information MLE, this approach sacrifices some efficiency by exploiting only a limited portion of the distributional information contained in the moment-generating function. See also Singleton (2001) and Liu (1997).
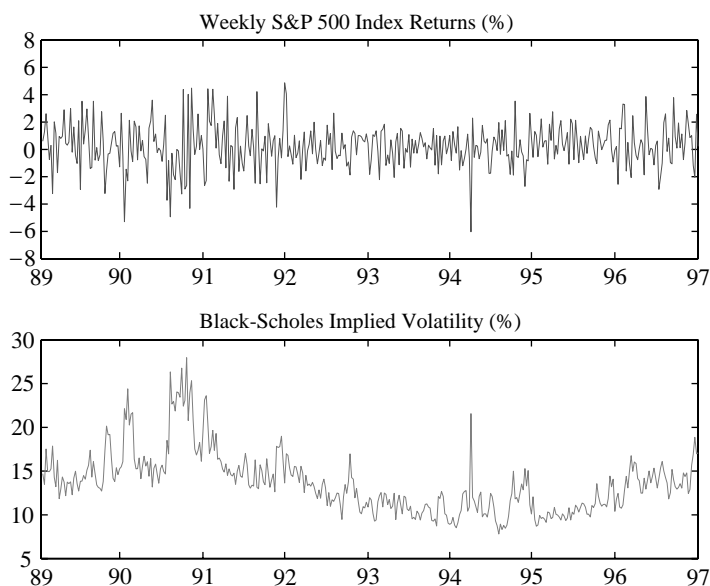
Fig. 1. Joint time series of weekly S&P 500 index returns and the near-the-money short-dated option prices.

significant changes from year to year in the trading volume of S&P 500 options. We also adjust the start time of the window so that the center of the window is at 10:15 a.m.

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 $\tau_n$ to expiration that is larger than 15 calendar days and as close as possible to 30 calendar days.[15] From the pool of options with the chosen time $\tau_n$ to expiration, we then select all options with a strike price $K_n$ nearest to the date-$n$ average of the S&P 500 index. If the remaining pool of options, with the chosen $\tau_n$ and $K_n$, contains multiple calls, we select one of these call options at random. Otherwise, a put option is selected at random.[16] 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. A valuable 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, accordingly, synchronized. The sample mean of $\{\tau_n\}$ is 31 days, with a sample

---

[15] Both time to expiration $\tau_n$ and sampling interval $\Delta$ are annualized, using a 365-calendar-day year and a 252-business-day year, respectively.

[16] We can 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 employing an additional contract variable. These two approaches are equivalent for our estimation strategy.
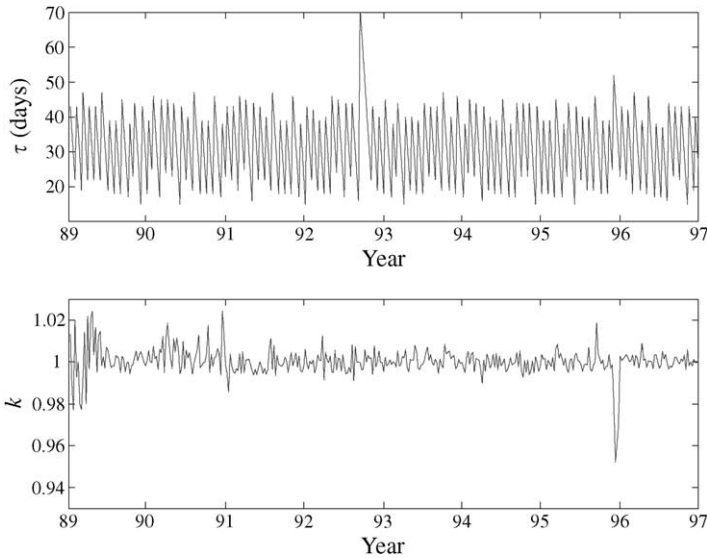
Fig. 2. Time series of contract variables: time-to-expiration $\tau$ and strike-to-spot ratios $k$.

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 Fig. 2.[17]

### 4.2. Time series of in-the-money short-dated calls

On each date $n$, we select an in-the-money call $C_n^{\mathrm{ITM}}$ 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 $C_n^{\mathrm{ITM}}$ 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 then convert the price to that of an ITM call using put/call parity. The sample mean of the strike-to-spot ratio $\{k_n^{\mathrm{ITM}}\}$ is 0.952, with a sample standard deviation of 0.007.

### 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 and December 1996, as measured by the Black–Scholes implied

---

[17] From September 11 to October 2, 1992, there is no October contract recorded in the BODB. This results in the "spike" in $\tau$ on the top panel of Fig. 2, because we have to use the November contracts for these observation days. Near the end of 1995, the exchange did not adjust its grid of strike prices to reflect the steady upward movement of the S&P 500 index. During this period, the highest strike price available was well below the spot S&P 500 index. This results in the "dip" in the strike-to-spot ratio on the bottom panel of Fig. 2.

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 10:00 and 11:00 a.m. For 1996, the time-window is reduced to 10:10–10:20 a.m., due to a surge in trading volume in 1996. Options with fewer 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. There is a total of 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

The estimation results are organized as follows. Section 5.1 focuses on the relative importance of jump- and volatility-risk premia in reconciling the joint time series $\{S_n, C_n\}$ of spot and option prices. Section 5.2 details a simultaneous estimation of the premia for jump and volatility risks, and evaluates the relative magnitude of these two types of risk premia, using an additional time series $\{C_n^{\text{ITM}}\}$ of in-the-money short-dated call options. Section 5.3 extends the analysis to the cross-sectional option data, providing further evidence for jump-risk premia. Finally, Section 5.4 examines possible model mis-specifications and their implications on our findings of jump-risk premia. The estimation results associated with the risk-free rates $r$ and dividend yields $q$, as well as the results of a Monte-Carlo study, are presented in Appendix E.

### 5.1. Reconciling spot and option dynamics

In order to examine the role of risk premia in reconciling spot and option dynamics, we focus on three nested models of Eqs. (2.1)–(2.5):

- The SVJ0 model: $\eta^v = 0$.
- The SV model: $\lambda = 0$.
- The SV0 model: $\lambda = 0$ and $\eta^v = 0$.

These nested models are chosen to represent three different risk-premium structures: jump-risk premia (SVJ0), volatility-risk premia (SV), and no risk premia (SV0).

For all three models, we perform joint estimations of their actual and risk-neutral dynamics using the time series $\{S_n, C_n\}$ of the S&P 500 index and the near-the-money short-dated option prices. (See data collection details in Section 4.1.) The estimation results are reported in Table 1 and the results of goodness-of-fit tests are summarized in Table 2. The goodness-of-fit tests are constructed directly from the

Table 1
IS-GMM estimates of three nested models[a]

|      | $\kappa_v$ | $\bar{v}$ | $\sigma_v$ | $\rho$ | $\eta^s$ | $\eta^v$ | $\lambda$ | $\mu$ (%) | $\sigma_J$ (%) | $\mu*$ (%) |
|------|-----------|-----------|-----------|--------|----------|----------|-----------|-----------|----------------|-----------|
| SVJ0 | 7.1       | 0.0134    | 0.28      | −0.52  | 3.1      | ≡0       | 27.1      | −0.3      | 3.25           | −18.0     |
|      | (1.9)     | (0.0029)  | (0.04)    | (0.07) | (2.9)    |          | (11.8)    | (1.7)     | (0.64)         | (1.6)     |
| SV   | 7.1       | 0.0137    | 0.32      | −0.53  | 8.6      | 7.6      | ≡0        | —         | —              | —         |
|      | (2.1)     | (0.0023)  | (0.03)    | (0.06) | (2.3)    | (2.0)    |           |           |                |           |
| SV0  | 5.3       | 0.0242    | 0.38      | −0.57  | 4.4      | ≡0       | ≡0        | —         | —              | —         |
|      | (1.9)     | (0.0044)  | (0.04)    | (0.05) | (1.8)    |          |           |           |                |           |

[a] Data: Weekly spot and options data, S&P 500 index, Jan. 1989–Dec. 1996.

Table 2
Goodness-of-fit tests

|                  | $\tilde{\varepsilon}$ | SV0         | SV         | SVJ0   |
|------------------|-----------------------|-------------|------------|--------|
| Individual tests | $y1$                  | 1.46        | −0.59      | 0.27   |
|                  | $y2$                  | −3.98**     | −1.56      | −0.60  |
|                  | $y3$                  | 0.77        | −0.29      | −0.65  |
|                  | $y4$                  | −1.45       | 0.27       | −0.36  |
|                  | $v1$                  | −1.91       | 1.80       | 0.95   |
|                  | $v2$                  | −2.28*      | 1.24       | 0.59   |
|                  | $yv$                  | 2.47*       | −0.10      | 0.60   |
| Joint tests      | All $y$               | 28.2**      | 8.1        | 1.8    |
|                  | $\chi^2$ (4)          | $(10^{-5})$ | (0.09)     | (0.77) |
|                  | All $v$               | 9.1*        | 11.4**     | 3.2    |
|                  | $\chi^2$ (2)          | (0.01)      | (0.003)    | (0.20) |
|                  | All                   | 59.9**      | 31.6**     | 7.6    |
|                  | $\chi^2$ (7)          | $(10^{-10})$| $(10^{-5})$| (0.37) |

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

heteroskedasticity-corrected version $\tilde{\varepsilon}$ of $\varepsilon$, defined by

$$\tilde{\varepsilon}_n^i = \frac{\varepsilon_n^i}{\sqrt{E_{(n-1)}\left(\varepsilon_n^i\right)^2}}, \quad i \in \{1, \ldots, 7\}.$$

We test the seven moment conditions, $E_{n-1}(\tilde{\varepsilon}_n) = 0$, both individually and jointly.[18]

As summarized in Table 2, both the SV0 and SV models are strongly rejected by the joint time-series data (with $p$-values of $10^{-10}$ and $10^{-5}$, respectively), while the SVJ0 model is not rejected (associated $p$-value = 0.37). To gain some insights to this

---

[18] The large-sample distribution of the test statistics is standard normal for the individual tests, and $\chi^2$ with $n$ degrees of freedom for a joint test on $n$ moment conditions. Appendix E provides further details on large-sample distributions of such test statistics.

result, we focus first on the SV0 model and on the moment condition

$$\mathrm{E}\big(\varepsilon_n^{y2}\big) = 0 \quad \text{with} \quad \varepsilon_n^{y2} = y_n^2 - M_2\big(V_{n-1}^{\vartheta}, \vartheta\big),$$

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 time period $\Delta$, we have $M_2(V_{n-1}^{\vartheta}, \vartheta) \approx V_{n-1}^{\vartheta}\Delta$, which leads to $\varepsilon_n^{y2} \approx y_n^2 - V_{n-1}^{\vartheta}\Delta$. For the SV0 model, the test statistics associated with $\mathrm{E}(\varepsilon_n^{y2})$ is significantly negative, indicating that the volatility realized in the spot market is significantly less than that observed in the options market (through the SV0 model).

One possible explanation is that the SV0 model has a very rigid risk-premium structure. In particular, investors' aversion to volatility uncertainty or jump risks is not incorporated. To examine the role of these two types of risk premia, we next focus on the SV and SVJ0 models. Table 1 shows that the volatility-risk premium coefficient $\eta^v$ is estimated to be positive and significantly different from zero for the SV model,[19] which is consistent with the findings of Guo (1998), Benzoni (1998), Poteshman (1998), Bakshi and Kapadia (2001), and Chernov and Ghysels (2000). Not examined in these studies or in the SV model is the role of jump-risk premia. Indeed, the results for the SVJ0 model imply large and significant premia for the jump-size uncertainty. Table 1 shows that the risk-neutral mean $\mu^*$ of the relative jump size is estimated to be $-18\%$, while its counterpart $\mu$ for the data-generating process is estimated to be $-0.3\%$. ($\mu^* - \mu$ is estimated at $-17.6\%$, with a standard error of $2.2\%$). This implies that, when weighted by aversion to large price movements, negative jumps are perceived to be more negative. 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.[20]

A revisit to the goodness-of-fit test associated with $\mathrm{E}(\varepsilon_n^{y2})$ reveals that it is no longer strongly violated for either the SV model or the SVJ0 model, suggesting that allowing for risk premia does reconcile, to some extent, the tension between the spot and option prices that arises in the SV0 model. This, however, does not imply that the risk premia implicit in option prices can be explained equally well by the premia for volatility risk or those for jump risk.

First, in contrast to the SVJ0 model, the SV model is still strongly rejected by the joint time-series data ($p$-value $= 10^{-5}$). Second, in contrast to the SVJ0 model, the estimated SV model implies an explosive risk-neutral volatility process, and severely over-prices long-dated options. To illustrate, we plot in Fig. 3 near-the-money

---

[19] Instead of testing the significance of the estimator for $\eta^v$, we can also perform a Lagrange-multiplier test of the SV0 model against the SV model using the moment condition $\mathrm{E}_n[\mathscr{H}_{n+1}(\eta^v)] = 0$, where $\mathscr{H}_{n+1}(\eta^v)$ denotes the "optimal" moment associated with $\eta^v$ as described in Section 3.2. In particular, this test is of the Lagrange-multiplier style in the sense that the moment condition $\mathrm{E}_n[\mathscr{H}_{n+1}(\eta^v)] = 0$, which is true under the alternative (the SV model), is tested using the parameter estimates associated with the Null (the SV0 model). The SV0 model (with $\eta^v = 0$) is rejected against the SV model ($p$-value $= 0.0002$).

[20] It should be noted that because we have set the jump-timing risk premium $\lambda^* - \lambda$ 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. We postpone a discussion on this issue to the next section.
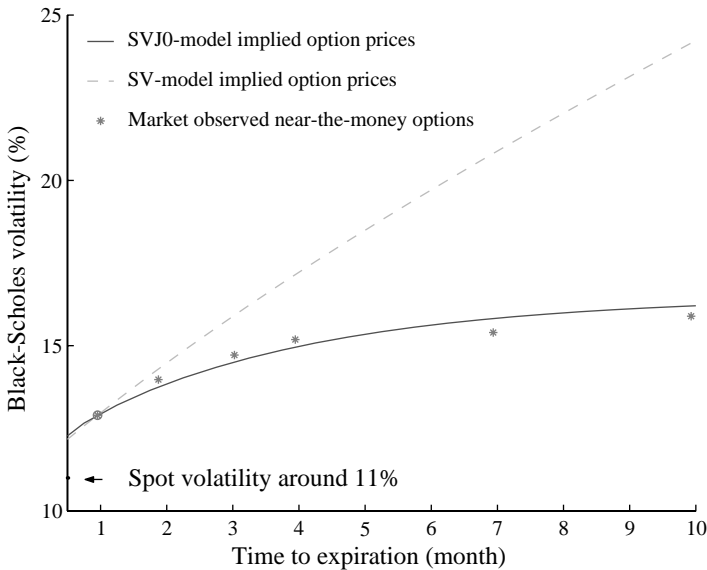
Fig. 3. On a medium-volatility day (November 22, 1996), the SVJ0 model (with jump-risk premia) is capable of pricing both short- and long-dated options, while the SV model (with volatility-risk premia) severely over-prices long-dated options.

options with varying degrees of maturities. The stars indicate the market observed prices, while the solid and dashed lines indicate the SVJ0- and SV-model implied option prices, respectively. The spot-market volatility is about 11% for this day, while the Black–Scholes implied-vol for the one-month contract is around 13%. Consistent with our earlier discussion, there is a gap between the volatility measured directly from the spot market and that implied by option prices (through either the Black–Scholes model or the SV0 model, which have essentially the same risk-premium structure). Both the SVJ0 and SV models are able to explain this gap – one by jump-risk premia and the other by volatility-risk premia. But their respective projections on the long-dated options are very different. In particular, for the SV model to explain the risk premia implicit in the one-month options, it requires a volatility-risk premium coefficient $\eta^v$ of such high magnitude that the estimated risk-neutral mean-reversion rate $\hat{\kappa}_v^* = \hat{\kappa}_v - \eta^v$ is negative, implying an explosive risk-neutral volatility process. Although this is not explicitly precluded by arbitrage arguments, it does cause the SV model to severely overprice long-dated options, as illustrated in Fig. 3. The jump-risk premia, on the other hand, can explain the risk premia implicit in one-month options without distorting long-dated options.[21]

---

[21] Quite intuitively, the term structure of volatility-risk premia is very different from that for jump-risk premia. In fact, it can be shown that as the maturity of an option approaches zero, the effect of volatility-risk premia diminishes, while that for jump-risk premia does not. It is closely related to the fact that volatility risk influences stock prices through (diffusive) Brownian motions.

Fig. 4 illustrates another important difference between the two types of risk premia considered in this paper. For short-dated options, the jump-risk premia respond quickly to market volatility, while the volatility-risk premia do not. For instance, Fig. 4 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 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.

Overall, the SVJ0 model clearly dominates the pure diffusion models, demonstrating the importance of the state-dependent jump-risk premia. We next examine the impact of the zero volatility-risk premia constraint ($\eta^v = 0$) imposed by the SVJ0 model. Letting $\mathscr{H}_n(\eta^v)$ denote the "optimal" moment associated with $\eta^v$, we test the SVJ0 model against the alternative that $\eta^v \neq 0$. Using the moment condition $E[\mathscr{H}_n(\Sigma^v)] = 0$, 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. (*p*-value = 0.55) This implies that introducing volatility-risk premia in addition to jump-risk premia will not result in any significant improvement in the goodness of fit.

In our model specification, we assume that jump arrival intensity is $\lambda V_t$, leaving out the possibility that there might be a constant component to the jump arrival intensity. We now test the specification of $\lambda V_t$ against the alternative $\lambda_0 + \lambda V_t$, for
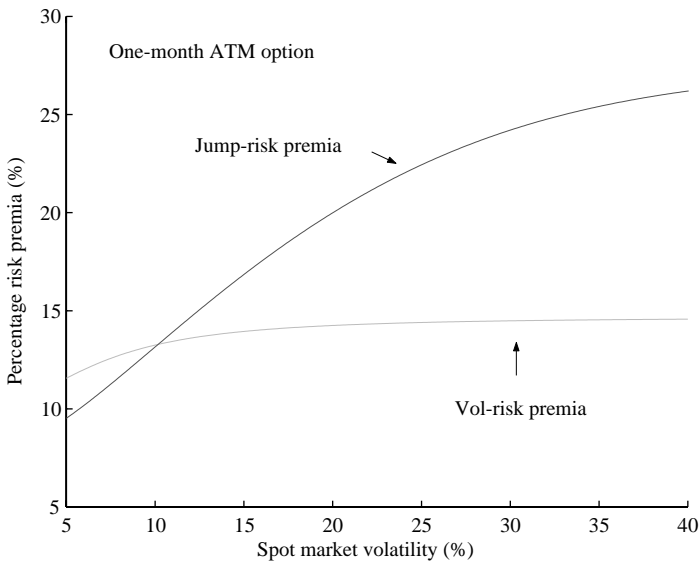


Fig. 4. Different responsiveness to market volatility. The percentage risk premia measure the amount of cents one has to pay for risk premia for every dollar invested in the one-month at-the-money options. The market volatility is in terms of the Black–Scholes volatility.

some constant $\lambda_0$. More specifically, we let $\mathcal{H}_n(\lambda_0)$ denote the "optimal" moment associated with $\lambda_0$, and use the moment condition $E[\mathcal{H}_n(\lambda_0)] = 0$, which is true under the alternative. We find that the SVJ0 model is not rejected against the alternative that $\lambda_0 \neq 0$ at traditional confidence levels.[22] (p-value = 0.12) This result implies that between the two components in jump-risk premia – one constant, the other state dependent, the state-dependent component plays a far more important role in reconciling the joint time-series data. In fact, in our first attempt to estimate jump-risk premia (not reported here), we adopt the specification of constant jump arrival intensity. For the purpose of reconciling the joint time-series data, this specification is problematic because it assigns the same amount of jump-risk premia irrespective of the market volatility. In such a setting, on days of high volatility, when options become more expensive, the percentage jump-risk premia effectively decreases. This partially explains the findings of Eraker (2000), who adopts the specification of constant jump-arrival intensity, and concludes jump-risk premia to be insignificant under this setting.

## 5.2. Simultaneous estimation of jump- and volatility-risk premia

In this section, we focus on a simultaneous estimation of the premia for jump and volatility risks, evaluating the relative magnitude of the two types of risk premia. For this, we focus on the SVJ model, which relaxes the constraint $\eta^v = 0$ from the SVJ0 model.

A direct estimation of the SVJ model using the joint time-series data $\{S_n, C_n\}$ reveals that our ability to pin down the relative magnitude of volatility-risk and jump-risk premia is limited.[23] Using the intuition that OTM/ITM options might have different sensitivities to the volatility and jump risks, we introduce an additional time series $\{C_n^{\mathrm{ITM}}\}$ of ITM calls (see data collection details in Section 4.2) in our model estimation. We assume that such ITM calls are priced with errors, and that the pricing errors are 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 to the bid/ask spread.[24] Under

---

[22] It should be noted that this test result does not imply that the constant component $\lambda_0$ 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 et al. (1999) report evidence in support of such a constant component. Johannes et al. (1998), on the other hand, provide evidence in support of the state-dependent component in jump arrival intensity. 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 $\mu$ is not significant.) 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 $\lambda$ plays a far more important role than the constant component $\lambda_0$.

[23] The estimation results for the SVJ model are not reported, but are available upon request.

[24] We choose ITM calls instead of OTM puts because there are overwhelmingly more quotes on ITM calls than on OTM puts. Consequently, it is easier to construct a time series of ITM-call quotes with strike-to-spot ratio as close to 0.95 as possible. This behavior of quote data is in direct contrast with the trade data—OTM puts are more liquid than ITM calls. Given that put-call parity holds relatively well, and that we scale the pricing error by the bid/ask spread, which should reflect liquidity, the choice between OTM puts and ITM calls is not expected to affect our results in any significant fashion.

these assumptions, letting $c_n^{\text{ITM}}$ and $k_n^{\text{ITM}}$ be the date-$n$ price-to-spot ratio and strike-to-spot ratio, respectively, of the in-the-money call option, we form the moment condition

$$\mathrm{E}\left(\mathscr{H}_n^{\text{ITM}}\right) = 0 \quad \text{with} \quad \mathscr{H}_n^{\text{ITM}} = \frac{c_n^{\text{ITM}} - f\left(V_n, \vartheta, r_n, q_n, \tau_n, k_n^{\text{ITM}}\right)}{\delta_n^{\text{ITM}}}, \tag{5.1}$$

where $\delta_n^{\text{ITM}}$ is the date-$n$ bid–ask spread (per unit spot price) and $f$ is the SVJ option pricing formula defined in Eq. (B.5). In addition to the 10 "optimal" moment conditions associated with the 10 SVJ-model parameters $\vartheta$, we introduce this moment condition to exploit the risk-premia information embedded in the ITM call (or OTM put) option prices.

Table 3 shows that both types of risk premia are estimated to be positive. The coefficient $\mu - \mu^*$ associated with the premia for jump-size uncertainty is estimated to be 18.4% with a standard error of 4%, while the coefficient $\eta^v$ associated with the premia for volatility risk is estimated to be 3 with a standard error of 2. Compared with the estimates for the SVJ0 model, which imposes the constraint $\eta^v = 0$, we see a reduction in jump-risk premia ($\lambda$ is reduced by half) accompanying the positive volatility-risk premia. By taking advantage of the additional time series of ITM call options, we can better pin down the jump-risk premia. The coefficient $\eta^v$ for the volatility-risk premia, however, still cannot be precisely measured. Searching 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. We also examine the goodness of fit by looking at the over-identifying restriction introduced by the additional moment condition Eq. (5.1). The $p$-value for that test is 0.58.

The SVJ model extends the SVJ0 model by allowing both the jump- and the volatility-risk premia to play a role in option pricing. For intuition purpose, we offer some approximate estimates of the breakdown between the two types of risk premia. Consider a one-month option traded at 10% market volatility. If the option is at the money, then about 55% of the overall risk premia is paid for jump risk and the rest for volatility risk. If the option is a 5% OTM put, then the jump-risk premium component increases to 80%, but for a 5% OTM call option the percentage decreases to 30%. In other words, investors are more worried about jump risk in OTM puts than OTM calls.

Similarly, the effect of risk premia also shows up in the underling stock prices. The time-$t$ instantaneous equity risk premium has two components: (1) $\eta^s V_t$ compensates

Table 3
IS-GMM estimates of the SVJ model[a]

| $\kappa_v$ | $\bar{v}$ | $\sigma_v$ | $\rho$ | $\eta^s$ | $\eta^v$ | $\lambda^* = \lambda$ | $\mu$ (%) | $\sigma_J$ (%) | $\mu^*$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| 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) |

[a] Weekly time-series $\{S_n, C_n\}$ of the S&P 500 index and the near-the-money short-dated options from January 1989 to December 1996. An additional time series $\{C_n^{\text{ITM}}\}$ of in-the-money calls are used.

for the usual diffusive return risk, and (2) $\lambda V_t(\mu - \mu^*)$ 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 the diffusive return risk is 5.5% per year (with a standard deviation of 3.4%), while that demanded for the 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 to 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.

Finally, recall that the SVJ model sets the jump-timing risk premium $\lambda^* - \lambda$ to zero. To gauge its impact on the estimated premium for the jump-size risk (measured in terms of $\mu^* - \mu$), we estimate $\lambda^*$ and $\lambda$ as two free parameters.[25] Our finding with respect to the premium for jump-size uncertainty remains robust: $\mu^* - \mu$ is estimated to be $-17.7\%$ with a standard error of 1.7%. The estimate for $\lambda$, however, is very different from those obtained under the constraint $\lambda^* = \lambda$. In particular, we find $\lambda^* < \lambda$, implying a negative jump-timing risk premium. This result could be due to the fact that different types of jumps coexist in the data. To ensure a non-negative jump-timing risk premium, we perform the same estimation under the constraint that $\lambda^* \geqslant \lambda$. The constraint is found to be binding and the estimates are close to those reported in Table 3. In short, the evidence in support of the premium for jump-size uncertainty is robust with respect to the constraint of zero jump-timing risk premium.

## 5.3. Implications on cross-sectional option prices

In this section, we extend our analysis to cross-sectional option data. (Details on data collection are given in Section 4.3.) Equipped with the time-series estimation results summarized in Tables 1 and 3, we examine the extent to which these models correctly price the cross-sectional options data observed in the market.

Table 4 summarizes the cross-sectional pricing errors, which are measured as the absolute differences between the model-implied and the market-observed option prices, both measured in terms of the Black–Scholes implied volatility. 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. As a gauge of how big the mis-pricings are in real terms, we also provide the average bid/ask spreads—the difference between offer and ask prices, each measured in terms of BS vol. For a pictorial exposition, Fig. 5 plots the volatility smiles (across different maturities) on a medium-volatility day, while Figs. 6 and 7, respectively, plot the volatility smiles on the most and least volatile days in our sample.

Overall, the ability of the SVJ0 model to capture the volatility smirks is quite remarkable, given only one option a day (marked by circles in Figs. 5–7) is used to obtain the SVJ0 model estimates, while the rest of the options in the cross-sectional data is purely out of sample. Moreover, our results indicate that the SVJ0 model can

---

[25] Details are available upon request.

Table 4
The SVJ model-implied pricing errors for the S&P 500 call and put options[a]

| | | 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$^{ITM}$ | | 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 | $60\leqslant\tau<180$ | 5.3 (−) | 4.2 (−) | 3.0 (−) | 2.1 (−) | 0.9 (−) | 0.5 (+) | 0.7 (−) | 0.6 (+) | 2.3 (−) |
| SV | | 2.8 (+) | 2.6 (+) | 4.2 (+) | 1.5 (+) | 2.1 (+) | 3.1 (+) | 1.5 (+) | 2.2 (+) | 3.2 (+) |
| SVJ0 | | 3.4 (−) | 3.6 (−) | 3.9 (−) | 0.6 (+) | 0.3 (+) | 0.4 (−) | 2.1 (+) | 1.1 (+) | 2.5 (−) |
| SVJ$^{ITM}$ | | 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 | $\tau\geqslant180$ | 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 (+) |
| 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 |

[a] *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.
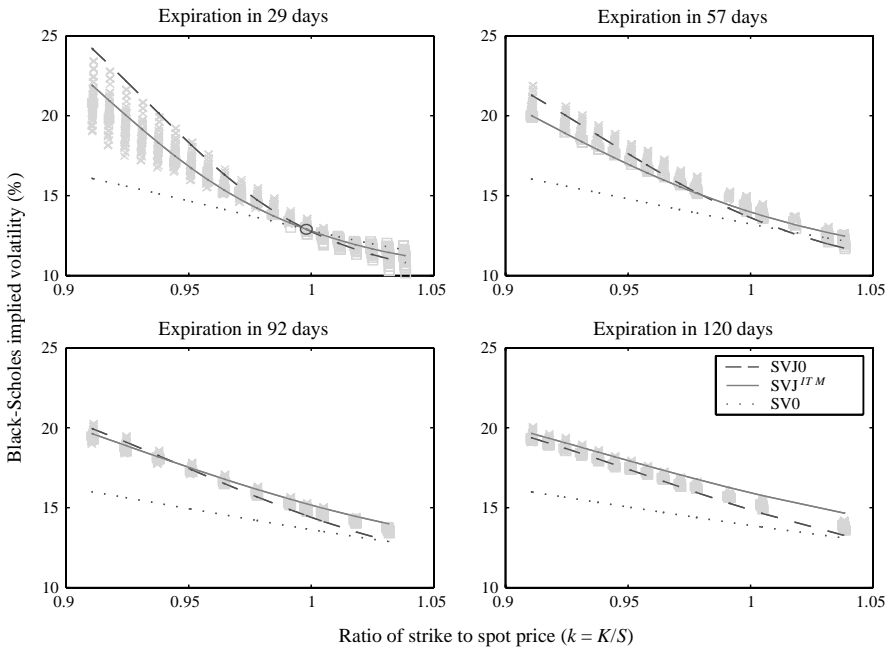
Fig. 5. Smiles curves on a medium-volatility day of the sample. All observations are observed between 10:10 and 10:20 a.m. on November 22, 1996. The call options are marked by ' × ', and the put options by '□'.

fit the steepness of the smile curves quite well on days of high, medium, and low volatilities. This is to be compared with the findings of Eraker (2000), who, using the specification of constant jump arrival intensity, finds that such a jump model cannot generate sufficiently steep smile curves on high-volatility days. This reinforces our earlier discussion in Section 5.1 on the importance of the state-dependent nature of the jump-risk premia. In particular, this state-dependent jump-risk premium is crucial not only for reconciling the spot and option dynamics, but also for explaining the changes over time of the smile curves.

While it can capture the volatility smirks relatively well, the SVJ0 model consistently under-prices medium and long-dated options on days of high volatility, and over-prices them on days of low volatility. This is closely related to the inability of the one-factor volatility model to capture the rich term structure of volatility implied by the data. As will be discussed in Section 5.4, the SVJ0 model reverts to its long-run mean too rapidly, compared with what is implied by the data. Consequently it under-prices long-dated options on high-volatility days and over-prices these options on low-volatility days.

On days of high and low volatilities, we also observe a large volume of quotes for deep-in-the-money puts, whose prices none of our models is 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 that suggested by the estimated models. A possible solution is to allow jumps in volatility,
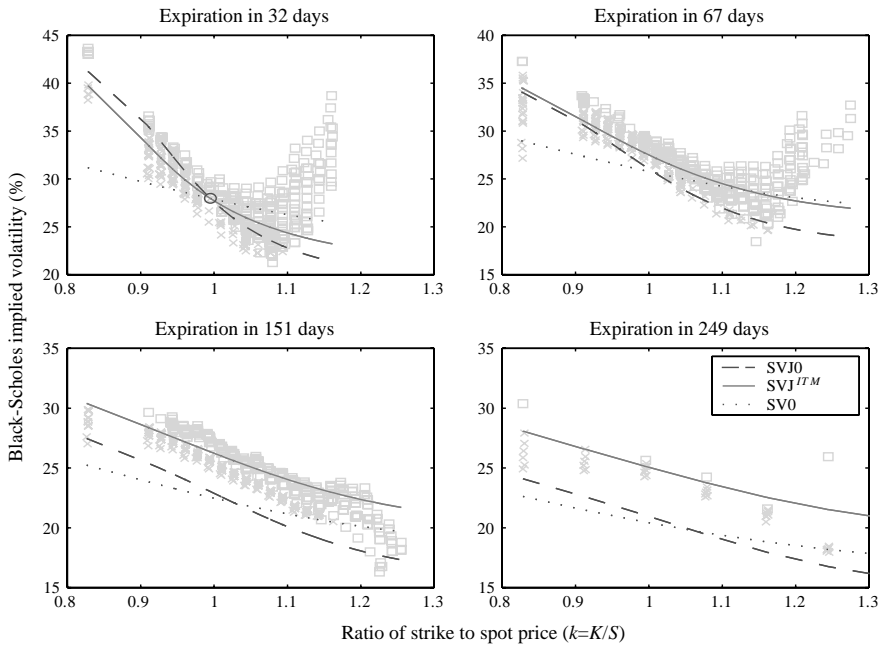
Fig. 6. Smiles curves on the most volatile day of the sample. All observations are observed between 10:00 and 11:00 a.m. on October 16, 1990. The call options are marked by ' × ,' and the put options by '□'.

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 the SV model is capable of explaining options across moneyness. This result is not novel. Fitting this class of stochastic volatility models directly to cross-sectional option prices, Bates (2000) and Bakshi et al. (1997) arrive at the same conclusion. Table 4 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.

## 5.4. Diagnostic tests on model mis-specifications

This section examines possible mis-specifications of the volatility dynamics.[26] For each of the parametric models considered in this paper, we obtain a time-series of

---

[26] Given that option prices are influenced by both the price dynamics and the market prices of risks, the diagnostic tests in this section and the goodness-of-fit tests in Table 2 are joint tests on the dynamics and the pricing kernel. The motivations for these tests, however, are different. In particular, the diagnostic tests designed in this section focus on the part of volatility dynamics that is less sensitive to the mis-specifications in pricing kernel. To illustrate this point, the diagnostic tests are performed for the four different models with very different specifications of pricing kernels.
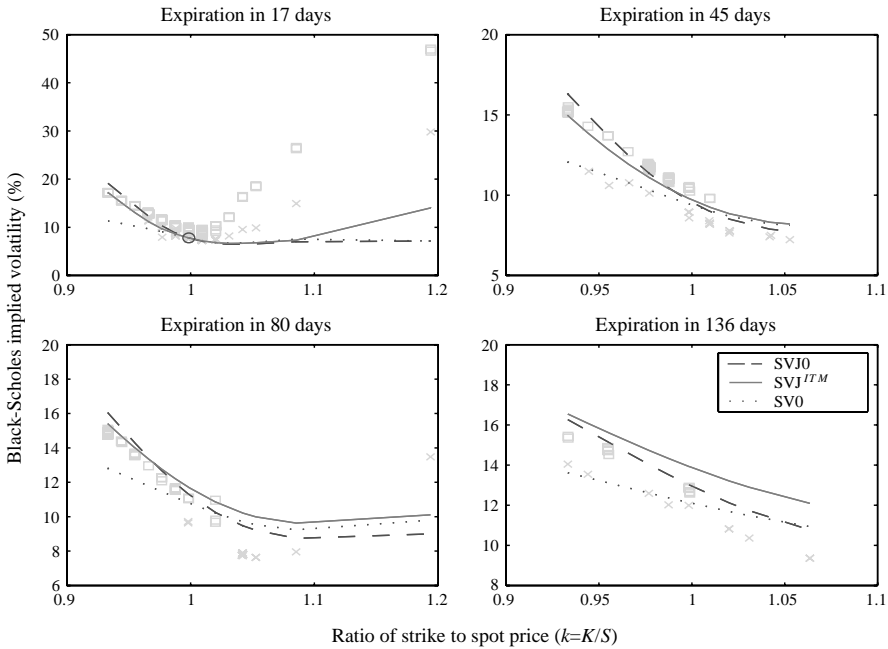
Fig. 7. Smiles curves on the least volatile day of the sample. All observations are observed between 10:00 and 11:00 a.m. on August 3, 1994. The call options are marked by ' × ,' and the put options by ' □ '.

option-implied volatility $\{V_n^{\vartheta}\}$. For example, Fig. 8 plots the series $\{\sqrt{V_n^{\vartheta}}\}$ for the SVJ0 model.[27] Taking advantage of the option-implied volatility series, our diagnostic tests focus on two issues: (1) the term structure of volatility and (2) jumps in volatility.

Given that there have been extensive empirical studies conducted on the stochastic-volatility models with jumps, it is instructive for us to first compare our model estimates—the subset associated with the data-generating process—with those reported in time-series studies using stock indices alone. Taking the SVJ0 model estimates (Table 1), for example, we have $\kappa_v = 7$, $\sqrt{\bar{v}} = 11\%$, $\sigma_v = 0.28$, and $\rho = -0.5$. This is compared with $\kappa_v = 3.2$, $\sqrt{\bar{v}} = 14\%$, $\sigma_v = 0.24$, and $\rho = -0.47$ reported by Eraker et al. (2000), and $\kappa_v = 3.4$, $\sqrt{\bar{v}} = 13\%$, $\sigma_v = 0.17$, and $\rho = -0.33$ reported by Andersen et al. (1998). Our estimate of the long-run mean of volatility is slightly lower because our sample period falls into a less volatile period than theirs. Overall, our estimates are of the same order of magnitude. This is quite remarkable, given that their volatility estimates rely on the time series of the S&P 500 index, while

[27] We plot $\{\sqrt{V_n^{\vartheta}}\}$ in Fig. 8 so that a direct comparison can be made with the Black–Scholes implied volatility (Fig. 1). While the time-series patterns are similar, the levels of inferred volatility are quite different. In particular, because of the different risk-premium structures of the two models, the volatilities inferred from using the Black–Scholes model are higher than those inferred from using the SVJ0 model.
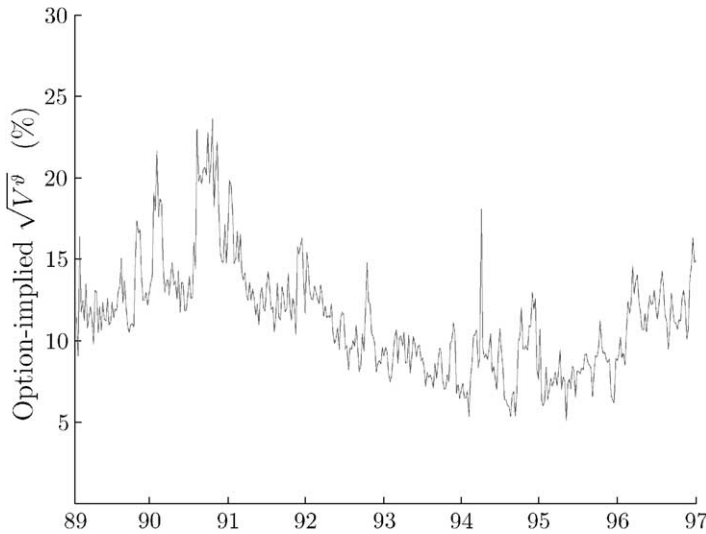
Fig. 8. The time series of $\{\sqrt{V_n^\vartheta}\}$ inferred from the time series $\{C_n\}$ of near-the-money short-dated options, using the SVJ0-model parameters $\vartheta$.

Table 5
Diagnostic tests on volatility dynamics

|  | SV0 | SV | SVJ0 | SVJ | SVJ$^{ITM}$ |
|---|---|---|---|---|---|
| $E\left(\varepsilon_n^v \varepsilon_{n+1}^v\right) = 0$ | −2.37* | −2.62** | −2.04* | −2.19* | −2.38* |
| $E\left(\varepsilon_n^{v3}\right) = 0$ | −0.69 | 0.85 | 0.70 | 0.54 | −0.59 |
| $E\left(\varepsilon_n^{v4}\right) = 0$ | −0.38 | 0.65 | 0.62 | 0.53 | 0.05 |
| $E\left(\tilde{\varepsilon}_n^{v3}\right) = 0$ | −0.58 | 1.46 | 2.17* | 2.59** | 0.13 |
| $E\left(\tilde{\varepsilon}_n^{v4}\right) = 0$ | 0.95 | 1.68 | 2.67** | 3.75** | 1.03 |

*Indicate significance under a 5% test.
**Indicate significance under a 1% test.
 The SVJ and SVJ$^{ITM}$ cases are for the same model except that, in the case of SVJ$^{ITM}$, one additional time series of ITM calls is used in model estimation.

ours depend heavily on option prices and on how risk premia are accounted for in option prices. The jump estimates, however, seem to be more diverse. As noted earlier, our ability to identify the jump component in the actual dynamics is quite limited by our relatively short sample.

Our first set of diagnostic tests indicates that the stochastic volatility model of Heston (1993) is not rich enough to capture the term structure of volatility implied by the data. (See also Stein, 1989; Bollerslev and Mikkelsen, 1996.) Letting $E_n(\varepsilon_{n+1}^{v1}) = 0$ be the moment condition associated with the first moment of volatility, our model specification implies that $\varepsilon^{v1}$ is serially uncorrelated, i.e., $E_n(\varepsilon_n^{v1} \varepsilon_{n+1}^{v1}) = 0$. As reported in Table 5, this hypothesis is strongly rejected. The sample estimate of

corr $(\varepsilon_n^{v1}, \varepsilon_{n+1}^{v1})$ is negative and significant. To see the implication of this finding, we recall that $\varepsilon_n^{v1} = V_n^\vartheta - M_5(V_{n-1}^\vartheta, \vartheta)$, 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 as[28]

$$\mathrm{corr}(\varepsilon_n^{v1}, \varepsilon_{n-1}^{v1}) = \mathrm{corr}(V_{n+1} - V_n e^{-\kappa_v \Delta t}, \ V_n - V_{n-1}e^{-\kappa_v \Delta t}).$$

Using the fact that the one-step auto-correlation $\mathrm{corr}(V_n, V_{n-1})$ is $\exp(-\kappa_v \Delta)$, we have $\mathrm{corr}(\varepsilon_n^{v1}, \varepsilon_{n-1}^{v1}) = e^{-\kappa_v \Delta t}[e^{-2\kappa_v \Delta t} - \mathrm{corr}(V_{n+1}, V_{n-1})]$. A negative and significant sample estimate of $\mathrm{corr}(\varepsilon_v^{v1}, \varepsilon_{n-1}^{v1})$ therefore indicates that the data call for $\mathrm{corr}(V_{n+1}, V_{n-1}) > \exp(-2\kappa_v \Delta t)$. This is in contrast to the model-prescribed two-step auto-correlation $\mathrm{corr}(V_{n+1}, V_{n-1}) = \exp(-2\kappa_v \Delta t)$. In other words, the stochastic volatility model is not capable of fitting one- and two-step auto-correlations simultaneously, the reason being that the model prescribes a term-structure of volatility, $\mathrm{corr}(V_n, V_{n+m}) = \exp(-m\kappa_v \Delta t)$, which "dies" too quickly relative to the data.

To accommodate a richer term structure of volatility, one solution is to allow for multiple volatility factors with different rates of mean-reversion.[29] An implementation of two-factor models, although feasible, is not considered here. This mis-specification, however, is not likely to influence our main empirical findings concerning the importance of jump-risk premia. This is because that our time-series estimation focuses mainly on short-dated options, the pricing of which is not significantly affected by such a mis-specification. For the purpose of pricing long-dated options, as well as for a better understanding of volatility-risk premia, however, accommodating a richer term structure of volatility is important. As reported in Section 5.3, this mis-specification results in under-priced long-dated options on high-volatility days, and over-priced long-dated options on low-volatility days.

We next focus on the higher moments of the volatility process, seeking evidence of jumps in volatility, as conjectured by Bates (2000). Let $\mathrm{E}_n(\varepsilon_{n+1}^{v3}) = 0$ and $\mathrm{E}_n(\varepsilon_{n+1}^{v4}) = 0$ be the moment conditions associated with the third and fourth moments of volatility. Table 5 reports the conditions $\mathrm{E}(\varepsilon_{n+1}^{v3}) = 0$ and $\mathrm{E}(\varepsilon_{n+1}^{v4}) = 0$, along with their respective heteroskedasticity-corrected versions, $\mathrm{E}(\tilde{\varepsilon}_{n+1}^{v3}) = 0$ and $\mathrm{E}(\tilde{\varepsilon}_{n+1}^{v4}) = 0$. Evidently, the heteroskedasticity-corrected ($\tilde{\varepsilon}$) tests have more asymptotic power than their respective uncorrected ($\varepsilon$) counterparts. For the SVJ0 and SVJ models, the sample estimates of the moment conditions $\mathrm{E}(\tilde{\varepsilon}_{n+1}^{v3}) = 0$ and $\mathrm{E}(\tilde{\varepsilon}_{n+1}^{v4}) = 0$ are found to be positive and significantly different from zero, indicating the possibility of jumps

---

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

[29] 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 et al. (2000). Indeed, there is an expanding empirical literature on estimating multi-factor volatility models. (See, for example, Andersen et al., 2000; Alizadeh et al., 2001; Chernov et al., 1999, 2000, and references therein.)

(with positive mean jump size) in the stochastic-volatility process 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.

Examples of jumps in stochastic volatility can be found in Duffie et al. (2000). Empirical findings with respect to such jumps-in-volatility models can be found in Eraker et al. (2000). Closely related is the issue that under the "square-root" specification, the volatility of volatility cannot increase fast enough during volatile markets, as documented by Jones (1999). Under our setting, a direct test on the volatility of volatility is the goodness-of-fit test associated with the moment condition $E(\varepsilon_n^{v2}) = 0$ (Table 2). Although there is no strong indication that this moment condition is violated for the SVJ0 model, a full resolution relies on an empirical study with a more relaxed volatility of volatility structure.[30]

How would such a mis-specification affect our findings with respect to the importance of jump-risk premia? In an unconditional sense, evidence for jump-risk premia can be found in the "gap" between the volatility implied by the near-the-money short-dated option prices and that observed directly from the spot market. More specifically, the role of jump-risk premia is to make such near-the-money options more expensive, closing the "gap" between the two. Adding jumps to the stochastic volatility, however, will result in a fatter-tailed return distribution, making the near-the-money options even cheaper and therefore widening the "gap." In other words, simply adding jumps in volatility will not replace the role of premia for price jumps. But like the premia for price jumps, one could also incorporate premia for volatility jumps, introducing an additional degree of freedom to explain option prices.[31] These empirical issues, however, are not addressed in this paper, and could be promising for future research.

## 6. Concluding remarks

In this paper, we examined how different risk factors are priced in the S&P 500 index options, and, in doing so, provided strong evidence in support of a jump-risk premium that is highly correlated with the market volatility. We found 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.

---

[30] The CEV model adopted by Jones (1999) is one example, although it does not provide an analytically tractable option pricing formula. Alternatively, multi-factor affine models can be considered.

[31] Statistically, jumps in volatility and jumps in price could result in similar price movements. Investors' aversion to these jump risks, however, could be quite different. For example, Liu et al. (2001) find that these two types of jumps have distinctively different implications for investors asset allocation. Potentially, one could also identify these two types of risk premia using options with different moneyness, which react differently to these two types of jump risks. For example, OTM puts are more sensitive to negative price jumps than OTM calls, and the price jump-risk premia documented in this paper play an important role in capturing this cross-sectional pattern. The cross-sectional analysis of Duffie et al. (2000) indicates that the role of risk-neutral volatility jumps differs from that of risk-neutral price jumps, indicating that the two types of risk premia could be important in capturing the cross-sectional behavior in different ways.

We conclude with 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 found 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 skewness,[32] 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.

## Appendix A. The state-price density

This appendix provides the state-price density that links the data-generating process defined in Eqs. (1) and (2) with the risk-neutral dynamics defined in Eqs. (4) and (5).

Consider a candidate state-price density $\pi$ of the form

$$\pi_t = \exp\left(-\int_0^t r_\tau \, d\tau\right) \varepsilon\left(-\int_0^t \zeta_\tau \, dW_\tau\right) \exp\left(\sum_{i, \, \tau_i \leqslant t} U_i^\pi\right), \tag{A.1}$$

where $\varepsilon(\cdot)$ denotes the stochastic exponential,[33] and where $\zeta$ are the market prices of the Brownian shocks in the price and volatility defined by

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

where $\eta^s$ and $\eta^v$ 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 $U_i^\pi$ are assumed to be *i.i.d.* normal with mean $\mu_\pi$ and variance $\sigma_\pi^2$, and are assumed to be independent of $W$, $W^r$, $W^q$, and inter-jump times. The random jump sizes $U_i^\pi$ and $U_i^s$ are allowed to be correlated (with constant $\rho_\pi$), but are assumed to be independent across different jump times. The most general form of jump-risk

---

[32] Harvey 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).

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

premia is obtained by treating $\mu_\pi$, $\sigma_\pi$, and $\rho_\pi$ as free parameters. In this paper, however, we constrain the mean relative jump size in the state-price density to be zero. That is, $\mu_\pi + \sigma_\pi^2/2 = 0$. This constraint is, in fact, translated to a zero jump-timing risk premium. (Similarly, if we were to turn off the correlation between $U_i^\pi$ and $U_i^s$ by letting $\rho_\pi = 0$, the jump-size risk premium would be zero.)

We now show that Eq. (A.1) indeed defines a state-price density. Let

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

be the total gain processes generated by holding one unit of the underlying security and one dollar in the bank account, respectively. For $\pi$ to be a state-price density, the deflated processes $\mathscr{S}^\pi = \pi \mathscr{S}$ and $\mathscr{B}^\pi = \pi \mathscr{B}$ are required to be local martingales. (One can show that this indeed rules out arbitrage opportunities involving $\mathscr{S}$ and $\mathscr{B}$, under natural conditions on dynamic trading strategies. See, for example, Appendix B.2 in Pan, 2000.) To see that $\mathscr{S}^\pi$ and $\mathscr{B}^\pi$ are indeed local martingales, we apply Ito's Formula

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

$$d\mathscr{B}_t^\pi = - \zeta_t^s \mathscr{B}_t^\pi \, dW_t^s - \zeta_t^v \mathscr{B}_t^\pi \, dW_t^v + \left[ \exp\left( U_{\mathscr{N}_t}^\pi \right) - 1 \right] \mathscr{B}_{t-}^\pi d\mathscr{N}_t,$$

where $\mathscr{N}_t$ 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 $\mathscr{S}^\pi$ and $\mathscr{B}^\pi$ are in fact local martingales, by using the fact that, for any $i \geqslant 1$, $U_i^\pi$ and $U_i^s$ are independent of $\{V_t\}$ and that $E[\exp(U_i^\pi + U_i^s) - 1] = \mu^*$ and $E[\exp(U_i^\pi) - 1] = 0$.

Finally, to link the state-price density $\pi$ with the risk-neutral dynamics defined in Eqs. (4) and (5), we define a density process $\xi_t = \pi_t \exp(\int_0^t r_s \, ds)$. Applying Ito's Formula, one can show that $\xi$ is a local martingale. If $\xi$ is actually a martingale, then $\xi$ uniquely defines an equivalent martingale measure $Q$. In fact, letting

$$W_t(Q) = W_t + \int_0^t \zeta_s \, ds, \quad 0 \leqslant t \leqslant T, \tag{A.3}$$

one can show that the dynamics of $(S, V)$ under $Q$ are indeed in the form of the risk-neutral dynamics defined Eqs. (4) and (5).

## Appendix B. Option pricing and numerical integration

This appendix provides option pricing under the risk-neutral dynamics specified in Eqs. (4) and (5). To facilitate our analyses of the constant component of the jump-arrival intensity, we assume the intensity to be $\lambda_0 + \lambda_1 V_t$, for some non-negative constants $\lambda_0$ and $\lambda_1$.

For any $c \in \mathbb{C}$, the time-$t$ conditional transform of $\ln S_T$, when well defined, is given by

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

Under certain integrability conditions (Duffie et al., 2000),

$$\psi^\vartheta(c,\, \upsilon,\, r,\, q,\, \tau) = \exp\left( \alpha(c,\, \tau,\, \vartheta,\, \theta_r,\, \theta_q) + \beta_\upsilon(c,\, \tau,\, \vartheta)\upsilon \right.$$
$$\left. + \beta_r(c,\, \tau,\, \theta_r)r + \beta_q(c,\, \tau,\, \theta_q)_q \right), \tag{B.1}$$

where $\alpha = \alpha_\upsilon + \alpha_r + \alpha_q$. Letting $\gamma_r^2 = \kappa_r^2 + 2(1-c)\sigma_r^2$ and $\gamma_q^2 = \kappa_q^2 + 2c\sigma_q^2$, the coefficients in Eq. (B.1) $\alpha_r$, $\beta_r$, $\alpha_q$, $\beta_q$, $\alpha_\upsilon$, and $\beta_\upsilon$ are defined by

$$\beta_r(c,\, t,\, \theta_r) = -\frac{2(1-c)\left(1 - \exp\left(-\gamma_r t\right)\right)}{2\gamma_r - (\gamma_r - \kappa_r)\left(1 - \exp\left(-\gamma_r t\right)\right)},$$

$$\alpha_r(c,\, t,\, \theta_r) = -\frac{\kappa_r \bar{r}}{\sigma_r^2}\left( (\gamma_r - \kappa_r)\tau + 2\ln\left[ 1 - \frac{\gamma_r - \kappa_r}{2\gamma_r}(1 - \mathrm{e}^{-\gamma_r \tau}) \right] \right), \tag{B.2}$$

$$\beta_q(c,\, t,\, \theta_q) = -\frac{2c\left(1 - \exp\left(-\gamma_q t\right)\right)}{2\gamma_q - (\gamma_q - \kappa_q)\left(1 - \exp\left(-\gamma_q t\right)\right)},$$

$$\alpha_q(c,\, t,\, \theta_q) = -\frac{\kappa_q \bar{q}}{\sigma_q^2}\left( (\gamma_q - \kappa_q)\tau + 2\ln\left[ 1 - \frac{\gamma_q - \kappa_q}{2\gamma_q}(1 - \mathrm{e}^{-\gamma_q \tau}) \right] \right), \tag{B.3}$$

$$\beta_\upsilon(c,\, t,\, \vartheta) = -\frac{a\left(1 - \exp\left(-\gamma_\upsilon t\right)\right)}{2\gamma_\upsilon - (\gamma_\upsilon + b)\left(1 - \exp\left(-\gamma_\upsilon t\right)\right)},$$

$$\alpha_\upsilon(c,\, t,\, \vartheta) = -\frac{\kappa_\upsilon^* \bar{\upsilon}^*}{\sigma_\upsilon^2}\left( (\gamma_\upsilon + b)\tau + 2\ln\left[ 1 - \frac{\gamma_\upsilon + b}{2\gamma_\upsilon}(1 - \mathrm{e}^{-\gamma_\upsilon \tau}) \right] \right)$$
$$+ \lambda_0 t\left( \exp\left( c\mu_J^* + \frac{c^2 \sigma_J^2}{2} \right) - 1 - c\mu^* \right), \tag{B.4}$$

where $b = \sigma_\upsilon \rho c - \kappa_\upsilon^*$, $a = c(1-c) - 2\lambda_1\left[\exp\left(c\mu_J^* + c^2\sigma_J^2/2\right) - 1 - c\mu^*\right]$, and $\gamma_\upsilon = \sqrt{b^2 + a\sigma_\upsilon^2}$. The parameters superscripted by $*$ denote the risk-neutral counterparts of those under the data-generating measure $P$. For example, $\kappa_\upsilon^* = \kappa_\upsilon - \eta^\upsilon$ and $\bar{\upsilon}^* = \kappa_\upsilon \bar{\upsilon}/\kappa_\upsilon^*$ 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 $\mu_J$. 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, $\sqrt{z} = |z|^{1/2}\exp\left(\mathrm{i}\arg(z)/2\right)$ and $\ln(z) = \ln|z| + \mathrm{i}\arg(z)$, where for any $z \in \mathbb{C}$, $\arg(z)$ is defined such that $z = |z|\exp\left(\mathrm{i}\arg(z)\right)$, with $-\pi < \arg(z) \leqslant \pi$.

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 $\tau_t$ can be calculated as

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

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

$$f(v, \vartheta, r, q, \tau, k) = \mathscr{P}_1 - k\mathscr{P}_2, \tag{B.5}$$

with

$$\mathscr{P}_1 = \frac{\psi(1, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}\big(\psi(1 - iu, v, r, q, \tau)e^{iu(\ln k)}\big)}{u} \, du,$$

$$\mathscr{P}_2 = \frac{\psi(0, v, r, q, \tau)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}\big(\psi(-iu, v, r, q, \tau)e^{iu(\ln k)}\big)}{u} \, du, \tag{B.6}$$

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

The improper probabilities $\mathscr{P}_1$ and $\mathscr{P}_2$ defined by Eq. (B.6) are key to determining the time-$t$ price $C_t$ of an option with time $\tau$ to expiration and strike-to-spot ratio $k$. This appendix provides a fast numerical scheme, with error analysis, for the inversion Eq. (B.6), assuming that the transform $\psi(c, v, r_t, q_t, \tau)$ defined by Eq. (B.1) 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 $\mathscr{P}_1 = \psi(1)\tilde{\mathscr{P}}_1$, where as can be seen from the CIR discount formula,

$$\psi(1) = \mathrm{E}_t\left[\exp\left(-\int_t^{t+\tau} q_s \, ds\right)\right] = \exp\big(\alpha_q(1, \tau, \theta_q) + \beta_q(1, \tau, \theta_q)q_t\big), \tag{B.7}$$

where $\alpha_q$ and $\beta_q$ are as defined in Eq. (B.3). Effectively, $\psi(1)$ is the dividend analogue of a "$\tau$-period bond price." Thus defined $\tilde{\mathscr{P}}_1$ is a real probability that can be calculated through the standard Lévy inversion formula

$$\tilde{\mathscr{P}}_1 = \mathrm{P}\big(\tilde{X}_1 \leqslant \bar{x}\big) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\operatorname{Im}\big(\tilde{\psi}_1(u) \exp(-iu\bar{x})\big)}{u} \, du, \tag{B.8}$$

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

$$\tilde{\psi}_1(u) = \frac{\psi(1 - iu) \exp(iu(r_t - q_t)\tau)}{\psi(1)}. \tag{B.9}$$

In practice, the Lévy inversion Eq. (B.8) is carried out via some form of numerical integration. Letting $I_1(u) = \operatorname{Im}(\tilde{\psi}_1(u) \exp(-iu\bar{x}))$ denote the integrand, we approximate by

$$\tilde{\mathscr{P}}_1 \approx \frac{1}{2} - \frac{1}{\pi} \sum_{n=0}^{[U_1/\Delta u_1]} \frac{I_1\big((n + 1/2)\Delta u_1\big)}{n + 1/2}, \tag{B.10}$$

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

$$\text{truncation error} = \left| \frac{1}{\pi} \int_{U_1}^{\infty} \frac{I_1(u)}{u} \, du \right| \leqslant \delta. \tag{B.11}$$

We can select a step size $\Delta u_1$ such that

$$\text{discretization error} \leqslant \max \left[ P\left( \tilde{X}_1 < \bar{x} - \frac{2\pi}{\Delta u_1} \right), \, P\left( \tilde{X}_1 > \bar{x} - \frac{2\pi}{\Delta u_1} \right) \right] \leqslant \delta, \tag{B.12}$$

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 $u$, $|I_1(u)| \leqslant \exp(-uA_1 + A_0)$ 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 $\delta$, we can therefore choose $U_1$ such that

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

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

$$P\left( |\tilde{X}_1 - \mu_{X_1}| > \frac{\sigma_{X_1}}{\sqrt{\delta}} \right) < \delta. \tag{B.13}$$

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 $\Delta u_1$ can be chosen such that

$$\frac{2\pi}{\Delta u_1} = \max(\bar{x} - \mu_{X_1}, \, \mu_{X_1} - \bar{x}) + \frac{\sigma_{X_1}}{\sqrt{\delta}}. \tag{B.14}$$

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 $E[\exp(u\tilde{X}_1)] = \tilde{\psi}_1(-iu)$, from which its mean and variance can be derived accordingly.

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

## Appendix C. Large-sample properties of IS-GMM estimators

An inherent feature of exchange-traded options is that certain contract variables, such as time $\tau_n$ to expiration and strike-to-spot ratio $k_n$, vary over time. As the option-implied stochastic volatility $V_n^\vartheta$ depends on $\tau_n$ 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 appendix, 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.[34]

### C.1. Stationarity assumption for contract variables

To motivate our assumptions for the contract variables, we recall from Fig. 2 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.

The nearly periodic feature of $\{\tau_n\}$ makes the usual mixing conditions difficult to justify. For example, suppose that $\{\tau_n\}$ is of the form $(40, 33, 26, 19, 40, 33, 26, 19, \ldots)$. Then on date $n$, depending on where we start initially, $\tau_n$ can be $40, 33, 26,$ or $19$. Effectively, this chain has an infinitely long "memory," contrary to the mixing property.[35] In this paper, we take an alternative approach, and assume that $\{\tau_n\}$ takes only finitely many outcomes, and satisfies a time-stationarity property (See Assumption C.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 $\tau_n$. 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.

An appropriate stationarity assumption for the dynamic behavior of the strike-to-spot ratio $\{k_n\}$ 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

---

[34] For exchange-traded derivatives, this situation of time-varying contract variables almost always arises. In over-the-counter markets, however, contract variables on regularly quoted derivative prices are usually constant over time. See Brandt and Santa-Clara (2001) for an application to over-the-counter derivatives.

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

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.

## C.2. Consistency

We start with a formal definition of the option-implied volatility introduced in Section 3.1. 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 Eq. (3), in that $\Xi$ is the maximal set for which a mapping $g : \Xi \to \mathbb{R}_+$ is uniquely defined by

$$f(g(c, \vartheta, r, q, \tau, k), \vartheta, r, q, \tau, k) = c \tag{C.1}$$

for all $(c, \vartheta, r, q, \tau, k) \in \Xi$. We suppose that the parameter space $\Theta$ 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 $\Theta$, 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 $\vartheta \in \Xi$, we can therefore define the date-$n$ option-implied volatility by

$$V_n^\vartheta = g(c_n, \vartheta, r_n, q_n, \tau_n, k_n). \tag{C.2}$$

We next establish the link between the option-implied volatility $V_n^\vartheta$ and the true volatility state variable $V_n$ by letting $V_n^\vartheta = v(V_n^\varDelta, \vartheta, r_n^\varDelta, q_n^\varDelta, \tau_n, k_n)$, where $v : \mathbb{R}_+ \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is defined by

$$v(v, \vartheta, r, q, \tau, k) = g(f(v, \vartheta_0, r, q, \tau, k), \vartheta, r, q, \tau, k), \tag{C.3}$$

where $g$ is defined by Eq. (C.1), using the fact that $c_n = f(V_n^\varDelta, \vartheta_0, r_n^\varDelta, q_n^\varDelta, \tau_n, k_n)$. We note that $v(v, \vartheta_0, r, q, \tau, k) = v$.

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 $\tau$, and letting $Y_n = \tau_{(n, n_v)}$ denote the "$n_v$-history" of $\tau$, we write

$$H(X_n, \vartheta, Y_n) = h\big(y_{(n, n_y)}, v(V_{(n, n_v)}, \vartheta, r_{(n, n_v)}, q_{(n, n_v)}, \tau_{(n, n_v)}, k_{(n, n_v)}), \vartheta\big), \tag{C.4}$$

where $r_{(n, n_v)} = [r_n, r_{n-1}, \ldots, r_{n-n_v+1}]$, and, analogously, $q_{(n, n_v)}, k_{(n, n_v)}$, and $\tau_{(n, n_v)}$ are the $n_v$-dimensional vectors consisting of $q_n, k_n, \tau_n$, 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 C.1. (Time stationarity of Y).* $\{Y_n\}$ has finitely many outcomes, denoted $\{1, 2, \ldots, I\}$. For each outcome $i$ and each positive integer $N$, let $A_N^{(i)} = \{n \leqslant N : Y_n = i\}$ 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

$$\lim_N \frac{\#A_N^{(i)}}{N} = w_i \quad \text{a.s.,} \tag{C.5}$$

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 $X_n = [y_{(n, n_v)}, V_{(n, n_y)}, r_{(n, n_v)}, q_{(n, n_v)}, k_{(n, n_v)}]$ is geometrically ergodic, since it includes only finitely many lags of the joint process.

*Assumption C.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

$$\sup_{\vartheta \in \Theta} \left| G_N^{(i)}(\vartheta) - G_\infty^{(i)}(\vartheta) \right| \to 0 \quad \text{a.s.} \tag{C.6}$$

Given the pointwise-SLLN portion of Assumption C.2, in order to establish the uniform SLLN of Assumption C.2, it is typical to assume some form of Lipschitz condition on $H(x, \vartheta, i)$ as a function of $\vartheta$. 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 C.1. (USLLN of $H(X, \vartheta, Y)$).* Under Assumptions C.1 and C.2, for each $\vartheta$, $G_\infty(\vartheta) = \lim_N G_N(\vartheta)$ exists, and

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

where $G_N(\vartheta)$, defined by Eq. (5), is the sample moment of the observation function.

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

*Assumption C.3 (Convergence of weighting matrices).* $\mathscr{W}_N \to \mathscr{W}_0$ almost surely for some constant symmetric positive-definite matrix $\mathscr{W}_0$.

Under Assumption C.3 and the conditions of Proposition C.1, the criterion function $C_N(\vartheta) = G_N(\vartheta)^\top \mathscr{W}_N G_N(\vartheta)$ converges almost surely to the asymptotic criterion function $C : \Theta \to \mathbb{R}$ defined by $C_N(\vartheta) = G_\infty(\vartheta)^\top \mathscr{W}_N G_\infty(\vartheta)$. In particular, we have $G(\vartheta_0) = 0$, given the moment condition Eq. (2), the pointwise-SLLN portion of Proposition C.1, and the fact that $V_n^{\vartheta_0} = V_n^\Delta$.

*Assumption C.4 (Uniqueness of minimizer).* $C(\vartheta) \neq G(\vartheta_0)$, $\vartheta \in \Theta$, $\vartheta \neq \vartheta_0$.

*Theorem C.1* (Strong consistency). *Under Assumptions C.1–C.4. the IS-GMM $\{\vartheta_N\}$ estimator converges to $\vartheta_0$ almost surely as $N \to \infty$.*

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

### C.2.1. Asymptotic normality

Next, we establish asymptotic normality for the IS-GMM estimator, allowing for time-varying contract variables. Because $v(v, \vartheta_0, r, q, \tau, k) = v$, the sample moment $G_N(\vartheta)$ evaluated at the true parameter $\vartheta_0$ 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 C.5* (CLT). $\sqrt{N}G_N(\vartheta_0)$ converges in distribution as $N \to \infty$ to a normal random vector with mean zero and some covariance matrix $\Sigma_0$.

This assumption follows immediately from the geometric ergodicity of $(y, V)$ and an assumption of integrability of $\left\| h(y_{(n, n_y)}, V_{(n, n_v)}) \right\|^{2+\delta}$, 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; 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 $\vartheta_0$, 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 $\vartheta$, defined by

$$d(\vartheta, X_n, Y_n) = \frac{\partial}{\partial \vartheta} h\left(y_{(n, n_y)}, V^\vartheta_{(n, n_v)}, \vartheta\right)$$

$$+ \sum_{i=n-n_v+1}^{n} \frac{\partial}{\partial v_i} h\left(y_{(n, n_y)}, V^\vartheta_{(n, n_v)}, \vartheta\right) g_\vartheta\left(c_i, \vartheta, r^A_i, q^A_i, \tau_i, k_i\right), \quad \text{(C.7)}$$

where $g_\vartheta(c, \vartheta, r, q, \tau, k) = \vartheta_g(c, \vartheta, r, q, \tau, k)/\partial\vartheta$, with $g$ defined by Eq. (C.1), and where $c_i = f(V^A_i, \vartheta_0, r^A_i, q^A_i, \tau_i, k_i)$. The first term on the right-hand side of Eq. (C.7) arises from the explicit dependence of $h$ on $\vartheta$, while the second term arises from the dependence of $h$ on $V_i$ and the dependence of $V^\vartheta_i = g(c_i, \vartheta, r^A_i, q^A_i, \tau_i, k_i)$ on $\vartheta$, for $i \in \{n - n_v + 1, \dots, n\}$. This second term is important in identifying risk-premium parameters such as $\eta^v$. Intuitively, such parameters are identified by exploring the option-pricing relation through $V^\vartheta$.

*Assumption C.6* (Convergence of "Jacobian estimator"). For some constant $(n_h \times n_\vartheta)$ matrix $d_0$ of rank $n_\vartheta$: (i) $1/N \sum_{n \leqslant N} d(\vartheta_0, X_n, Y_n)$ converges in probability as $N \to \infty$ to $d_0$. (ii) For any $\{\vartheta_n\}$ converging in probability as $n \to \infty$ to $\vartheta_0$, $1/N \sum_{n \leqslant N} d(\vartheta_n, X_n, Y_n)$ converges in probability as $N \to \infty$ to $d_0$.

Part (i) of Assumption C.6 follows from geometric ergodicity of $X$, independence and time-stationary of $Y$, and integrability (over the stationary distribution of $X$) of $d(\vartheta_0, X_n, i)$, for each $i$. Given that part (i) holds, part (ii) follows from assuming first-moment continuity (as in Hansen, 1982) of $d(\vartheta_n, X_n, Y_n)$ at $\vartheta_0$.

*Theorem C.2* (Asymptotic normality). *Under Assumptions C.1–C.6, $\sqrt{N}(\vartheta_n - \vartheta_0)$ converges in distribution as $N \to \infty$ to a normal random vector with mean zero and covariance matrix*

$$\Lambda = \left(d_0^\top \mathscr{W}_0 d_0\right)^{-1} d_0^\top \mathscr{W}_0 \Sigma_0 \mathscr{W}_0 d_0 \left(d_0^\top \mathscr{W}_0 d_0\right)^{-1}. \tag{C.8}$$

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

For the usual two-step GMM of Hansen, 1982, under which the distance matrices are chosen so that $\mathscr{W}_0 = \Sigma_0^{-1}$, 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 $n_\vartheta$, and $\mathscr{W}_0$ is the identify matrix), so $\Lambda = d_0^\top \Sigma_0 (d_0^\top)^{-1}$.

## Appendix D. A recursive formula for conditional moments

To facilitate our analyses of the constant component of the jump-arrival intensity, we assume the intensity to be $\lambda_0 + \lambda_1 V_t$, for some non-negative constants $\lambda_0$ and $\lambda_1$. Under certain integrability conditions (Duffie et al., 2000), the date-$n$ conditional moment-generating function of $(y_{n+1}, V_{n+1})$ can be derived as $\phi(u_y, u_v, V_n)$, with $\phi$ defined by

$$\phi(u_y, u_v, v) = \exp\left(A(u_y, u_v) + B(u_y, u_v)v\right), \tag{D.1}$$

where, letting $b = \sigma_v \rho u_y - \kappa_v$, $a = -u_y^2 - 2u_y[\eta^s - 1/2 - \lambda_1 \mu^*] - 2\lambda_1(\exp(u_y \mu_J + u_y^2 \sigma_J^2/2) - 1)$, and $\gamma = \sqrt{b^2 + a\sigma_v^2}$, the coefficients $A$ and $B$ are defined by

$$B(u_y, u_v) = -\frac{a(1 - \exp(-\gamma\Delta)) - u_v[2\gamma - (\gamma - b)(1 - \exp(-\gamma\Delta))]}{2\gamma - (\gamma + b)(1 - \exp(-\gamma\Delta)) - u_v \sigma_v^2(1 - \exp(-\gamma\Delta))},$$

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

From the explicitly known moment-generating function $\phi$, conditional moments of $(y, V)$ can be derived using Eq. (8). In this section, we provide the following

recursive scheme that turns out to be very useful in practice,[36]

$$\mathrm{E}_t\left(y_{t+\Delta}^0 V_{t+\Delta}^m\right) = \sum_{j=0}^{m-1} C_{m-1}^j \mathrm{E}_t\left(y_{t+\Delta}^0 V_{t+\Delta}^j\right) p_{y,v}^{(0,m-j)}(V_t), \quad m \geqslant 1,$$

$$\mathrm{E}_t\left(y_{t+\Delta}^n V_{t+\Delta}^m\right) = \sum_{i=0}^{n-1} \sum_{j=0}^{m} C_{n-1}^i C_m^j \mathrm{E}_t\left(y_{t+\Delta}^i V_{t+\Delta}^j\right) p_{y,v}^{(n-i,m-j)}(V_t),$$
$$n \geqslant 1, \ m \geqslant 0, \tag{D.3}$$

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

$$p_{y,v}^{(i,j)}(V_t) = A_{y,v}^{(i,j)} + B_{y,v}^{(i,j)} V_t, \tag{D.4}$$

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 $i \geqslant 0$ and $j \geqslant 0$. 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

$$B_{y,v}^{(0,m)} = \frac{m}{2} f_0 \sigma_v^2 B_{y,v}^{(0,m-1)}, \quad m \geqslant 2, \quad B_{y,v}^{(n,0)} = \lambda_1 J_n f_0 - \frac{1}{2}\sum_{i=1}^{n-1} C_n^i B_{y,v}^{(i,0)} f_{n-i}, \quad n \geqslant 3,$$

$$B_{y,v}^{(n,1)} = \frac{1}{2}\kappa f_0 f_n + \frac{1}{2}\sigma_v^2 f_0 B_{y,v}^{(n,0)} - \frac{1}{2}\sum_{i=1}^{n-1} C_n^i f_i B_{y,v}^{(n-i,1)}, \quad n \geqslant 2,$$

$$B_{y,v}^{(n,m)} = \frac{m}{2}\sigma_v^2 f_0 B_{y,v}^{(n,m-1)} - \frac{1}{2}\sum_{i=1}^{n} C_n^i f_i B_{y,v}^{(n-i,m)}, \quad n \geqslant 1, \quad m \geqslant 2,$$

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 $n \geqslant 3$) are the moments of the jump amplitude. The coefficients $f_i$ and $g_i$ are given by

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

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

where

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

---

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

and

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

$$\gamma_2 = -\gamma_1^2 - \left(\frac{\sigma_v}{\kappa}\right)^2(1 - \rho^2 + \lambda_1 J_2),$$

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

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

$$A_{y,v}^{(0,n)} = \frac{n-1}{2}\sigma_v^2 f_0 A_{y,v}^{(0,n-1)}, \, n \geqslant 2, \quad A_{y,v}^{(n,0)} = \lambda_0 J_n^0\Delta - \frac{\kappa\bar{v}}{\sigma_v^2}\left(\kappa\gamma_n\Delta + \hat{f}_n - \hat{g}_n\right), \, n \geqslant 2,$$

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

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

where for $n \geqslant 1$, $\hat{C}_n^0(m) = m$, $\hat{C}_n^n(m) = 1$, 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

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

## Appendix E. Appendix to Section 5

### E.1. Tests of moment conditions

Our tests of moment conditions follow from the tests of orthogonality conditions developed in Eichenbaum et al. (1988), and are also closely related to the Hansen (1982) test of over-identifying restrictions (Lemma 4.1 of Hansen, 1982). Let $E_n(\varepsilon_{n+1}) = 0$ 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 $E_n(\mathscr{H}_{n+1}) = 0$. To test $E_n(\varepsilon_{n+1}) = 0$ we construct its sample analogue by

$$\mathscr{G}_N\left(\hat{\vartheta}_N\right) = \frac{1}{N}\sum_{n \leqslant N}\varepsilon_n\left(\hat{\vartheta}_N\right), \tag{E.1}$$

where $\varepsilon_n(\hat{\vartheta}_N)$ denotes evaluating the moments $\varepsilon$ at the IS-GMM estimator $\hat{\vartheta}_N$. Using arguments similar to those following Assumption C.5 in Section 3, one can show that, under typical technical regularity conditions, $\sqrt{N}\mathscr{G}_N(\hat{\vartheta}_0)$ is asymptotically normal. Applying a standard mean-value expansion,

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

where $\bar{\vartheta}_N^j$ is can be shown between $\vartheta_0^j$ and $\vartheta_N^j$, for $j \in \{1, \ldots, n_\vartheta\}$. Moreover, for sufficiently large $N$ and with probability arbitrarily close to one, we can write

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

where $G_N = (N)^{-1}\Sigma_n\mathscr{H}_n$ is the sample analogue of the "optimal" moments. We know that $\partial G_N(\bar{\vartheta}_N)/\partial\vartheta$ converges to a constant full-rank matrix $d_0$ in probability, under Assumption C.6, using the fact that $\hat{\vartheta}_N$ is estimated under an exactly identified IS-GMM setting.

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

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

where $\approx^a$ means "asymptotically equivalent in distribution to." Thus, $\mathscr{G}_N(\hat{\vartheta}_N)$ is asymptotically normal with some covariance matrix $\Omega$. An estimator $\Omega_N$ of $\Omega$ can be obtained by estimating the covariance matrix of the right-hand side of (E.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 $\sqrt{N}\mathscr{G}_N^i(\hat{\vartheta}_N)/\sqrt{(\Omega_N)_{ii}}$ is asymptotically standard normal. We can test any subgroup of moment conditions, indexed by $I$, by using the fact that, in large sample, $N(\mathscr{G}_N^I(\hat{\vartheta}_N))^\top/((\Omega_N)_I)^{-1}\mathscr{G}_N^I(\hat{\vartheta}_N)$ is distributed as a $\chi^2$ random variable with $\#(I)$ degrees of freedom.

### E.2. Estimation of interest rates and dividend yields

For the purpose of estimating the respective parameter vectors $\theta_r$ and $\theta_q$ of the short-rate process $r$ and the dividend-rate process $q$ defined by (3), 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 $\Delta$, and taking advantage of the fact that the conditional density of $q_n$ given $q_{n-1}$ is that of a non-central $\chi^2$ (Feller, 1951; Cox et al., 1985), we estimate $\theta_q$ 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 et al., 1985)

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

Table 6
ML estimates of interest rates $r$ and dividend yields $q$ [a]

| $\kappa_r$ | $\bar{r}$ | $\sigma_r$ | $\kappa_q$ | $\bar{q}$ | $\sigma_q$ |
|---|---|---|---|---|---|
| 0.20 | 0.058 | 0.0415 | 0.24 | 0.025 | 0.0269 |
| (0.15) | (0.016) | (0.0009) | (0.33) | (0.011) | (0.0004) |

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

Table 7
A Monte Carlo study[a]

|  | $\kappa_v$ | $\bar{v}$ | $\sigma_v$ | $\rho$ | $\eta^s$ | $\eta^v$ | $\lambda$ | $\sigma_J$ (%) | $\mu^*$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| True | 6.5 | 0.0150 | 0.30 | −0.50 | 3.5 | 3.0 | 12.0 | 3.0 | −19 |
| Mean | 6.6 | 0.0153 | 0.30 | −0.50 | 3.4 | 3.5 | 12.8 | 3.8 | −19 |
| Std | 1.4 | 0.0028 | 0.02 | 0.04 | 3.0 | 2.6 | 3.5 | 2.6 | 3 |

[a] The mean and standard deviation are calculated from a sample of 100 Monte Carlo simulations. For each simulation and estimation, the mean relative jump size $\mu$ is fixed at −0.8. Data: exactly the same data structure as the time series $S_n$, $C_n$, $C_n^{ITM}$ used for the SVJ model estimation in Table 3, with the state variables $S_n$, $V_n$ simulated using the "true" model parameters.

where $\theta_r^0$ denotes the true parameter vector, and where $\alpha_r$ and $\beta_r$ are as defined in (B.2). The one-period conditional density $p^R(\cdot|R_{n-1}; \theta_r)$ of $R_n$ given $R_{n-1}$ is therefore given by

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

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 $\chi^2$.

The ML estimates of $\theta_r$ and $\theta_q$ 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.

### E.3. A Monte-Carlo simulation study

We perform a Monte-Carlo study to examine the small sample properties of the IS-GMM estimators. Using the "true" model parameters summarized in Table 7, we simulate a joint time series of state variables $(S,\ V)$ with weekly frequency and of the same length as the real data, from which we price time series $\{Cn,\ C_n^{ITM}\}$ of near-the-money and in-the-money short-dated call options. The data structure, including option maturities and moneyness, is exactly the same as that used for the SVJ$^{ITM}$ model estimation in Table 3.

Given that jump parameters are hard to pin down under the actual dynamics, we set the mean relative jump size $\mu$ at −0.8% throughout this Monte-Carlo study. The results obtained from 100 simulations are summarized in Table 7. Comparing the

Monte Carlo results with the IS-GMM results reported in Table 3 for SVJ$^{ITM}$, we see that the standard errors obtained from the large-sample theory work reasonably well in the small sample. It should be noted, however, that the standard errors reported for the jump-related estimates are less accurate than those associated with the diffusive moments.

## References

Aït-Sahalia, Y., Lo, A., 2000. Nonparametric risk management and implied risk aversion. Journal of Econometrics 94, 9–51.

Aït-Sahalia, Y., Wang, Y., Yared, F., 2001. Do option markets correctly price the probabilities of movement of the underlying asset?. Journal of Econometrics 102, 67–110.

Alizadeh, S., Brandt, M., Diebold, F., 2001. High- and low-frequency exchange rate volatility dynamics: range-based estimation of stochastic volatility models. Working paper, Wharton School, University of Pennsylvania.

Andersen, T., Benzoni, L., Lund, J., 1998. Estimating jump-diffusions for equity returns. Working paper, Kellogg Graduate School of Management, Northwestern University.

Andersen, T., Bollerslev, T., Diebold, F., Ebens, H., 2001. The distribution of stock return volatility. Journal of Financial Economics 61.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2000. The distribution of realized exchange rate volatility. Journal of the American Statistical Association, forthcoming.

Andrews, D.W.K., 1987. Consistency in nonlinear econometric models: a generic uniform law of large numbers. Econometrica 55, 1465–1471.

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

Bakshi, G., Kapadia, N., 2001. Delta-hedged gains and the pricing of volatility risk. Working paper, University of Massachusetts, Amherst.

Bakshi, G., Madan, D., 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, Kellog Graduate School of Management, Northwestern University.

Black, F., 1976. Studies of stock price volatility changes. Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, pp. 177–181.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Bollerslev, T., Mikkelsen, H.O., 1996. Modeling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151–184.

Brandt, M., Santa-Clara, P., 2001. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. Journal of Financial Economics, forthcoming.

Brémaud, P., 1981. Point Processes and Queues, Martingale Dynamics. Springer, New York.

Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 1999. A new class of stochastic volatility models with jumps: theory and estimation. Working paper, Columbia University, University of North Carolina, and Duke University.

Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2000. Alternative models for stock price dynamics. Working paper, Columbia University, University of North Carolina, and Duke University.

Chernov, M., Ghysels, E., 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, 407–458.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

Das, S.R., Sundaram, R.K., 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.

Detemple, J., Selden, L., 1991. A general equilibrium analysis of option and stock market interactions. International Economic Review 32, 279–303.

Doob, J., 1953. Stochastic Processes. Wiley, New York.

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

Duffie, D., Singleton, K., 1993. Simulated moments estimation of markov models of asset prices. Econometrica 61, 929–952.

Duffie, D., Singleton, K., 1999. Modeling term structure of defaultable bonds. Review of Financial Studies 12, 687–720.

Eichenbaum, M.S., Hansen, L.P., Singleton, K.J., 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., 2000. Do stock prices and volatility jump? reconciling evidence from spot and option prices. Working paper, Graduate School of Business, University of Chicago.

Eraker, B., Johannes, M.S., Polson, N.G., 2000. The impact of jumps in returns and volatility. Working paper, University of Chicago.

Feller, W., 1951. Two singular diffusion problems. Annals of Mathematics 54, 173–182.

Gallant, A.R., Tauchen, G., 1998. Reprojecting partially observed systems with application to interest rate diffusions. Journal of American Statistical Association 93, 10–24.

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

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, Princeton.

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.

Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263–1295.

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.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.

Jackwerth, J.C., 2000. Recovering risk aversion from options prices and realized returns. Review of Financial Studies 13, 433–451.

Jackwerth, J.C., Rubinstein, M., 1996. Recovering probability distributions from option prices. Journal of Finance 51, 1611–1631.

Johannes, M., Kumar, R., Polson, N., 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.

Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risk assets. The Journal of Finance 31, 1085–1100.

Liu, J., 1997. Generalized method of moments estimation of affine diffusion processes. Working paper, Graduate School of Business, Stanford University.

Liu, J., Longstaff, F., Pan, J., 2001. Dynamic asset allocation with event risk. Working paper, UCLA Anderson School and MIT Sloan School.

Lucas, R.E., 1978. Asset prices in an exchange economy. Econometrica 46, 1429–1445.

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

Pan, J., 2000. Jump-Diffusion Models of Asset Prices: Theory and Empirical Evidence. Ph. D. Thesis, Graduate School of Business, Stanford University.

Pastrorello, S., Renault, E., Touzi, N., 1996. Statistical inference for random variance option pricing. Working paper, Dipartimento di Scienze Economiche, Università di Bologna, Italy.

Pham, H., Touzi, N., 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-Champaign.

Protter, P., 1990. Stochastic Integration and Differential Equations. Springer, New York.

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

Rosenberg, J., Engle, R., 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., 2001. Estimation of affine asset pricing models using the empirical characteristic function. Journal of Econometrics 102, 111–141.

Stein, E., Stein, J., 1991. Stock price distributions with stochastic volatility: an analytic approach. Review of Financial Studies 4, 725–752.

Stein, J., 1989. Overreactions in the options market. Journal of Finance 44, 1011–1023.