Journal of Financial Econometrics, 2006, Vol. 4, No. 3, 353–384

Leverage and Volatility Feedback Effects in High-Frequency Data

Tim Bollerslev
Duke University

Julia Litvinova
The Brattle Group

George Tauchen
Duke University

ABSTRACT

We examine the relationship between volatility and past and future returns using high-frequency aggregate equity index data. Consistent with a prolonged "leverage" effect, we find the correlations between absolute high-frequency returns and current and past high-frequency returns to be significantly negative for several days, whereas the reverse cross-correlations are generally negligible. We also find that high-frequency data may be used in more accurately assessing volatility asymmetries over longer daily return horizons. Furthermore, our analysis of several popular continuous-time stochastic volatility models clearly points to the importance of allowing for multiple latent volatility factors for satisfactorily describing the observed volatility asymmetries.

KEYWORDS: high-frequency data, leverage effect, stochastic volatility models, temporal aggregation, volatility asymmetry, volatility feedback effect

One of the striking empirical regularities to emerge from the burgeoning literature on volatility modeling over the past two decades concerns the apparent asymmetry in the relationship between equity market returns and

We thank Bjorn Eraker, Pete Kyle, and other participants in the Duke Financial Econometrics Lunch Group for many useful discussions and comments. We also thank an associate editor for the Journal and two anonymous referees for many helpful suggestions, which have significantly improved this article. Bollerslev's work was supported by a grant from the NSF to the NBER. Address correspondence to Tim Bollerslev, Department of Economics, Duke University, Box 90097, Durham, NC 27708, or e-mail: boller@econ.duke.edu.

doi:10.1093/jjfinec/nbj014
Advance Access publication May 16, 2006

volatility.[1] At the same time, there is little agreement concerning the fundamental cause(s) behind the observed asymmetry.

Early influential studies by Black (1976) and Christie (1982) attributed the asymmetric return–volatility relationship to changes in financial leverage, or debt-to-equity ratios. However, the magnitude of the effect of a decline in current prices on future volatilities appears too large to be explained solely by changes in financial leverage [see, e.g., Figlewski and Wang (2001)]. Furthermore, counter to a leverage-based explanation, the asymmetry is generally larger for aggregate market index returns than that for individual stocks [see, e.g., Kim and Kon (1994), Tauchen, Zhang, and Liu (1996), and Andersen et al. (2001)].

The other leading explanation for the volatility asymmetry rests on a time-varying risk premium, or volatility feedback effect, as discussed by, for example, French, Schwert, and Stambaugh (1987) and Campbell and Hentschel (1992). If volatility is priced, an anticipated increase in volatility would raise the required rate of return, in turn necessitating an immediate stock-price decline to allow for higher future returns. Therefore, the causality underlying the volatility feedback effect runs from volatility to prices, as opposed to the leverage effect that hinges on the reverse causal relationship.[2]

On comparing the magnitude of the two effects, Bekaert and Wu (2000) and Wu (2001) argued that the volatility feedback effect dominates the leverage effect empirically. However, many other studies [see, e.g., Nelson (1991), Engle and Ng (1993), and Glosten, Jagannathan, and Runkle (1993)] have found that volatility increases more following negative returns than positive returns and that the relationship between expected returns and volatility is insignificant, or even negative empirically.[3]

Meanwhile, the existing empirical literature has focused almost exclusively on the trade-offs observed over daily or longer return horizons. However, as discussed above, from an empirical perspective the fundamental difference between the leverage and volatility feedback explanations lies in the *causality*: the leverage effect explains why a negative return leads to higher subsequent volatility, whereas the volatility feedback effect justifies how an increase in

[1] Anecdotal evidence like the heightened volatility following the October 1987 stock market crash and the more recent turmoil following Russia's default and the Long-Term Capital Management (LTCM) debacle in September 1998, as well as the relatively low volatility accompanying the rapid run-up in prices during the recent tech bubble, are all consistent with this asymmetry.

[2] Although the term is arguably a misnomer, following the extant literature we will continue to refer to a negative correlation between volatility and past and current returns as a "leverage" effect. Tauchen (2005) has recently shown how stochastic volatility and dynamic leverage effects can arise endogenously within the context of a self-contained general equilibrium model in which the sign and the magnitude of the return–volatility correlation depends directly on the coefficient of risk aversion and the intertemporal rate of substitution of a representative consumer.

[3] More recent powerful statistical inference based on long return histories or more efficient volatility measures constructed from high-frequency intraday returns by Lundblad (2005), and Bali and Peng (2006) and Ghysels, Santa-Clara, and Valkanov (2005), respectively, do suggest a significant positive risk-return trade-off relationship.

volatility may result in negative returns. These causal relationships may, of course, appear immediate in lower-frequency data, and hence, the two explanations are indistinguishable. Therefore, this may help explain the often-inconclusive and sometimes-conflicting results reported in the extant literature.

We seek to improve on this through the use of high-frequency five-minute S&P 500 futures returns to differentiate more clearly between the two competing effects. Of course, volatility is not directly observable. Relying on the absolute high-frequency returns as a simple volatility proxy, our results reveal a new and striking prolonged negative correlation between the volatility and the current and lagged returns, which lasts for several days. In contrast, the correlations between the returns and the lagged volatility are all close to zero. These results therefore support the notion of a highly significant prolonged leverage effect at the intradaily level. At the same time, we also observe a very strong contemporaneous correlation between the high-frequency returns and their absolute value, which as noted above could be interpreted as evidence in favor of an almost instantaneous volatility feedback effect.

Taking the analysis one step further, we demonstrate how the high-frequency data can also be used in more accurately measuring the asymmetry in the return–volatility relationship over longer daily horizons.[4] In particular, we derive a simple aggregation formula that under reasonable assumptions decomposes the correlations between the squared daily returns and future and past daily returns into the underlying high-frequency correlations. In contrast to the noisy sample cross-correlations constructed directly from the daily returns, the corresponding daily cross-correlations implied from the intraday returns over the identical time span reveal a much clearer picture in regard to the pertinent volatility asymmetries. Consistent with the leverage effect observed in the high-frequency correlations, the *implied* daily correlations between the squared returns and the current and one- to three-day-lagged returns are all significantly negative.[5]

These new findings naturally raise the question of whether some of the popular stochastic volatility models in the existing literature are compatible with the observed empirical dependencies. The formulation and estimation of continuous-time stochastic volatility models has been an extremely active area of research over the past decade [see, e.g., the recent survey in Tauchen (2004)]. Although the importance of allowing for a contemporaneous leverage effect in the form of a negative correlation between the latent volatility factor(s) and the

[4] This mirrors the use of high-frequency data in more effectively measuring volatility dependencies over longer-run interdaily horizons in Bollerslev and Wright (2000), and the aforementioned use of high-frequency-based realized volatility-type measures in Ghysels, Santa-Clara, and Valkanov (2005) in assessing the risk-return trade-off over multiple days.

[5] Similar asymmetric cross-correlation patterns based on long time series of daily equity returns and averages across multiple return series have previously been documented in independent work in the ecophysics literature by Bouchaud, Matacz, and Potters (2001), Perello, Masoliver, and Anento (2004), and Perello, Masoliver, and Bouchaud (2004). The discrete-time linear ARCH (LARCH) model recently proposed by Giraitis et al. (2004) is also in part motivated by these patterns.

innovations driving the returns in these types of models is now well established, the dynamic implications of doing so have not yet been fully explored.

To this end, we begin by showing that even though the popular one-factor Heston (1993) affine stochastic volatility model with a contemporaneous leverage effect implies a geometrically declining pattern in the cross-correlations between the squared returns and the lagged returns, it is not able to reproduce the corresponding slow decay observed in the actual data. The one-factor stochastic volatility model has, of course, previously been rejected in the literature, and our estimation results based on daily data and the Efficient Method of Moments (EMM) estimation procedure of Gallant and Tauchen (1996) also soundly refute the model. In response to this, more complicated multifactor stochastic volatility models, which break the tight link between tail thickness and volatility persistence inherent in the one-factor model, have recently been proposed in the literature. In line with the results reported by Chernov et al. (2003), we find that a two-factor logarithmic stochastic volatility model does a good job in terms of describing the dependencies in the daily returns when judged by the EMM goodness-of-fit test. However, on simulating artificial high-frequency returns from the estimated daily model, we find that it is not able to reproduce the prolonged leverage effect, or asymmetric cross-correlations, observed in the high-frequency data. In contrast, on estimating the same model with actual high-frequency hourly returns, we find that the resulting model estimates effectively capture the observed cross-correlation patterns at daily, hourly, and five-minute return frequencies. Thus, even though the two-factor hourly model formally fails the EMM goodness-of-fit test, our results clearly illustrate the importance of incorporating multiple volatility factors to describe satisfactorily the dynamic inter- and intraday leverage effects documented here. At a more general level, the results in this article directly highlight the practical importance and utility of using high-frequency intraday data along with proper aggregation formulas in the formulation of richer and empirically more realistic continuous-time models capable of describing salient data features across varying sampling frequencies and return horizons.

The rest of this article is organized as follows. Section 1 describes the striking high-frequency return–volatility cross-correlation patterns. This section also demonstrates how the high-frequency data may be used in more effectively assessing the cross-correlations over longer return horizons through a simple-to-implement temporal aggregation formula. Section 2 discusses the cross-correlations from the popular one-factor Heston model, along with the pertinent volatility asymmetries implied by our daily EMM-based estimates for that model. The daily and hourly estimates for the two-factor logarithmic stochastic volatility models and the accompanying cross-correlation patterns for different return frequencies are discussed in Section 3. Section 4 concludes. All of the mathematical proofs and specific details concerning the EMM estimation results are relegated to three Appendixes.

1 VOLATILITY ASYMMETRIES AND CROSS-CORRELATION PATTERNS

1.1 Data

Our empirical investigations are based on high-frequency tick-by-tick S&P 500 futures data from the Chicago Mercantile Exchange (CME), spanning the time period from January 4, 1988, through March 9, 1999, for 2757 trading days. The S&P 500 futures market is among the most actively traded markets in the world. During the sample period, 7,145,360 transactions occurred, and the average time between two consecutive transactions was only around nine seconds. Transaction costs in futures market are generally also much lower than that in the corresponding cash markets. Indeed, several studies have documented that futures markets tend to lead their underlying cash markets in terms of the price discovery process; see, for example, Hasbrouck (2003). This is particularly important in the present context, as we are concerned with measuring potentially very short-lived lead–lag effects.

The regular trading hours of the CME S&P 500 futures contract extends from 09:30 to 16:15 EST. However, because the cash market in New York closes at 16:00 EST, we restrict our analysis to that time interval. Also, owing to the unusually high volatility at the opening, we omit the first five minutes of each trading day.[6] All in all, this leaves us with 77 five-minute intervals per day. We calculate the continuously compounded returns over each of the five-minute intervals by the logarithmic difference between the two tick prices immediately before the respective five-minute marks; see Dacorogna et al. (2001) for further discussion and verification of this approach. We also construct corresponding "hourly" returns by summing 11 successive five-minute returns, leaving us with seven "hourly" returns per trading day.

Summary statistics for the five-minute, hourly, and daily returns are presented in Table 1. Consistent with the absence of any confounding market microstructure effects and the lack of serial correlation in the high-frequency five-minute returns, the standard deviation of the daily returns is approximately

Table 1 Summary statistics for S&P 500 futures returns, 1988–1999

	Mean	Median	Standard deviation	Skewness	Kurtosis	Sample size
Five-minute	2.16e−06	0.00e−07	0.00095	−0.029	74.947	212,289
Hourly	2.37e−05	1.08e−04	0.00302	−1.068	26.424	19,299
Daily	1.66e−04	1.13e−04	0.00839	−0.055	17.865	2757

[6] All of our results are robust to the inclusion of this five-minute return, although some of the lines in graphs presented below appear slightly more jagged. Details are available upon request.

equal to the standard deviation of the five-minute returns times $\sqrt{77}$. Also, the standard deviation of the hourly returns approximately equals $\sqrt{11}$ times that of the five-minute returns. Although the distribution of the hourly returns appears to be skewed to the left, the sample skewness coefficients for the five-minute and daily returns are both close to zero. Of course, the sample kurtosis is much greater than the normal value of 3 for all three sampling frequencies. The time-series plot of the daily returns in Figure 1 also clearly shows the familiar volatility clustering effect, along with a few occasional very large absolute returns.

We next turn to a discussion of the corresponding cross-correlation patterns and pertinent leverage and volatility feedback effects observed at the daily, hourly, and five-minute sampling frequencies.

1.2 Daily versus High-Frequency Patterns

Volatility is, of course, not directly observable. To circumvent this, we follow the standard approach of relying on the squared or absolute returns over the relevant time interval as a simple, albeit noisy, expost volatility proxy; see, for example,

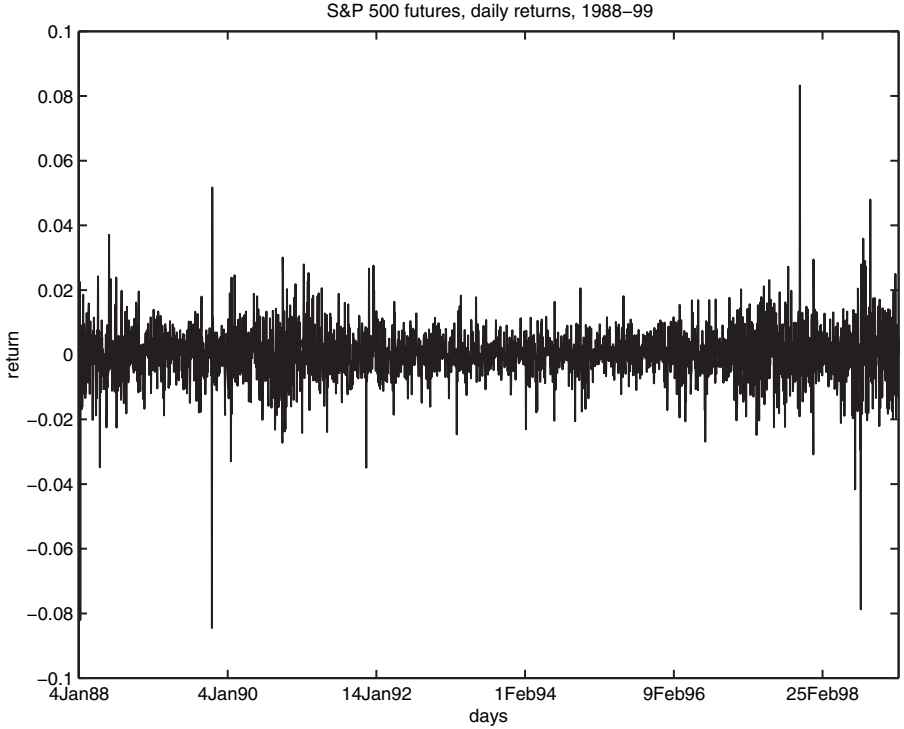

Figure 1 Daily S&P 500 futures return, 1988–1999.

the discussion in Andersen and Bollerslev (1998a).[7] In particular, let $R_{t,t+h} \equiv p_{t+h} - p_t$ denote the continuously compounded returns from time t to $t+h$, where we normalize the unit time interval to one day for notational convenience. The middle panel in Figure 2 then plots the sample cross-correlations between the daily squared returns and the returns for leads and lags ranging from -20 to 20 days,[8]

$$\text{corr}(R^2_{t,t+1}, R_{t-j,t-j+1}) \qquad j = -20, \ldots, 20.$$

The lag 1 and 2 cross-correlations are both negative and, according to the simple Bartlett $1.96/\sqrt{T} \approx 0.038$ confidence bands included as an informal guide, both significant at conventional levels.[9] Surprisingly, however, the contemporaneous cross-correlation is much smaller (in an absolute sense), and the overall pattern conveyed by the daily correlations is somewhat erratic and difficult to interpret.[10] This is, of course, entirely consistent with the conflicting and inconclusive results reported in the extant literature discussed above.

Consider now instead the cross-correlations for the high-frequency absolute five-minute returns depicted in the top panel in the figure

$$\text{corr}(|R_{t,t+\triangle}|, R_{t-j\triangle,t-(j-1)\triangle}) \qquad j = -400, \ldots, 400,$$

where $\triangle = 1/77$ corresponds to the five-minute interval.[11] The figure reveals a distinct and very slow decay in the values of the cross-correlations for positive

[7] Other volatility proxies based on all of the tick-by-tick could in principle be used. However, even though the S&P 500 futures market is one of the most liquid markets in the world, market microstructure frictions, such as price discreteness and nonsynchronous trading, significantly affect the observed price changes at the very highest intradaily sampling frequencies. This in turn complicates the calculation of corresponding volatility measures, and the development of more sophisticated methods for doing so is currently a very active area of research; see, for example, Hansen and Lunde (2006), and the many references therein. Building on the ideas developed in that literature, it might be possible to sharpen the empirical results reported on below.

[8] Following the arguments in Yu (2005b), it would be interesting to extend the present analysis to explore the patterns in the *conditional* cross-correlations, $\text{corr}(R^2_{t,t+1}, R_{t-j,t-j+1} | R_{t-j,t-j+1})$. A nonparametric kernel-based procedure, as adopted by Johannes (2004) in the estimation of conditional kurtosis measures, could be used in doing so.

[9] The corresponding Generalized Method of Moments (GMM)-based heteroskedasticity robust standard errors for the lag 1 and 2 cross-correlations equal 0.0485 and 0.0321, respectively.

[10] A very similar picture was obtained by using the absolute daily returns, $|R_{t,t+1}|$, in place of the squared returns as an alternative volatility proxy. However, for the ease of comparison with our later analytical results, we only present the graph for the daily squared return correlations. A very similar pattern was also obtained by using the daily close-to-close return instead of the open-to-close return. All of these pictures are available upon request.

[11] Owing to the heavy tails in the high-frequency return distribution, we will focus on the more robust absolute return volatility measure, but the same qualitative, albeit somewhat more noisy, pictures are obtained with the high-frequency squared returns.

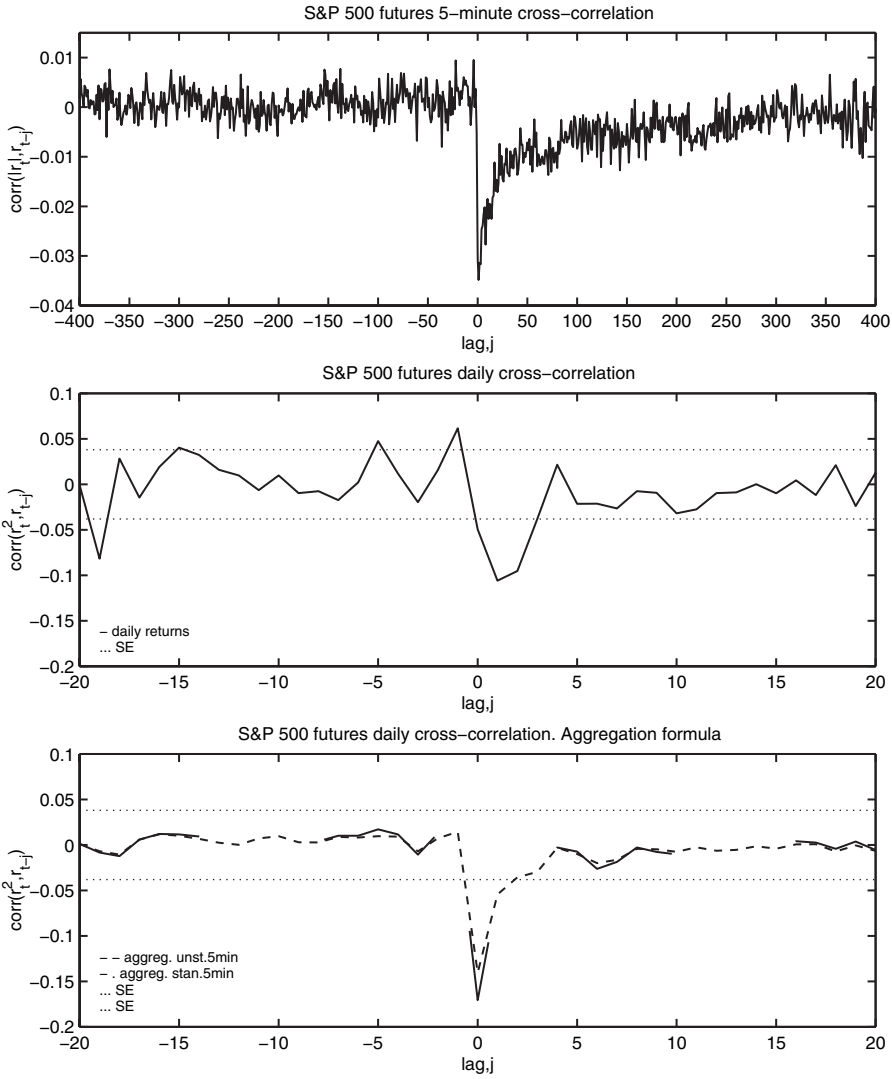

Figure 2 S&P 500 sample cross-correlations. The top panel shows the cross-correlations between the five-minute raw and absolute returns. The middle panel depicts the cross-correlations between the daily and the daily squared returns. The bottom panel shows the implied daily cross-correlations between the daily returns and the daily squared returns calculated from the aggregation formula in Section 1.3. The long dashed line gives the correlations implied by the standardized five-minute returns, whereas the short dashed line is based on the raw five-minute returns. The dotted lines in the two lower panels give the conservative 95% Bartlett confidence bands under the null hypothesis of zero correlations. All of the calculations are based on S&P 500 futures data from 1988 to 1999.

j's, whereas all of the reverse correlations for $j < 0$ tend to be scattered around zero. The contemporaneous effect is clearly the largest, but the impact of current return shocks on the future volatilities lasts for at least five days, or $5 \times 77 = 385$ five-minute intervals. These results strongly support the notion of a prolonged leverage effect, whereas there is little or no evidence for a delayed volatility feedback effect coming from the high-frequency data. The highly significant negative contemporaneous correlation may, of course, be interpreted as evidence in favor of an instantaneous volatility feedback effect.

The contrast between the much sharper picture revealed by the high-frequency cross-correlations in the top panel and the noisy daily correlations in the center panel is striking and indirectly suggests that the high-frequency data could be used in more effectively assessing the volatility asymmetries over the longer-run interdaily return horizons. We next discuss a simple-to-implement temporal aggregation formula explicitly designed to accomplish this goal.

1.3 Temporal Aggregation and Implied Daily Patterns

To facilitate the discussion of the temporal aggregation of the intraday returns, let

$$r_{t,i} \equiv R_{t+(i-1)\triangle,t+i\triangle} \equiv p_{t+i\triangle} - p_{t+(i-1)\triangle} \qquad t = 1,2,\ldots,T, \qquad i = 1,2,\ldots,n,$$

denote the return over the ith high-frequency interval on day t, where by assumption the number of nonoverlapping intraday return intervals is equal to the integer $n = \triangle^{-1}$, and by definition $r_{t,i+n} \equiv r_{t+1,i}$. Also, for notational convenience, denote the daily returns by the single subscript

$$r_t \equiv R_{t,t+1} \equiv \sum_{i=1}^{n} r_{t,i} \qquad t = 1,2,\ldots,T.$$

The lemma below then establishes a simple way of calculating the daily cross-correlations from the corresponding high-frequency cross-correlations. This lemma relies on the following two assumptions, both of which may be shown to hold well empirically.[12]

Assumption 1 *The high-frequency returns are third-order stationary in the sense that*

$$\mathrm{cov}(r_{t,i+h}^2, r_{t-\tau,h+1}) = \mathrm{cov}(r_{t,i}^2, r_{t-\tau,1}).$$

[12] We refer to Litvinova (2004) for further discussion and empirical verification of the two assumptions.

Assumption 2 *The expected value of the product of three high-frequency returns equals zero, unless two of the returns span the same time interval:*

$$E(r_{t,i}r_{\tau,j}r_{s,h})\begin{cases} \neq 0 & \text{if } t = \tau,\ i = j \text{ or } t = s,\ i = h \\ & \text{or } s = \tau,\ h = j \\ = 0 & \text{otherwise .} \end{cases}$$

Lemma *Given Assumptions 1 and 2, the cross-covariances for the daily returns may be expressed as:*

$$\tau \neq 0: \quad \text{cov}\,(r_t^2,\, r_{t-\tau}) = \sum_{i=-n}^{n} (n - |i|)\ \text{cov}(r_{t,\,i+1}^2, r_{t-\tau,1}), \tag{1}$$

$$\tau = 0: \quad \text{cov}\,(r_t^2, r_t) = 3\sum_{i=-n}^{n} (n - |i|)\ \text{cov}\,(r_{t,i+1}^2, r_{t,1}) - 2n\ \text{cov}(r_{t,1}^2, r_{t,1}). \tag{2}$$

The proof of the lemma is straightforward, and given in Appendix A.1.[13]

Turning to the corresponding empirical results, the long dashed line in the bottom panel in Figure 2 shows the implied daily cross-correlations calculated from the raw high-frequency five-minute returns. The daily correlations now indicate a highly significant contemporaneous leverage effect, accompanied by a fairly rapid and then gradual decay to zero for $j > 0$. Moreover, consistent with the picture for the high-frequency correlations in the top panel in the same figure, all of the correlations for $j < 0$ are now very close to zero. Therefore, the figure clearly highlights the benefit of using the high-frequency data in more accurately measuring and assessing the sources of the volatility asymmetry over daily and longer return horizons.

Meanwhile, it is well known that high-frequency returns are characterized by strong systematic intraday volatility patterns and that these patterns can sometimes obscure the dynamic dependencies gleaned from the autocorrelations for the raw absolute or squared intraday returns [see, e.g., the discussion in Andersen and Bollerslev (1997, 1998b)]. Hence, to guard against any spurious influences arising from the intraday volatility pattern, we also report the implied daily cross-correlations calculated from the five-minute returns in which we first remove the average intraday pattern,

[13] A bivariate extension of the temporal aggregation formula in which the cross-market correlations between the daily squared returns in one market and the daily lagged returns in another market are inferred from the underlying high-frequency returns may be derived in a similar manner; see Litvinova (2004) for further details.

$$\hat{r}_{t,i} = r_{t,i}\left(\sum_{\tau=1}^{T} r_{\tau,i}^2\right)^{-1/2} \qquad t = 1, 2, \ldots, T, \quad i = 1, 2, \ldots, n.$$

The implied daily correlations based on these standardized returns, indicated by the chain line in the bottom panel in the figure, fall almost on top of the ones based on the raw high-frequency returns. In sharp contrast to the daily correlations in the middle panel, the peak in the cross-correlations occurs very clearly at $j = 0$. Thus, the results for the standardized returns only further reinforce our earlier conclusion of a strong daily contemporaneous leverage effect.

We next evaluate the ability of some of the popular continuous-time stochastic volatility models analyzed in the existing literature for satisfactorily describing this striking empirical regularity.

2 ONE-FACTOR HESTON MODEL

The one-factor affine stochastic volatility model, first studied in the context of option pricing by Scott (1987) and later analyzed more formally by Heston (1993), is arguably the most widely used continuous-time stochastic volatility model in finance. The model is most easily expressed in stochastic differential equation form as

$$\begin{aligned} dp_t &= (\mu + cV_t)dt + \sqrt{V_t}\, dW_{1t} \\ dV_t &= (\alpha + \beta V_t)dt + \sigma\sqrt{V_t}\, dW_{2t} \end{aligned} \tag{3}$$

where W_{1t} and W_{2t} denote two possibly correlated Brownian motions, say $\text{corr}(dW_{1t}, dW_{2t}) = \rho$. This correlation between the two Brownian motions directly allows for a contemporaneous leverage effect, and ρ is often simply referred to as the "leverage parameter;" see also the related discussion for discrete-time stochastic volatility models in, for example, Harvey and Shephard (1996) and Yu (2005a). The parameter $-\beta$ dictates the speed of mean reversion in the (latent) volatility factor, whereas $-\alpha/\beta$ determines the unconditional long-run average volatility, and the volatility-of-volatility parameter σ is most directly related to the tails of the return distribution.[14]

To relate the implications from the Heston model to the sample cross-correlations discussed in the previous section, consider the continuously compounded returns from time t to $t + \triangle$ implied by the model

$$R_{t,t+\triangle} = p_{t+\triangle} - p_t = \mu\triangle + c\int_t^{t+\triangle} V_u du + \int_t^{t+\triangle} \sqrt{V_u} dW_{1u}.$$

[14] For the process to be stationary, these three parameters must satisfy: $\alpha > 0$, $\beta < 0$, and $\sigma^2 \leq 2\alpha$.

Then, as formally shown in Appendix A.2 for $k = 0,1,2,\ldots$,

$$\text{cov}(R^2_{t+(k-1)\triangle,t+k\triangle},R_{t-\triangle,t}) = -\rho\sigma\frac{\alpha}{\beta}a^2_{\triangle}(1+\beta a_{\triangle})^{k-1}, \tag{4}$$

where $a_{\triangle} = (1-e^{\beta\triangle})/(-\beta)$.[15] Moreover, it is easy to see that regardless of the value of $\triangle$, the covariances between the squared return and future returns implied by the Heston model are all equal to zero; that is, $\text{cov}(R^2_{t+(k-1)\triangle,t+k\triangle},R_{t-\triangle,t}) = 0$ for $k = -1,-2,\ldots$.

Thus, in terms of the daily cross-correlations between the squared return and the lagged returns depicted in the two bottom panels in Figure 2, this implies that $\text{corr}(R^2_{t,t+1},R_{t-j,t-j+1})$ for $j = 0,1,2,\ldots$ should decrease geometrically with j, with the rate of decay determined solely by the value of the volatility mean reversion parameter, β. Of course, the actual magnitude of the effect further depends on the values of α, σ, and ρ.

Similar closed-form expressions are unfortunately not available for the more elaborate continuous-time models analyzed below, nor for the corresponding cross-correlations for the absolute high-frequency returns, $\text{corr}(|R_{t,t+\triangle}|, R_{t-j\triangle,t-(j-1)\triangle})$, also depicted in the figures. Hence, in the following we instead rely on long, artificially simulated samples from the different model specifications to infer numerically the model-implied cross-correlations.

2.1 Daily Estimates and Implied Volatility Asymmetries

The transition density for discretely sampled observations from continuous-time stochastic volatility models and consequently their likelihood functions are generally not available in closed form. To circumvent this, we rely on the EMM estimation techniques of Gallant and Tauchen (1996), which under appropriate conditions approximate the (infeasible) maximum likelihood estimates arbitrarily well. Further details concerning the EMM estimation approach and the actual parameter estimates for the Heston model are given in Appendix A.3.

The parameter estimates for the daily S&P 500 returns analyzed here are directly in line with the values reported in the existing literature for other markets and time periods. In particular, the very small negative value for β implies a strong degree of volatility persistence, whereas the value for ρ suggests a highly significant contemporaneous leverage effect. Focusing instead on the cross-correlations implied by the estimates, the dashed line in the upper panel in Figure 3 shows the daily correlations between the squared returns and the returns at leads and lags ranging from -20 to 20 based on a simulated sample of 100,000 "days" from the estimated

[15] For small values of $\triangle$, $a_{\triangle} = (1-e^{\beta\triangle})/(-\beta) \approx (1-1-\beta\triangle)/(-\beta) = \triangle$, so that the formula approximately equals $-\rho\sigma(\alpha/\beta)\triangle^2(1+\beta\triangle)^{k-1}$.

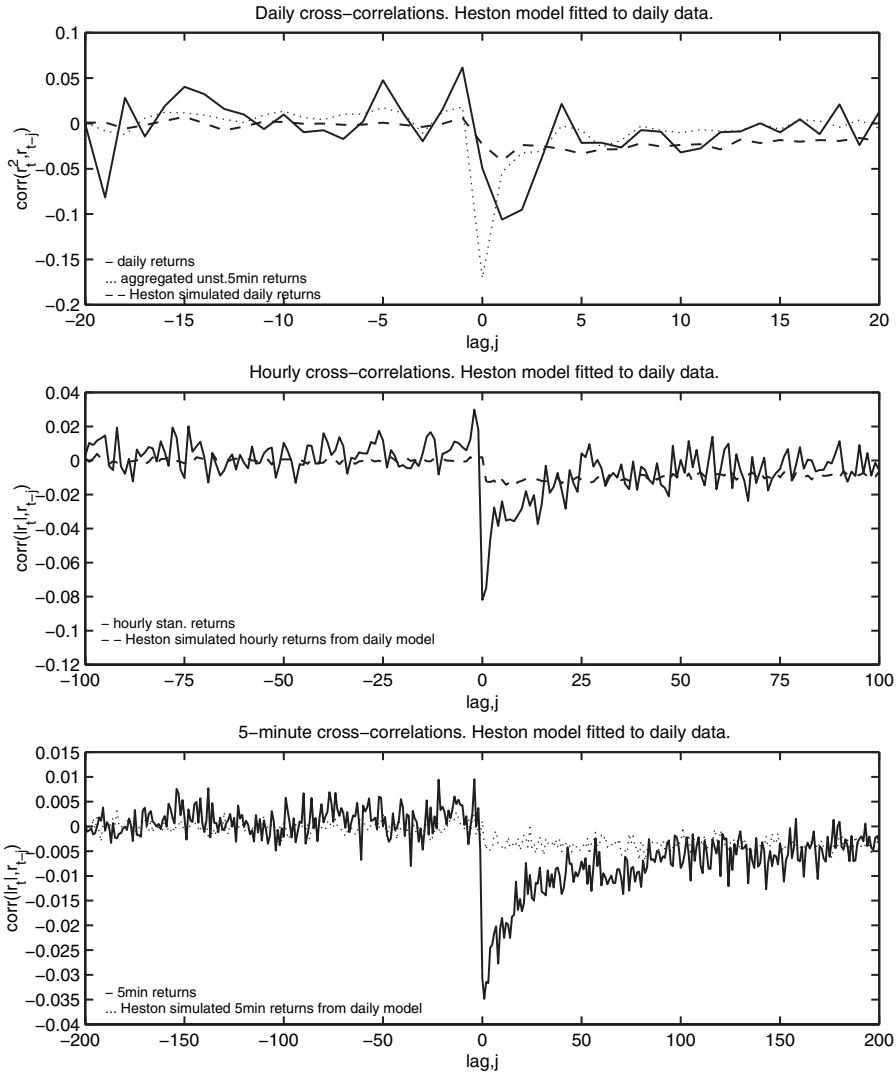

Figure 3 Daily Heston model cross-correlations. The top panel shows the daily cross-correlations calculated from the raw daily returns (solid line) and the aggregation formula for the five-minute returns (dotted line), along with the corresponding correlations implied by a long simulation from the estimated Heston model (dashed line). The middle panel shows the cross-correlations between the hourly returns and the absolute hourly returns (solid line), and the corresponding correlations implied by the Heston model (dashed line). The bottom panel reports the five-minute return sample cross-correlations (solid line) and model-implied estimates (dotted line). All of the calculations are based on S&P 500 futures data from 1988 to 1999.

model.[16] The figure also shows the corresponding sample cross-correlations for the daily S&P 500 returns (solid line), as well as the implied daily correlations (dotted line) calculated from the high-frequency five-minute returns and the temporal aggregation formula in Section 1.3. As is immediately evident, the daily model-implied correlations are essentially flat and completely miss the strong contemporaneous, or lag $j = 0$, effect that is clearly visible with the help of the high-frequency data and the aggregation formula.

This marked difference between the model-implied correlations and the information in the high-frequency data is further underscored by the middle and bottom panels in the same figure. The middle panel, in particular, shows the cross-correlations between the hourly absolute and raw returns for leads and lags ranging from -100 to 100 hours, or approximately 14 trading days. Comparing the solid line for the sample cross-correlations for the actual hourly returns, in which we have standardized the returns by the average standard deviation for that particular hour of the day as discussed above, with the dashed line implied by the daily estimates for the Heston model, reveals a marked difference for small values of $j > 0$. The bottom panel in the figure shows an even greater difference between the sample (solid line) and model-implied (dotted line) cross-correlations for the absolute five-minute returns. It is obvious that the Heston model simply is not capable of reproducing the very pronounced leverage effect seen in the data.

The Heston model has, of course, previously been soundly rejected in the literature, and the overall goodness-of-fit test for the model reported in the table in Appendix A.3 $[\chi^2(8) = 29.95]$ not surprisingly also strongly rejects the model.[17] Nonetheless, as far as the leverage effect and the asymmetry in the return–volatility relationship are concerned, the failure of the model only becomes visible when utilizing the high-frequency data, either directly as in the bottom two panels in Figure 3, or indirectly through the more accurate daily sample correlations implied by the high-frequency data as depicted in the top panel.

We next turn to a discussion of one of the more complicated multifactor stochastic volatility models that have recently been proposed in the literature in response to the widely documented rejections of the Heston and other one-factor models.

---

[16] Although the cross-correlations between the returns and the squared returns in the Heston model are available in closed-form, that is not the case for any of the other cross-correlations analyzed below. Hence, for compatibility reasons we rely on numerical simulations throughout. To enhance the numerical stability, we simulate the square root of the volatility process, $Z_t = \sqrt{V_t}$, which by *Ito*'s lemma has a constant diffusion term

$$dZ_t = -\frac{\beta}{2Z_t}\left(-\frac{\alpha}{\beta}+\frac{\sigma^2}{2\beta}-Z_t^2\right)dt+\frac{1}{2}\sigma dW_t.$$

See Durham and Gallant (2002) for further discussion along these lines.

[17] EMM estimation of the Heston model with hourly returns and the same auxiliary model as described in Section 4.2, resulted in equally strong rejections $[\chi^2(19) = 83.40]$. Further details concerning these results are available in Litvinova (2004).

3 TWO-FACTOR STOCHASTIC VOLATILITY MODEL

The addition of a second (latent) volatility factor has the potential of breaking the tight link between tail thickness and volatility persistence inherent in the one-factor stochastic volatility models.[18] More important in the present context, as we document below, the addition of a second volatility factor also allows for more flexible asymmetric return–volatility dependencies, as manifest in empirically more realistic cross-correlations patterns between the squared or absolute returns and the lagged returns.[19]

The extensive empirical analysis by Chernov et al. (2003) suggests that a logarithmic two-factor stochastic volatility model (LL2V1 in their terminology) provides a good description of the dynamic dependencies in the time series of daily Dow Jones Industrial Average (DJIA) returns over the 1953–1999 time period. Thus, we follow their lead in estimating the identical model for the S&P 500 futures return analyzed here. Specifically, in a slight change of notation,

$$\begin{aligned} dp_t &= (\alpha_{10} + \alpha_{12}\mu_t)dt + \text{s-exp}(\beta_{10} + \beta_{13}V_{1t} + \beta_{14}V_{2t}) \\ &\quad (\psi_{11}dW_{1t} + \psi_{13}dW_{3t} + \psi_{14}dW_{4t}) \\ d\mu_t &= \alpha_{22}\mu_t dt + dW_{2t} \\ dV_{1t} &= \alpha_{33}V_{1t}dt + dW_{3t} \\ dV_{2t} &= \alpha_{44}V_{2t}dt + (1 + \beta_{44}V_{2t})dW_{4t} \end{aligned} \tag{5}$$

where $\psi_{11} \equiv \sqrt{1 - \psi_{13}^2 - \psi_{14}^2}$, and s-exp(·) denotes the exponential function spliced with appropriate growth conditions to ensure the existence of a unique stationary solution; see Appendix A in Chernov et al. (2003) for a more detailed discussion. The μ_t factor represents a stochastic drift, whereas the instantaneous volatility is modeled as an exponential function of a linear combination of the two volatility factors, V_{1t} and V_{2t}. The simple one-factor Heston model discussed in the previous section is thus obtained by equating β_{10} to zero, replacing s-exp(·) with the square-root function, and "switching off" the drift, μ_t, and the second volatility factor, V_{2t}. Meanwhile, our model estimates discussed below clearly indicate the importance of incorporating a second volatility factor to simultaneously allow for both strong volatility persistence and more rapid short-lived volatility movements.

The volatility asymmetry implied by the LL2V1 model in Equation (5) is closely related to the value of the three correlation parameters, ψ_{11}, ψ_{13}, and ψ_{14}.

[18] The incorporation of discontinuous discrete jumps, as in, for example, Andersen, Benzoni, and Lund (2002) and Eraker, Johannes, and Polson (2003), affords an alternative approach for breaking this link.

[19] The analysis in Gallant, Hsu, and Tauchen (1999) also suggests that the two-factor stochastic volatility model can closely approximate long-memory-type dependencies in the volatility for fairly long lags.

In particular, it is possible to show that the instantaneous (conditional) correlation between the logarithmic price changes and the changes in the spot variance equals

$$\operatorname{corr}(dp_t, \beta_{13} dV_{1t} + \beta_{14} dV_{2t}) = \frac{\beta_{13}\psi_{13} + \beta_{14}\psi_{14}(1 + \beta_{44} V_{2t})}{\sqrt{\beta_{13}^2 + \beta_{14}^2 (1 + \beta_{44} V_{2t})^2}} dt. \tag{6}$$

Thus, in contrast to the time-invariant ρ coefficient in the Heston model, the contemporaneous leverage effect in the two-factor model is state-dependent. Of course, it is difficult to calibrate this instantaneous correlation directly to the data. Hence, we next provide a characterization of the two-factor model's implications in regard to the cross-correlations discussed in the preceding two sections. We begin by analyzing the implications of the model estimates obtained with daily data.

3.1 Daily Estimates and Implied Volatility Asymmetries

The EMM estimation results for the two-factor model based on the daily S&P 500 returns detailed in Appendix A.3 confirm that the model does a much better job of accounting for the dynamic dependencies in the daily data. None of the *t*-ratio diagnostics for the model are significant, and the overall chi-square goodness-of-fit test $[\chi^2(2) = 2.000]$ is also supportive of the LL2V1 specification. Therefore, our results are directly in line with the previous evidence for the daily DJIA data reported in Chernov et al. (2003). At the same time, it is noteworthy that on evaluating the expression in Equation (6) at the estimated parameter values and the stationary mean of the V_{2t} process, the instantaneous leverage effect implied by the model equals -0.372, which is extremely close to the previously estimated value of -0.374 for ρ in the one-factor Heston model.

Meanwhile, the cross-correlations implied by the two-factor model depicted in Figure 4 obviously look very different from the corresponding correlations for the Heston model in Figure 3. Most noticeable, the patterns in the daily correlations from the LL2V1 model in the top panel now closely match the sharp negative peak at lag 0 and the gradual decay to zero observed in the implied daily sample correlations calculated from the high-frequency returns. The LL2V1 model also provides a closer fit to the high-frequency sample absolute return cross-correlations in the two bottom panels. However, even though the model is able to match the slow multiday decay in the five-minute sample correlations, it still misses the pronounced spike at zero and the accompanying fairly rapid decay for the first two hours thereafter (lag 0 to 24). Hence, in our last set of empirical results we investigate how well the two-factor model estimated directly to the high-frequency data does in terms of describing these intraday dynamic features.

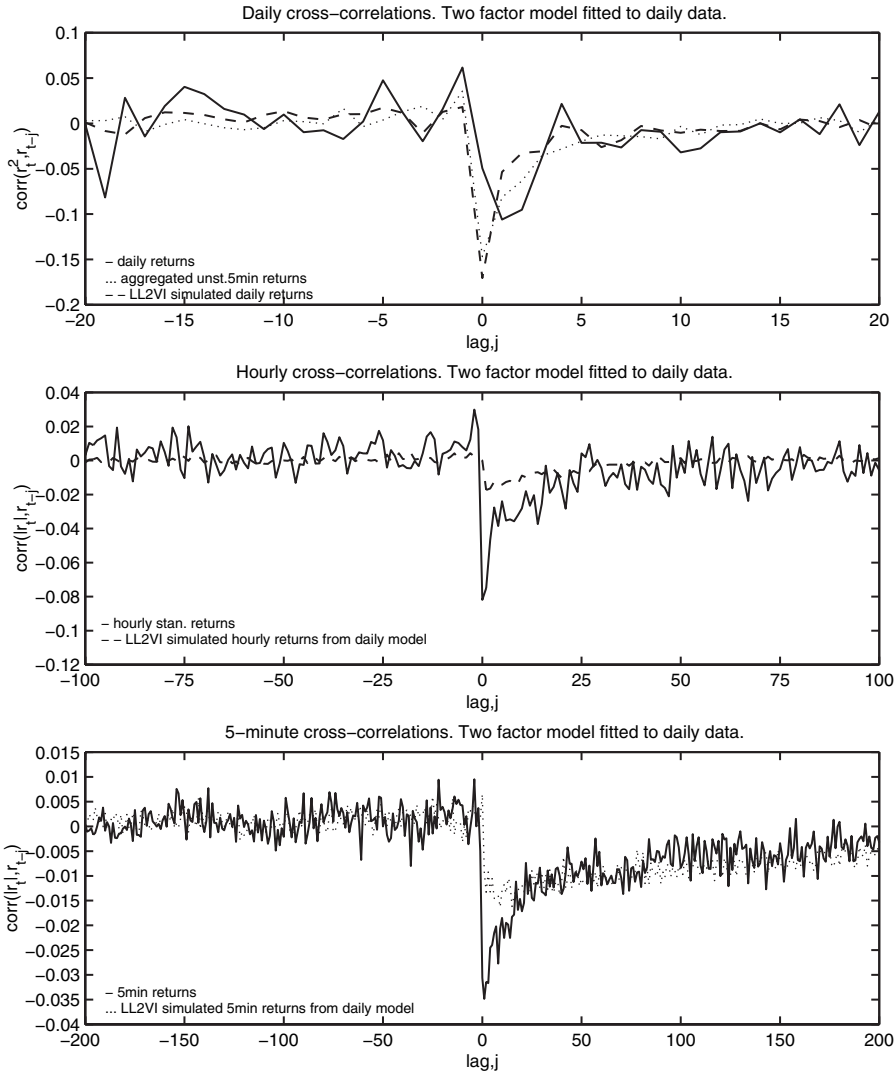

Figure 4 Daily two-factor model cross-correlations. The top panel shows the daily cross-correlations calculated from the raw daily returns (solid line) and the aggregation formula for the five-minute returns (dotted line), along with the corresponding correlations implied by a long simulation from the estimated daily two-factor model (dashed line). The middle panel shows the cross-correlations between the hourly returns and the absolute hourly returns (solid line), and the corresponding correlations implied by the two-factor model (dashed line). The bottom panel reports the five-minute return sample cross-correlations (solid line) and model-implied estimates (dotted line). All of the calculations are based on S&P 500 futures data from 1988 to 1999.

3.2 Hourly Estimates and Implied Volatility Asymmetries

The estimation of a satisfactory model for high-frequency five-minute returns over a long 12-year sample involving more than 200,000 return observations, let alone estimation with the underlying nonsynchronous tick-by-tick transaction prices, presents some formidable challenges; see, for example, the discussion in Andersen and Bollerslev (1998b) and Rydberg and Shephard (2003) and the many references therein. As a compromise, we therefore restrict our attention to the smaller sample of hourly returns consisting of *only* 19,299 observations. Also, as previously noted, high-frequency returns are subject to very strong intraday volatility patterns. Rather than augmenting the two-factor LL2V1 model to accommodate this important feature of the data, we instead base our hourly model estimates on the standardized hourly returns in which we divide each of the returns with the sample standard deviation for that particular hour of the day.

The details concerning the resulting EMM estimates are again reported in Appendix A.3. Consistent with the component structure advocated by Engle and Lee (1999), the auxiliary model for the hourly returns used in the estimation is comprised of a GARCH(2,2) leading term augmented with an eight-order Hermite expansion. The actual point estimates for the half-lives for the two volatility factors ($-\log 2/\alpha_{33}$ and $-\log 2/\alpha_{44}$) equal 13.5 and 0.1 days, respectively. These are both considerably less persistent than the corresponding daily model estimates. Moreover, in contrast to the daily estimates discussed above, and the previous daily results reported in Chernov et al. (2003), where the instantaneous correlations between the innovation to the price process and each of the two volatility factors are both around -0.35, the first slowly mean-reverting volatility factor now has a large negative correlation of -0.79, whereas the second quickly mean-reverting volatility factor has a small *positive* correlation of 0.12.[20] These results thus suggest that a positive return shock is on average associated with an immediate, albeit very small and short-lived, increase in volatility coming from the transient component, and a much larger and immediate decrease in volatility associated with the more persistent component. In other words, the model exclusively attributes the leverage effect to the long-term volatility factor.

The dynamic implications of the hourly model estimates in terms of the intraday absolute return cross-correlations patterns are depicted in the bottom two panels in Figure 5. Although the two-factor model estimated with daily returns does not account for the sharp peak at lag 0, the hourly model generally provides a close match to the sample cross-correlations. More important, the daily cross-correlations from the simulated hourly model also closely match the implied daily sample correlations in the top panel. Hence, from the perspective of satisfactorily describing the volatility asymmetries inherent in the high-frequency

[20] Durham (2004) also found the correlation between returns and the persistent volatility component in an alternative two-factor representation for daily S&P 500 returns to be -0.58, whereas the correlation between returns and the transient volatility factor in his model is estimated to be 0.08.

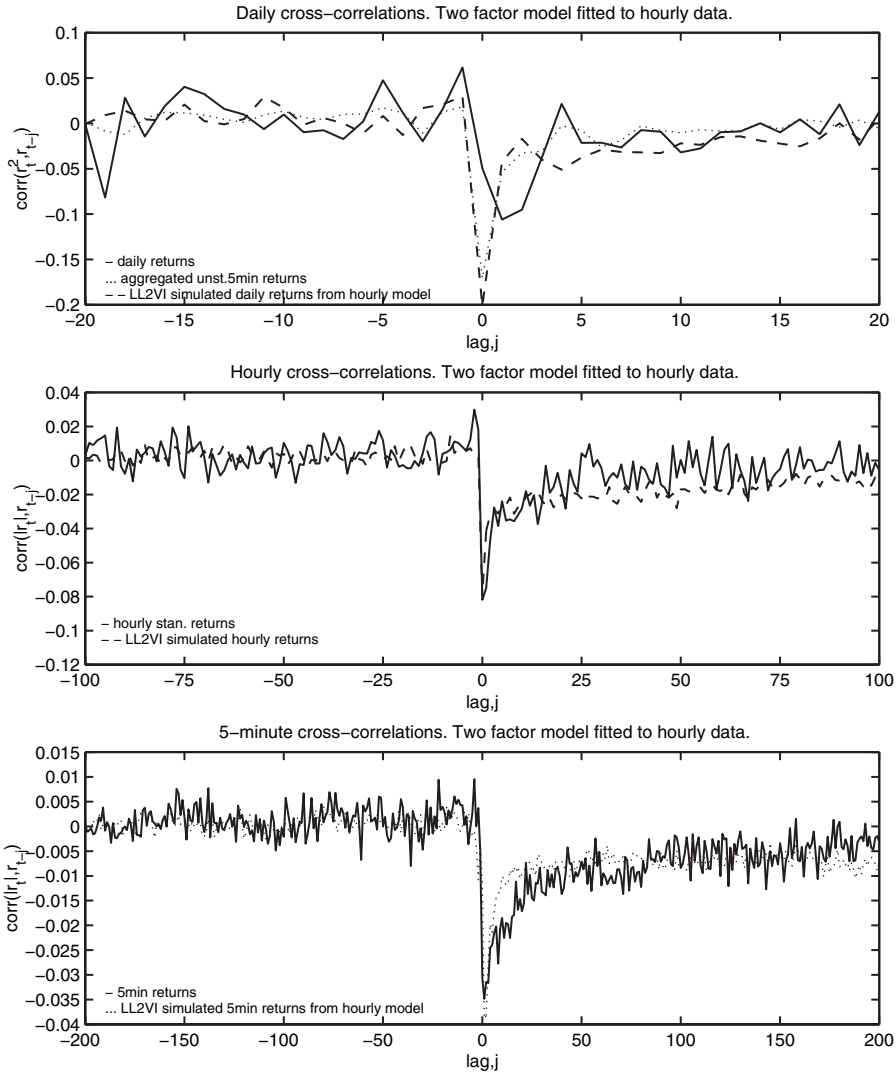

Figure 5 Hourly two-factor model cross-correlations. The top panel shows the daily cross-correlations calculated from the raw daily returns (solid line) and the aggregation formula for the five-minute returns (dotted line), along with the corresponding correlations implied by a long simulation from the estimated hourly two-factor model (dashed line). The middle panel shows the cross-correlations between the hourly returns and the absolute hourly returns (solid line), and the corresponding correlations implied by the two-factor model (dashed line). The bottom panel reports the five-minute return sample cross-correlations (solid line) and model-implied estimates (dotted line). All of the calculations are based on S&P 500 futures data from 1988 to 1999.

data, the LL2V1 model calibrated to best fit the hourly return dynamics may be seen as a success.

Meanwhile, many of the *t*-ratio diagnostics for the hourly model estimates reported in Table A3 are significant, and the overall fit of the model is also strongly rejected by the EMM chi-square goodness-of-fit test $[\chi^2(13) = 63.53]$. This is, of course, not entirely surprising and indirectly suggests the need of incorporating additional volatility factors and/or jump processes to account fully for the complex dynamic dependencies in the high-frequency data. Alternative long-memory continuous-time stochastic volatility models have also been proposed by Brockwell and Marquardt (2005), Comte and Renault (1998), and Comte, Coutin, and Renault (2004), among others.[21] In spite of the very different long-run implications from these fractionally integrated models, as previously noted, the pertinent hyperbolic decay patterns may typically be approximated fairly closely for moderately large lags by the superposition of just a few volatility factors. The estimation of such models is beyond the scope of this article, but it is the focus of much current research.[22] The volatility asymmetries documented here should thus provide a new important yardstick by which to measure the empirical success of any of these more complicated models.

4 CONCLUSION

We find striking new evidence on the negative correlation between stock market movements and stock market volatility over fine intraday sampling frequencies. In particular, it appears that a sharp decline in the market over a five-minute interval is typically associated with a rise in market volatility that persists for up to several days after the initial event. These findings are generally consistent with those of previous studies using daily data; nonetheless, we also develop new aggregation formulas showing how the high-frequency data may be used in much more precisely estimating the daily cross-correlation patterns in comparison with previously available estimates obtained from daily data alone.

We also present estimation results for specific continuous-time stochastic volatility models, where the observed market data are viewed as observations on a discretely sampled diffusion. We work at two different sampling frequencies. The first is the usual daily frequency, where we obtain results about the apparent empirical success of a two-factor volatility model with leverage effects consistent with those documented in the existing literature. Because the estimated model operates in continuous time, we can readily "intrapolate" to a finer time scale than the daily observation frequency actually used in estimating the

[21] Related discrete-time long-memory stochastic volatility models have previously been developed and estimated by Breidt, Crato, and de Lima (1998) and Harvey (1998).

[22] Interesting ongoing work along these lines includes the ADS-ACD model in Ryan (2006), the Lévy-driven stochastic volatility models in Todorov (2005), and the two-factor stochastic volatility model explicitly incorporating market microstructure noise estimated by Tran (2006).

model and determine the model's ability to predict the observed patterns for finer intradaily data. Doing so, we find that the two-factor model estimated with daily data is unable to account for the actual within-day patterns of covariation between returns and volatility. Re-estimating the model on hourly returns, we find that it does much better in terms of these same diagnostics, although it is formally rejected by a general chi-squared model misspecification test, in turn suggesting that further elaboration of the model would be needed in future studies.

In this regard, our work clearly documents the power of using the high-frequency data for measuring and assessing the importance of the volatility asymmetry effect. At the same time, our work raises a serious conundrum for economic modeling of stock market data. The conventional explanation for the volatility asymmetry effect is risk-based, as described in, for example, Campbell and Hentschel (1992) and reviewed in Tauchen (2005). However, the existing economic models are thought of as applying at much coarser time intervals, say monthly or quarterly; the mechanisms of the risk-based explanation might be expected to work themselves out only relatively slowly. Thus, it is not at all clear that a risk-based explanation can adequately account for the strong asymmetry we detect at the intraday level. Evidently, the risk-based models need to be reformulated, possibly in a more flexible continuous-time framework, and confronted directly to the empirical findings of this article.

APPENDIX A

A.1 Proof of Temporal Aggregation Formula

Consider first the case where $\tau \neq 0$. By standard arguments

$$
\begin{aligned}
\operatorname{cov}(r_t^2, r_{t-\tau}) &= \operatorname{cov}\left(\left(\sum_{i=1}^{n} r_{t,i}\right)^2, \sum_{j=1}^{n} r_{t-\tau,j}\right) \\
&= \operatorname{cov}\left(\sum_{i=1}^{n} r_{t,i}^2, \sum_{j=1}^{n} r_{t-\tau,j}\right) + 2\operatorname{cov}\left(\sum_{i<k} r_{t,i} r_{t,k}, \sum_{j=1}^{n} r_{t-\tau,j}\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{cov}\left(r_{t,i}^2, r_{t-\tau,j}\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{cov}\left(r_{t,i-j+1}^2, r_{t-\tau,1}\right) \\
&= \sum_{i=1}^{n}\sum_{k=i-n}^{i-1} \operatorname{cov}\left(r_{t,k+1}^2, r_{t-\tau,1}\right) \\
&= \sum_{i=-n}^{n} (n - |i|)\operatorname{cov}(r_{t,i+1}^2, r_{t-\tau,1}),
\end{aligned}
$$

where the third and fourth equality follows by Assumptions 2 and 1, respectively, and the second to last equality is obtained by substituting $k = i - j$.

Now consider the case where $\tau = 0$. As before

$$\begin{aligned}\operatorname{cov}(r_t^2, r_t) &= \operatorname{cov}\left(\left(\sum_{i=1}^{n} r_{t,i}\right)^2, \sum_{j=1}^{n} r_{t,j}\right) \\ &= \operatorname{cov}\left(\sum_{i=1}^{n} r_{t,i}^2, \sum_{j=1}^{n} r_{t,j}\right) + 2\operatorname{cov}\left(\sum_{i<k} r_{t,i} r_{t,k}, \sum_{j=1}^{n} r_{t,j}\right).\end{aligned}$$

The last term equals zero by Assumption 2 unless $j = i$ or $j = k$. Further rearranging this term:

$$\begin{aligned}&2\sum_{j=1}^{n}\sum_{i=1}^{n-1}\sum_{k=i+1}^{n} \operatorname{cov}\left(r_{t,i} r_{t,k}, r_{t,j}\right) \\ &\quad = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \left[\operatorname{cov}\left(r_{t,i} r_{t,j}, r_{t,i}\right) + \operatorname{cov}\left(r_{t,i} r_{t,j}, r_{t,j}\right)\right] \\ &\quad = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \operatorname{cov}(r_{t,1} r_{t,\underbrace{(j-i}_{k}+1)}, r_{t,1}) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \operatorname{cov}(r_{t,\underbrace{(i-j}_{-k}+1)} r_{t,1}, r_{t,1}) \\ &\quad = 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} [\operatorname{cov}(r_{t,1} r_{t,k+1}, r_{t,1}) + \operatorname{cov}(r_{t,-k+1} r_{t,1}, r_{t,1})] \\ &\quad = 2\sum_{j=1}^{n} (n-j)\left[\operatorname{cov}\left(r_{t,1} r_{t,j+1}, r_{t,1}\right) + \operatorname{cov}\left(r_{t,-j+1} r_{t,1}, r_{t,1}\right)\right] \\ &\quad = 2\sum_{j=-n}^{n} (n - |j|)\operatorname{cov}(r_{t,1} r_{t,j+1}, r_{t,1}) - 2n \operatorname{cov}(r_{t,1}^2, r_{t,1}),\end{aligned}$$

where the second equality follows by Assumption 1. Thus,

$$\begin{aligned}&\operatorname{cov}\left(\left(\sum_{i=1}^{n} r_{t,i}\right)^2, \sum_{j=1}^{n} r_{t,j}\right) \\ &\quad = \sum_{i=-n}^{n} (n - |i|)\operatorname{cov}(r_{t,i+1}^2, r_{t,1}) + 2\sum_{i=-n}^{n} (n - |i|)\operatorname{cov}(r_{t,1} r_{t,i+1}, r_{t,1}) - 2n \operatorname{cov}(r_{t,1}^2, r_{t,1}),\end{aligned}$$

and because

$$\begin{aligned}\text{cov}(r_{t,1}r_{t,i+1},r_{t,1}) &= E[r_{t,1}r_{t,i+1}r_{t,1}] - E[r_{t,1}r_{t,i+1}]E[r_{t,1}] \\ &= E[r^2_{t,-i+1}r_{t,1}] - E[r^2_{t,-i+1}]E[r_{t,1}] \\ &= \text{cov}(r^2_{t,-i+1},r_{t,1}),\end{aligned}$$

the expression in the lemma follows by substitution.

A.2 Cross-Correlations in the Heston Model

For easy reference to earlier results in Bollerslev and Zhou (2002, 2006), we follow the slightly different notation therein,

$$\begin{aligned}dp_t &= (\mu + cV_t)dt + \sqrt{V_t}\,dB_t \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}\,dW_t,\end{aligned} \tag{A1}$$

where $\text{corr}(dB_t,dW_t) = \rho$. The continuously compounded returns from time t to $t+\triangle$ is then given by

$$R_{t,t+\triangle} = p_{t+\triangle} - p_t = \mu\triangle + c\int_t^{t+\triangle} V_u du + \int_t^{t+\triangle} \sqrt{V_u}dB_u.$$

For simplicity, assume that $\mu = c = 0$. It follows then by *Ito*'s lemma that

$$R^2_{t,t+\triangle} = 2\int_t^{t+\triangle} R_{u,u+\triangle}\sqrt{V_u}dB_u + \int_t^{t+\triangle} V_u du.$$

Moreover, because $E\left(\int_{t+(n-1)\triangle}^{t+n\triangle} R_{u,u+\triangle}\sqrt{V_u}dB_u|\mathcal{F}_t\right) = 0,$

$$\begin{aligned}&E\left(\int_{t+(n-1)\triangle}^{t+n\triangle} R_{u,u+\triangle}\sqrt{V_u}dB_u \int_{t-\triangle}^{t} \sqrt{V_u}dB_u\right) \\ &\quad = E\left[E\left(\int_{t+(n-1)\triangle}^{t+n\triangle} R_{u,u+\triangle}\sqrt{V_u}dB_u|\mathcal{F}_t\right)\int_{t-\triangle}^{t} \sqrt{V_u}dB_u\right] = 0.\end{aligned}$$

Hence, the unconditional cross-correlations of interest may be expressed as

$$
\begin{aligned}
&\operatorname{cov}(R^2_{t+(n-1)\triangle,t+n\triangle}, R_{t-\triangle,t}) \\
&\quad = 2\operatorname{cov}\left(\int_{t+(n-1)\triangle}^{t+n\triangle} R_{u,u+\triangle}\sqrt{V_u}dB_u, \int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right) \\
&\qquad + \operatorname{cov}\left(\int_{t+(n-1)\triangle}^{t+n\triangle} V_u du, \int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right) \\
&\quad = E\left(\int_{t+(n-1)\triangle}^{t+n\triangle} V_u du \int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right) \\
&\quad = E\left[E_{t+(n-1)\triangle}\left(\int_{t+(n-1)\triangle}^{t+n\triangle} V_u du\right)\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right] \\
&\quad = E\left(a_\triangle V_{t+(n-1)\triangle}\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right),
\end{aligned}
\tag{A2}
$$

where the last equality follows from the result in Bollerslev and Zhou (2002),

$$
\begin{aligned}
E_t\left(\int_t^T V_s ds\right) &= V_t\frac{1}{\kappa}(1-e^{-\kappa(T-t)}) + \theta(T-t) - \frac{\theta}{\kappa}(1-e^{-\kappa(T-t)}) \\
&= a_{T-t}V_t + b_{T-t}.
\end{aligned}
$$

Now, using Equation (A2) in the Appendix of Bollerslev and Zhou (2006), it follows that for $n = 1$,

$$
E\left(V_t\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right) = \rho\sigma\theta a_\triangle. \tag{A3}
$$

Consider now the expression for $n = 2$,

$$
\begin{aligned}
&E\left(V_{t+\triangle}\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)\\
&\quad=E\left[\left(V_t+\int_{t}^{t+\triangle}\kappa(\theta-V_u)du+\int_{t}^{t+\triangle}\sigma\sqrt{V_u}dW_u\right)\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right]\\
&\quad=E\left(V_t\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)+E\left(\int_{t}^{t+\triangle}\kappa(\theta-V_u)du\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)\\
&\qquad+\sigma E\left(\int_{t}^{t+\triangle}\sqrt{V_u}dW_u\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right).
\end{aligned}
$$

The last term obviously equals zero. The second term may alternatively be expressed as

$$
\begin{aligned}
&E\left(\int_{t}^{t+\triangle}\kappa(\theta-V_u)du\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)\\
&\quad=E\left(\int_{t}^{t+\triangle}-\kappa V_u du\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)\\
&\quad=-\kappa E\left[E_t\left(\int_{t}^{t+\triangle}V_u du\right)\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right]\\
&\quad=-\kappa a_{\triangle}E\left(V_t\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right).
\end{aligned}
$$

Collecting terms, we thus have

$$
E\left(V_{t+\triangle}\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)=(1-\kappa a_{\triangle})E\left(V_t\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right).
$$

By analogous arguments it follows that in general

$$
E\left(V_{t+(n-1)\triangle}\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right)=(1-\kappa a_{\triangle})^{n-1}E\left(V_t\int_{t-\triangle}^{t}\sqrt{V_u}dB_u\right). \tag{A4}
$$

Substituting (A4) into the general expression for the cross-correlations in (A2) and utilizing (A3) thus yields

$$\text{cov}(R^2_{t+(n-1)\triangle,t+n\triangle},R_{t-\triangle,t}) = (1-\kappa a_\triangle)^{n-1}\rho\sigma\theta a^2_\triangle. \tag{A5}$$

Finally, note that for small values of $\triangle$, $a_\triangle = (1-e^{-\kappa\triangle})/\kappa \approx (1-1+\kappa\triangle)/\kappa = \triangle$, so that the latter expression may be approximated by

$$\text{cov}(R^2_{t+(n-1)\triangle,t+n\triangle},R_{t-\triangle,t}) \approx (1-\kappa\triangle)^{n-1}\rho\sigma\theta\triangle^2. \tag{A6}$$

A.3 EMM Estimation Results

Several alternative estimation procedures have been developed in the literature for the estimation of continuous-time latent stochastic volatility processes over the past decade. The EMM estimation techniques of Gallant and Tauchen (1996) allow for the calculation of estimates that are arbitrarily close to maximum likelihood. Intuitively, the technique works by using the score from an easy-to-estimate sufficiently general auxiliary model to define the moment conditions to be equated to zero in a long simulation from the structural model to be estimated by appropriately choosing the underlying model parameters; for further discussion and documentation of the free software used in our estimation, see Gallant and Tauchen (1993, 1998).

The auxiliary model for the daily S&P 500 returns used here consists of a GARCH(1,1) leading term augmented with a nonparametric error density in the form of an eight-degree Hermit expansion. This particular model was chosen by the Bayesian information criterion (BIC) in a semi-nonparametric (SNP) series expansion, and it closely mirrors the auxiliary models used in many other applications. The resulting parameter estimates for the Heston model are reported in the top part of Table A1. The estimates are based on a simulated sample of length 100,000 "days," with an Euler discretization grid of $1/\triangle = 24$ steps per day. The *t*-ratio diagnostics associated with the different parameters in the auxiliary model are in turn reported in the bottom part of the table. All of the *t*-statistics associated with the even powers in the Hermite expansion are significant, indicating that the Heston model is not able to describe satisfactorily the conditional return distribution. The significant *t*-statistic associated with the first-order autoregressive coefficient also indicates that the constant drift specification in the simple Heston model is too restrictive.

The daily estimates for the two-factor LL2V1 model reported in Table A2 are based on the same auxiliary GARCH(1,1) model augmented an eight-order Hermite expansion used in the estimation of the daily Heston model. In contrast to the results in Table A1, none of the *t*-ratio diagnostics for the LL2V1 model are significant. The overall chi-square goodness-of-fit test for the model reported in the top panel also indicates that the model does a good job in terms of describing

the distributional characteristics of the daily returns as summarized by the auxiliary model.

The last set of EMM estimation results in Table A3 pertains to the two-factor LL2V1 model estimated with hourly returns. Because the LL2V1 model is not designed to account for the strong intraday volatility patterns, we standardized the hourly returns used in the estimation of the model by dividing each of the returns with the unconditional sample standard deviation for that particular hour of the day, as described in Section 1.3 in this article. The BIC used in selecting the auxiliary model in the SNP series expansion indicates that a GARCH(2,2) model with a nonparametric error density represented by an eight-degree Hermit expansion is the preferred specification. The estimation of the model is based on a simulated sample of length 100,000 "days," with $1/\triangle = 77$ steps per "day." With our working definition of an "hour" as 11 five-minute intervals, this leaves us with seven "hours" per "day." Many of the t-ratio diagnostics reported in the bottom panel

Table A1 EMM estimation results and diagnostics for one-factor Heston model based on daily S&P 500 futures returns, 1988–1999.

	Parameter	Estimate	SE
	μ	0.0174	0.0131
	α	0.0152	0.0056
	β	−0.0259	0.0099
	σ	0.1160	0.0259
	ρ	−0.3738	0.1278
	χ^2_8	29.9475	
	p-value	0.0002	
	t-ratio diagnostics		
Location function			
	b_0	psi(1)	0.363
	b_1	psi(2)	−3.030
Scale function			
	τ_0	tau(1)	0.933
	τ_{1a}	tau(2)	0.570
	τ_{1g}	tau(3)	0.752
Hermite polynomial			
	$a_{0,1}$	A(2)	0.022
	$a_{0,2}$	A(3)	2.383
	$a_{0,3}$	A(4)	−0.504
	$a_{0,4}$	A(5)	3.398
	$a_{0,5}$	A(6)	−0.750
	$a_{0,6}$	A(7)	3.231
	$a_{0,7}$	A(8)	−0.799
	$a_{0,8}$	A(9)	2.913

Table A2 EMM estimation results and diagnostics for two-factor LL2VI model based on daily S&P 500 futures returns, 1988–1999.

	Parameter	Estimate	SE
	α_{10}	0.0001	0.0002
	α_{12}	−0.1823	0.0671
	α_{22}	−45.8384	7.7956
	α_{33}	0.0000	0.0000
	α_{44}	−0.2931	0.3279
	β_{10}	−6.2619	0.6007
	β_{13}	0.0084	0.0043
	β_{14}	0.4011	0.2351
	β_{44}	0.0793	0.1782
	ψ_{13}	−0.3407	0.2435
	ψ_{14}	−0.3652	0.2157
	χ^2_2	1.9998	
	p-value	0.3679	
	t-ratio diagnostics		
Location function			
	b_0	psi(1)	0.113
	b_1	psi(2)	−0.350
Scale function			
	τ_0	tau(1)	0.610
	τ_{1a}	tau(2)	0.608
	τ_{1g}	tau(3)	0.525
Hermite polynomial			
	$a_{0,1}$	A(2)	−0.098
	$a_{0,2}$	A(3)	0.034
	$a_{0,3}$	A(4)	0.195
	$a_{0,4}$	A(5)	−0.018
	$a_{0,5}$	A(6)	0.451
	$a_{0,6}$	A(7)	−0.042
	$a_{0,7}$	A(8)	0.545
	$a_{0,8}$	A(9)	0.085

Table A3 EMM estimation results and diagnostics for two-factor LL2VI model based on hourly S&P 500 futures returns, 1988–1999.

Parameter	Estimate	SE
α_{10}	0.0002	0.0001
α_{12}	0.0002	0.0003
α_{22}	−0.0246	0.1045
α_{33}	−0.0515	0.0142
α_{44}	−8.2755	2.1099

continued

Table A3 *(continued)*

	Parameter	Estimate	SE
	β_{10}	−5.4861	0.0372
	β_{13}	0.1166	0.0182
	β_{14}	−3.3383	0.3897
	β_{44}	2.6562	0.0988
	ψ_{13}	−0.7919	0.0569
	ψ_{14}	0.1159	0.0266
	χ^2_{13}	63.5282	
	p-value	<0.01	
	t-ratio diagnostics		
Location function			
	b_0	psi(1)	−0.760
	b_1	psi(2)	−1.163
Scale function			
	τ_0	tau(1)	1.766
	τ_{1a}	tau(3)	2.286
	τ_{2a}	tau(2)	0.329
	τ_{1g}	tau(4)	2.094
	τ_{2g}	tau(5)	0.830
Hermite polynomial			
	$a_{1,0}$	A(2)	2.075
	$a_{0,1}$	A(3)	0.047
	$a_{1,1}$	A(4)	2.344
	$a_{0,2}$	A(5)	1.106
	$a_{1,2}$	A(6)	0.364
	$a_{0,3}$	A(7)	−0.213
	$a_{1,3}$	A(8)	3.829
	$a_{0,4}$	A(9)	1.768
	$a_{1,4}$	A(10)	−0.031
	$a_{0,5}$	A(11)	−0.491
	$a_{1,5}$	A(12)	2.710
	$a_{0,6}$	A(13)	2.524
	$a_{1,6}$	A(14)	−0.486
	$a_{0,7}$	A(15)	−0.776
	$a_{1,7}$	A(16)	1.696
	$a_{0,8}$	A(17)	2.926

of Table A3 are significant, and the overall chi-square test in the top panel also strongly rejects the hourly LL2V1 model.

Received May 25, 2006; revised February 20, 2006; accepted March 9, 2006.

References

Andersen, T. G., L. Benzoni, and J. Lund. (2002). "An Empirical Investigation of Continuous-Time Equity Return Models." *Journal of Finance 57*, 1239–1284.

Andersen, T. G., and T. Bollerslev. (1997). "Intraday Periodicity and Volatility Persistence in Financial Markets." *Journal of Empirical Finance 4*, 115–158.

Andersen, T. G., and T. Bollerslev. (1998a). "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts." *International Economic Review 39*, 885–905.

Andersen, T. G., and T. Bollerslev. (1998b). "Deutschemark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies." *Journal of Finance 53*, 219–265.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens. (2001). "The Distribution of Stock Return Volatility." *Journal of Financial Economics 61*, 43–76.

Bali, T. G., and L. Peng. (2006). "Is there a Risk-Return Tradeoff? Evidence from High-Frequency Data." Forthcoming in *Journal of Applied Econometrics*.

Bekaert, G., and G. Wu. (2000). "Asymmetric Volatility and Risk in Equity Markets." *The Review of Financial Studies 13*, 1–42.

Black, B. (1976). "Studies of Stock Price Volatility Changes." *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics*, 177–181.

Bollerslev, T., and J. H. Wright. (2000). "Semiparametric Estimation of Long-Memory Volatility Dependencies: The Role of High-frequency Data." *Journal of Econometrics 98*, 81–106.

Bollerslev, T. and H. Zhou. (2002). "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility." *Journal of Econometrics 109*, 33–65.

Bollerslev, T. and H. Zhou. (2006). "Volatility Puzzles: A Unified Framework for Gauging Return-Volatility Regressions." *Journal of Econometrics 131,* 123–150.

Bouchaud, J.-P., A. Matacz, and M. Potters. (2001). "Leverage Effect in Financial Markets: The Retarded Volaitlity Model." *Physical Review Letters 87*, 228701–228704.

Breidt, F. J., N. Crate, and P. de Lima. (1998). "On the Detection and Estimation of Long Memory in Stochastic Volatility." *Journal of Econometrics 83*, 325–348.

Brockwell, P. J, and T. Marquardt. (2005). "Lévy-Driven and Fractionally Integrated ARMA Processes with Continuous Time Parameter." *Statistica Sinica 15*, 477–494.

Campbell, J. Y., and L. Hentschel. (1992). "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics 31*, 281–331.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen. (2003). "Alternative Models for Stock Price Dynamics." *Journal of Econometrics 116*, 225–257.

Christie, A. C. (1982). "The Stochastic Behavior of Common Stock Variances—Value, Leverage and Interest Rate Effects." *Journal of Financial Economics 3*, 145–166.

Comte, F. L. Coutin, and E. Renault. (2004). "Fractional Stochastic Volatility Models." Working Paper, Department of Economics, University of North Carolina.

Comte, F., and E. Renault. (1998). "Long-Memory in Continuous-Time Stochastic Volatility Models." *Mathematical Finance 8*, 291–323.

Dacorogna, M., R. Gençay, U. Müller, R. Olsen, and O. Pictet. (2001). *An Introduction to High-Frequency Finance*. San Diego: Academic Press.

Durham, G. B. (2004). "Monte Carlo Methods for Estimating, Smoothing, and Filtering One and Two-Factor Stochastic Volatility Models." Working Paper, Department of Finance, University of Colorado.

Durham, G. B., and A. R. Gallant. (2002). "Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes." *Journal of Business and Economic Statistics 20*, 297–316.

Engle, R. F., and G. J. Lee. (1999). "A Permanent and Transitory Component Model of Stock Return Volatility." In R. F. Engle and H. L. White (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger*. Oxford, UK: Oxford University Press.

Engle, R. F., and V. K. Ng. (1993). "Measuring and Testing the Impact of News on Volatility." *Journal of Finance 48*, 1749–1778.

Eraker, B., M. S. Johannes, and N. G. Polson. (2003). "The Impact of Jumps in Volatility and Returns." *Journal of Finance 58*, 1269–1300.

Figlewski, S., and X. Wang. (2001). "Is the 'Leverage Effect' a Leverage Effect?" Working Paper, Department of Finance, New York University.

French, K. R., G. W. Schwert, and R. F. Stambaugh. (1987). "Expected Stock Returns and Volatility." *Journal of Financial Economics 19*, 3–30.

Gallant, A. R., C.-T. Hsu, and G. Tauchen. (1999). "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance." *Review of Economics and Statistics 81*, 617–631.

Gallant, A. R., and G. Tauchen. (1993). "SNP: A Program for Nonparametric Time Series Analysis, Version 8.3, User's Guide." Program Documentation, Department of Economics, Duke University.

Gallant, A. R., and G. Tauchen. (1996). "Which Moments to Match?" *Econometric Theory 12*, 657–681.

Gallant, A. R., and G. Tauchen. (1998). "EMM: A Program for Efficient Method of Moments Estimation, Version 1.4, User's Guide." Program Documentation, Department of Economics, Duke University.

Ghysels, E., P. Santa-Clara, and R. Valkanov. (2005). "There is a Risk-Return Tradeoff After All." *Journal of Financial Economics 76*, 509–548.

Giraitis, L., R. Leipus, P. M. Robinson, and D. Surgailis. (2004). "LARCH, Leverage, and Long Memory." *Journal of Financial Econometrics 2*, 177–210.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. (1993). "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance 48*, 1779–1801.

Hansen, P. R., and A. Lunde. (2006). "Realized Variance and Market Microstructure Noise." *Journal of Business and Economic Statistics 24,* 127–161.

Harvey, A. C. (1998). "Long Memory in Stochastic Volatility." In J. Knight and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*. Oxford, UK: Butterworth-Heineman.

Harvey, A. C., and N. Shephard. (1996). "The Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns." *Journal of Business and Economic Statistics 14*, 429–434.

Hasbrouck, J. (2003). "Intraday Price Formation in U.S. Equity Index Markets." *Journal of Finance 58*, 2375–2400.

Heston, S. L. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies 6*, 327–347.

Johannes, M. (2004). "The Statistical and Economic Role of Jumps." *Journal of Finance 59*, 227–260.

Kim, D., and S. J. Kon. (1994). "Alternative Models for the Conditional Heteroskedasticity of Stock Returns." *Journal of Business 67*, 563–598.

Litvinova, J. (2004). "Volatility Asymmetry in High Frequency Data." PhD Thesis, Department of Economics, Duke University.

Lundblad, Christian T. (2005). "The Risk Return Tradeoff in the Long Run: 1836–2003." Working Paper, Department of Finance, Indiana University.

Nelson, Daniel B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica 59*, 347–370.

Perello, J., J. Masoliver, and N. Anento. (2004). "A Comparison between Several Correlated Stochastic Volatility Models." *Physica A 344*, 134–137.

Perello, J., J. Masoliver, and J.-P. Bouchaud. (2004). "Multiple Time Scales in Volatility and Leverage Correlations: A Stochastic Volatility Model." *Applied Mathematical Finance 11*, 27–50.

Ryan, F. (2006). "Estimation and Simulation of the S&P 500 Futures Index: The ADS-ACD Model." Work in Progress, Department of Economics, Duke University.

Rydberg, T. H., and N. Shephard. (2003). "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models." *Journal of Financial Econometricss 1*, 2–25.

Scott, L. O. (1987). "Option Pricing when the Variance Changes Randomly: Theory, Estimation and an Application." *Journal of Financial and Quantitative Analysis 22*, 419–438.

Tauchen, G. (2004). "Recent Developments in Stochastic Volatility: Statistical Modelling and General Equilibrium." Working Paper, Department of Economics, Duke University.

Tauchen, G. (2005). "Stochastic Volatility in General Equilibrium." Working Paper, Department of Economics, Duke University.

Tauchen, G., H. Zhang, and M. Liu. (1996). "Volume, Volatility and Leverage: A Dynamic Analysis." *Journal of Econometrics 74*, 177–208.

Todorov, V. (2005). "Econometric Analysis of Jump-Driven Stochastic Volatility Models." Working Paper, Department of Economics, Duke University.

Tran, D. T. (2006). "Stochastic Volatility, Intraday Seasonality, and Market Microstructure Noise." Work in Progress, Department of Economics, Duke University.

Wu, G. (2001). "The Determinants of Asymmetric Volatility." *Review of Financial Studies 14*, 837–859.

Yu, J. (2005a). "On Leverage in a Stochastic Volatility Model." *Journal of Econometrics 127*, 165–178.

Yu, J. (2005b). "Is No News Good News? Reconciling Evidence from ARCH and Stochastic Volatility Models." Working Paper, Department of Economics, Singapore Management University.