*Econometrica*, Vol. 71, No. 2 (March, 2003), 579–625

# MODELING AND FORECASTING REALIZED VOLATILITY

BY TORBEN G. ANDERSEN, TIM BOLLERSLEV, FRANCIS X. DIEBOLD, AND PAUL LABYS[1]

We provide a framework for integration of high-frequency intraday data into the measurement, modeling, and forecasting of daily and lower frequency return volatilities and return distributions. Building on the theory of continuous-time arbitrage-free price processes and the theory of quadratic variation, we develop formal links between realized volatility and the conditional covariance matrix. Next, using continuously recorded observations for the Deutschemark/Dollar and Yen/Dollar spot exchange rates, we find that forecasts from a simple long-memory Gaussian vector autoregression for the logarithmic daily realized volatilities perform admirably. Moreover, the vector autoregressive volatility forecast, coupled with a parametric lognormal-normal mixture distribution produces well-calibrated density forecasts of future returns, and correspondingly accurate quantile predictions. Our results hold promise for practical modeling and forecasting of the large covariance matrices relevant in asset pricing, asset allocation, and financial risk management applications.

KEYWORDS: Continuous-time methods, quadratic variation, realized volatility, high-frequency data, long memory, volatility forecasting, density forecasting, risk management.

## 1. INTRODUCTION

THE JOINT DISTRIBUTIONAL characteristics of asset returns are pivotal for many issues in financial economics. They are the key ingredients for the pricing of financial instruments, and they speak directly to the risk-return tradeoff central to portfolio allocation, performance evaluation, and managerial decision-making. Moreover, they are intimately related to the fractiles of conditional portfolio return distributions, which govern the likelihood of extreme shifts in portfolio value and are therefore central to financial risk management, figuring prominently in both regulatory and private-sector initiatives.

The most critical feature of the conditional return distribution is arguably its second moment structure, which is empirically the dominant time-varying characteristic of the distribution. This fact has spurred an enormous literature on

[1] This research was supported by the National Science Foundation. We are grateful to Olsen and Associates, who generously made available their intraday exchange rate data. For insightful suggestions and comments we thank three anonymous referees and the Co-Editor, as well as Kobi Bodoukh, Sean Campbell, Rob Engle, Eric Ghysels, Atsushi Inoue, Eric Renault, Jeff Russell, Neil Shephard, Til Schuermann, Clara Vega, Ken West, and seminar participants at BIS (Basel), Chicago, CIRANO/Montreal, Emory, Iowa, Michigan, Minnesota, NYU, Penn, Rice, UCLA, UCSB, the June 2000 Meeting of the WFA, the July 2001 NSF/NBER Conference on Forecasting and Empirical Methods in Macroeconomics and Finance, the November 2001 NBER Meeting on Financial Risk Management, and the January 2002 North American Meeting of the Econometric Society. The views expressed by Paul Labys are strictly his own and do not necessarily reflect the views or opinions of Charles River Associates.

the modeling and forecasting of return volatility.[2] Over time, the availability of data for increasingly shorter return horizons has allowed the focus to shift from modeling at quarterly and monthly frequencies to the weekly and daily horizons. Forecasting performance has improved with the incorporation of more data, not only because high-frequency volatility turns out to be highly predictable, but also because the information in high-frequency data proves useful for forecasting at longer horizons, such as monthly or quarterly.

In some respects, however, progress in volatility modeling has slowed in the last decade. First, the availability of truly high-frequency intraday data has made scant impact on the modeling of, say, daily return volatility. It has become apparent that standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data, and models specified directly for the intraday data generally fail to capture the longer interdaily volatility movements sufficiently well. As a result, standard practice is still to produce forecasts of daily volatility from daily return observations, even when higher-frequency data are available. Second, the focus of volatility modeling continues to be decidedly very low-dimensional, if not universally univariate. Many multivariate ARCH and stochastic volatility models for time-varying return volatilities and conditional distributions have, of course, been proposed (see, for example, the surveys by Bollerslev, Engle, and Nelson (1994) and Ghysels, Harvey, and Renault (1996)), but those models generally suffer from a curse-of-dimensionality problem that severely constrains their practical application. Consequently, it is rare to see substantive applications of those multivariate models dealing with more than a few assets simultaneously.

In view of such difficulties, finance practitioners have largely eschewed formal volatility modeling and forecasting in the higher-dimensional situations of practical relevance, relying instead on ad hoc methods, such as simple exponential smoothing coupled with an assumption of conditionally normally distributed returns.[3] Although such methods rely on counterfactual assumptions and are almost surely suboptimal, practitioners have been swayed by considerations of feasibility, simplicity, and speed of implementation in high-dimensional environments.

Set against this rather discouraging background, we seek to improve matters. We propose a new and rigorous framework for volatility forecasting and conditional return fractile, or value-at-risk (VaR), calculation, with two key properties. First, it effectively exploits the information in intraday return data, *without having to explicitly model the intraday data*, producing significant improvements in predictive performance relative to standard procedures that rely on daily data alone. Second, it achieves a simplicity and ease of implementation, that, for example, holds promise for high-dimensional return volatility modeling.

[2] Here and throughout, we use the generic term "volatilities" in reference both to variances (or standard deviations) and covariances (or correlations). When important, the precise meaning will be clear from the context.

[3] This approach is exemplified by the highly influential "RiskMetrics" of J. P. Morgan (1997).

We progress by focusing on an empirical measure of daily return variability called *realized volatility*, which is easily computed from high-frequency intraperiod returns. The theory of quadratic variation suggests that, under suitable conditions, realized volatility is an unbiased and highly efficient estimator of return volatility, as discussed in Andersen, Bollerslev, Diebold, and Labys (2001) (henceforth ABDL) as well as in concurrent work by Barndorff-Nielsen and Shephard (2002a, 2001).[4] Building on the notion of continuous-time arbitrage-free price processes, we advance in several directions, including rigorous theoretical foundations, multivariate emphasis, explicit focus on forecasting, and links to modern risk management via modeling of the entire conditional density.

Empirically, by treating volatility as observed rather than latent, our approach facilitates modeling and forecasting using simple methods based directly on observable variables.[5] We illustrate the ideas using the highly liquid U.S. dollar ($), Deutschemark (DM), and Japanese yen (¥) spot exchange rate markets. Our full sample consists of nearly thirteen years of continuously recorded spot quotations from 1986 through 1999. During that period, the dollar, Deutschemark, and yen constituted the main axes of the international financial system, and thus spanned the majority of the systematic currency risk faced by large institutional investors and international corporations.

We break the sample into a ten-year "in-sample" estimation period, and a subsequent two-and-a-half-year "out-of-sample" forecasting period. The basic distributional and dynamic characteristics of the foreign exchange returns and realized volatilities during the in-sample period have been analyzed in detail by ABDL (2000a, 2001).[6] Three pieces of their results form the foundation on which the empirical analysis of this paper is built. First, although raw returns are clearly leptokurtic, returns standardized by realized volatilities are approximately Gaussian. Second, although the distributions of realized volatilities are clearly right-skewed, the distributions of the logarithms of realized volatilities are approximately Gaussian. Third, the long-run dynamics of realized logarithmic volatilities are well approximated by a fractionally-integrated long-memory process.

Motivated by the three ABDL empirical regularities, we proceed to estimate and evaluate a multivariate model for the logarithmic realized volatilities: a fractionally-integrated Gaussian vector autoregression (VAR). Importantly,

[4] Earlier work by Comte and Renault (1998), within the context of estimation of a long-memory stochastic volatility model, helped to elevate the discussion of realized and integrated volatility to a more rigorous theoretical level.

[5] The direct modeling of observable volatility proxies was pioneered by Taylor (1986), who fit ARMA models to absolute and squared returns. Subsequent empirical work exploiting related univariate approaches based on improved realized volatility measures from a heuristic perspective includes French, Schwert, and Stambaugh (1987) and Schwert (1989), who rely on daily returns to estimate models for monthly realized U.S. equity volatility, and Hsieh (1991), who fits an AR(5) model to a time series of daily realized logarithmic volatilities constructed from 15-minute S&P500 returns.

[6] Strikingly similar and hence confirmatory qualitative findings have been obtained from a separate sample consisting of individual U.S. stock returns in Andersen, Bollerslev, Diebold, and Ebens (2001).

our approach explicitly permits measurement errors in the realized volatilities. Comparing the resulting volatility forecasts to those obtained from currently popular daily volatility models and more complicated high-frequency models, we find that our simple Gaussian VAR forecasts generally produce superior forecasts. Furthermore, we show that, given the theoretically motivated and empirically plausible assumption of normally distributed returns conditional on the realized volatilities, the resulting lognormal-normal mixture forecast distribution provides conditionally well-calibrated density forecasts of returns, from which we obtain accurate estimates of conditional return quantiles.

In the remainder of this paper, we proceed as follows. We begin in Section 2 by formally developing the relevant quadratic variation theory within a standard frictionless arbitrage-free multivariate pricing environment. In Section 3 we discuss the practical construction of realized volatilities from high-frequency foreign exchange returns. Next, in Section 4 we summarize the salient distributional features of returns and volatilities, which motivate the long-memory trivariate Gaussian VAR that we estimate in Section 5. In Section 6 we compare the resulting volatility point forecasts to those obtained from more traditional volatility models. We also evaluate the success of the density forecasts and corresponding VaR estimates generated from the long-memory Gaussian VAR in conjunction with a lognormal-normal mixture distribution. In Section 7 we conclude with suggestions for future research and discussion of issues related to the practical implementation of our approach for other financial instruments and markets.

## 2. QUADRATIC RETURN VARIATION AND REALIZED VOLATILITY

We consider an $n$-dimensional price process defined on a complete probability space, $(\Omega, \mathcal{F}, P)$, evolving in continuous time over the interval $[0, T]$, where $T$ denotes a positive integer. We further consider an information filtration, i.e., an increasing family of $\sigma$-fields, $(\mathcal{F}_t)_{t\in[0,T]} \subseteq \mathcal{F}$, which satisfies the usual conditions of $P$-completeness and right continuity. Finally, we assume that the asset prices through time $t$, including the relevant state variables, are included in the information set $\mathcal{F}_t$.

Under the standard assumptions that the return process does not allow for arbitrage and has a finite instantaneous mean, the asset price process, as well as smooth transformations thereof, belong to the class of special semi-martingales, as detailed by Back (1991). A fundamental result of stochastic integration theory states that such processes permit a unique canonical decomposition. In particular, we have the following characterization of the logarithmic asset price vector process, $p = (p(t))_{t\in[0,T]}$.

PROPOSITION 1: *For any n-dimensional arbitrage-free vector price process with finite mean, the logarithmic vector price process, p, may be written uniquely as the sum of a finite variation and predictable mean component,* $A = (A_1, \dots, A_n)$, *and* *a local martingale,* $M = (M_1, \ldots, M_n)$. *These may each be decomposed into a continuous sample-path and jump part,*

$$p(t) = p(0) + A(t) + M(t) = p(0) + A^c(t) + \Delta A(t) + M^c(t) + \Delta M(t), \tag{1}$$

*where the finite-variation predictable components,* $A^c$ *and* $\Delta A$, *are respectively continuous and pure jump processes, while the local martingales,* $M^c$ *and* $\Delta M$, *are respectively continuous sample-path and compensated jump processes, and by definition* $M(0) \equiv A(0) \equiv 0$. *Moreover, the predictable jumps are associated with genuine jump risk, in the sense that if* $\Delta A(t) \neq 0$, *then*

$$P[\mathrm{sgn}(\Delta A(t)) = -\mathrm{sgn}(\Delta A(t) + \Delta M(t))] > 0, \tag{2}$$

*where* $\mathrm{sgn}(x) \equiv 1$ *for* $x \geq 0$ *and* $\mathrm{sgn}(x) \equiv -1$ *for* $x < 0$.

Equation (1) is standard; see, for example, Protter (1992, Chapter 3). Equation (2) is an implication of the no-arbitrage condition. Whenever $\Delta A(t) \neq 0$, there is a predictable jump in the price—the timing and size of the jump is perfectly known (just) prior to the jump event— and hence there is a trivial arbitrage (with probability one) unless there is a simultaneous jump in the martingale component, $\Delta M(t) \neq 0$. Moreover, the concurrent martingale jump must be large enough (with strictly positive probability) to overturn the gain associated with a position dictated by $\mathrm{sgn}(\Delta A(t))$.

Proposition 1 provides a general characterization of the asset return process. We denote the (continuously compounded) return over $[t-h, t]$ by $r(t, h) = p(t) - p(t-h)$. The cumulative return process from $t = 0$ onward, $r = (r(t))_{t \in [0,T]}$, is then $r(t) \equiv r(t, t) = p(t) - p(0) = A(t) + M(t)$. Clearly, $r(t)$ inherits all the main properties of $p(t)$ and may likewise be decomposed uniquely into the predictable and integrable mean component, $A$, and the local martingale, $M$. The predictability of $A$ still allows for quite general properties in the (instantaneous) mean process; for example, it may evolve stochastically and display jumps. Nonetheless, the continuous component of the mean return must have smooth sample paths compared to those of a nonconstant continuous martingale—such as a Brownian motion—and any jump in the mean must be accompanied by a corresponding predictable jump (of unknown magnitude) in the compensated jump martingale, $\Delta M$. Consequently, there are two types of jumps in the return process, namely, predictable jumps where $\Delta A(t) \neq 0$ and equation (2) applies, and purely unanticipated jumps where $\Delta A(t) = 0$ but $\Delta M(t) \neq 0$. The latter jump event will typically occur when unanticipated news hit the market. In contrast, the former type of predictable jump may be associated with the release of information according to a predetermined schedule, such as macroeconomic news releases or company earnings reports. Nonetheless, it is worth noting that any slight uncertainty about the precise timing of the news (even to within a fraction of a second) invalidates the assumption of predictability and removes the jump in the mean process. If there are no such perfectly anticipated news releases, the predictable, finite variation mean return, $A$, may still evolve stochastically, but it will have continuous sample paths. This constraint is implicitly invoked in the vast majority of the continuous-time models employed in the literature.[7]

Because the return process is a semi-martingale it has an associated quadratic variation process. Quadratic variation plays a critical role in our theoretical developments. The following proposition enumerates some essential properties of the quadratic return variation process.[8]

PROPOSITION 2: *For any n-dimensional arbitrage-free price process with finite mean, the quadratic variation $n \times n$ matrix process of the associated return process, $[r, r] = \{[r, r]_t\}_{t\in[0,T]}$, is well-defined. The* $i$th *diagonal element is called the quadratic variation process of the* $i$th *asset return while the* $ij$th *off-diagonal element,* $[r_i, r_j]$, *is called the quadratic covariation process between asset returns i and j. The quadratic variation and covariation processes have the following properties*:

(i) *For an increasing sequence of random partitions of* $[0, T], 0 = \tau_{m,0} \leq \tau_{m,1} \leq \dots$, *such that* $\sup_{j\geq 1}(\tau_{m,j+1} - \tau_{m,j}) \to 0$ *and* $\sup_{j\geq 1} \tau_{m,j} \to T$ *for* $m \to \infty$ *with probability one*, *we have that*

$$\lim_{m\to\infty}\left\{\sum_{j\geq 1}[r(t\wedge\tau_{m,j})-r(t\wedge\tau_{m,j-1})][r(t\wedge\tau_{m,j})-r(t\wedge\tau_{m,j-1})]'\right\}\to[r,r]_t, \tag{3}$$

*where* $t \wedge \tau \equiv \min(t, \tau), t \in [0, T]$, *and the convergence is uniform on* $[0, T]$ *in probability.*

(ii) *If the finite variation component*, $A$, *in the canonical return decomposition in Proposition* 1 *is continuous*, *then*

$$[r_i, r_j]_t = [M_i, M_j]_t = [M_i^c, M_j^c]_t + \sum_{0\leq s\leq t} \Delta M_i(s)\Delta M_j(s). \tag{4}$$

The terminology of quadratic variation is justified by property (i) of Proposition 2. Property (ii) reflects the fact that the quadratic variation of continuous finite-variation processes is zero, so the mean component becomes irrelevant for the quadratic variation.[9] Moreover, jump components only contribute to the quadratic covariation if there are simultaneous jumps in the price path for the $i$th and $j$th asset, whereas the squared jump size contributes one-for-one to the quadratic variation. The quadratic variation process measures the realized sample-path variation of the squared return processes. Under the weak auxiliary

[7] This does not appear particularly restrictive. For example, if an announcement is pending, a natural way to model the arrival time is according to a continuous hazard function. Then the probability of a jump within each (infinitesimal) instant of time is zero—there is no discrete probability mass—and by arbitrage there cannot be a predictable jump.

[8] All of the properties in Proposition 2 follow, for example, from Protter (1992, Chapter 2).

[9] In the general case with predictable jumps the last term in equation (4) is simply replaced by $\sum_{0\leq s\leq t} \Delta r_i(s)\Delta r_j(s)$, where $\Delta r_i(s) \equiv \Delta A_i(s) + \Delta M_i(s)$ explicitly incorporates both types of jumps. However, as discussed above, this case is arguably of little interest from a practical empirical perspective.

condition ensuring property (ii), this variation is exclusively induced by the innovations to the return process. As such, the quadratic covariation constitutes, in theory, a unique and invariant ex-post realized volatility measure that is essentially model free. Notice that property (i) also suggests that we may approximate the quadratic variation by cumulating cross-products of high-frequency returns.[10] We refer to such measures, obtained from actual high-frequency data, as *realized volatilities*.

The above results suggest that the quadratic variation is the dominant determinant of the return covariance matrix, especially for shorter horizons. Specifically, the variation induced by the genuine return innovations, represented by the martingale component, locally is an order of magnitude larger than the return variation caused by changes in the conditional mean.[11] We have the following theorem which generalizes previous results in ABDL (2001).

THEOREM 1: *Consider an n-dimensional square-integrable arbitrage-free logarithmic price process with a continuous mean return, as in property* (ii) *of Proposition* 2. *The conditional return covariance matrix at time $t$ over $[t, t+h]$, where $0 \le t \le t+h \le T$, is then given by*

$$\text{(5)} \qquad \operatorname{cov}(r(t+h,h) \mid \mathcal{F}_t) = E([r,r]_{t+h} - [r,r]_t \mid \mathcal{F}_t) + \Gamma_A(t+h,h) + \Gamma_{AM}(t+h,h) + \Gamma'_{AM}(t+h,h),$$

*where $\Gamma_A(t+h,h) = \operatorname{cov}(A(t+h) - A(t) \mid \mathcal{F}_t)$ and $\Gamma_{AM}(t+h,h) = E(A(t+h)[M(t+h) - M(t)]' \mid \mathcal{F}_t)$.*

PROOF: From equation (1), $r(t+h,h) = [A(t+h) - A(t)] + [M(t+h) - M(t)]$. The martingale property implies $E(M(t+h) - M(t) \mid \mathcal{F}_t) = E([M(t+h) - M(t)]A(t) \mid \mathcal{F}_t) = 0$, so, for $i, j \in \{1, \dots, n\}$, $\operatorname{cov}([A_i(t+h) - A_i(t)], [M_j(t+h) - M_j(t)] \mid \mathcal{F}_t) = E(A_i(t+h)[M_j(t+h) - M_j(t)] \mid \mathcal{F}_t)$. It therefore follows that $\operatorname{cov}(r(t+h,h) \mid \mathcal{F}_t) = \operatorname{cov}(M(t+h) - M(t) \mid \mathcal{F}_t) + \Gamma_A(t+h,h) + \Gamma_{AM}(t+h,h) + \Gamma'_{AM}(t+h,h)$. Hence, it only remains to show that the conditional covariance of the martingale term equals the expected value of the quadratic variation. We proceed by verifying the equality for an arbitrary element of the covariance matrix. If this is the $i$th diagonal element, we are studying a univariate square-integrable martingale and by Protter (1992, Chapter II.6, Corollary 3), we have $E(M_i^2(t+h)) = E([M_i, M_i]_{t+h})$, so

$$\begin{aligned} \operatorname{var}(M_i(t+h) - M_i(t) \mid \mathcal{F}_t) &= E([M_i, M_i]_{t+h} - [M_i, M_i]_t \mid \mathcal{F}_t) \\ &= E([r_i, r_i]_{t+h} - [r_i, r_i]_t \mid \mathcal{F}_t), \end{aligned}$$

[10] This has previously been discussed by Comte and Renault (1998) in the context of estimating the spot volatility for a stochastic volatility model corresponding to the derivative of the quadratic variation (integrated volatility) process.

[11] This same intuition underlies the consistent filtering results for continuous sample path diffusions in Merton (1980) and Nelson and Foster (1995).

where the second equality follows from equation (4) of Proposition 2. This confirms the result for the diagonal elements of the covariance matrix. An identical argument works for the off-diagonal terms by noting that the sum of two square-integrable martingales remains a square-integrable martingale and then applying the reasoning to each component of the polarization identity,

$$[M_i, M_j]_t = 1/2([M_i + M_j, M_i + M_j]_t - [M_i, M_i]_t - [M_j, M_j]_t).$$

In particular, it follows as above that

$$\begin{aligned} &E([M_i, M_j]_{t+h} - [M_i, M_j]_t \mid \mathcal{F}_t) \\ &\quad = 1/2[\mathrm{var}([M_i(t+h) + M_j(t+h)] - [(M_i(t) + M_j(t)] \mid \mathcal{F}_t) \\ &\qquad - \mathrm{var}(M_i(t+h) - M_i(t) \mid \mathcal{F}_t) - \mathrm{var}(M_j(t+h) - M_j(t) \mid \mathcal{F}_t)] \\ &\quad = \mathrm{cov}([M_i(t+h) - M_i(t)], [M_j(t+h) - M_j(t)] \mid \mathcal{F}_t). \end{aligned}$$

Equation (4) of Proposition 2 again ensures that this equals $E([r_i, r_j]_{t+h} - [r_i, r_j]_t \mid \mathcal{F}_t)$. *Q.E.D.*

Two scenarios highlight the role of the quadratic variation in driving the return volatility process. These important special cases are collected in a corollary that follows immediately from Theorem 1.

COROLLARY 1: *Consider an n-dimensional square-integrable arbitrage-free logarithmic price process, as described in Theorem* 1. *If the mean process,* $\{A(s) - A(t)\}_{s\in[t,\,t+h]}$, *conditional on information at time t is independent of the return innovation process,* $\{M(u)\}_{u\in[t,\,t+h]}$, *then the conditional return covariance matrix reduces to the conditional expectation of the quadratic return variation plus the conditional variance of the mean component, i.e., for* $0 \le t \le t+h \le T$,

$$\mathrm{cov}(r(t+h, h) \mid \mathcal{F}_t) = E([r, r]_{t+h} - [r, r]_t \mid \mathcal{F}_t) + \Gamma_A(t+h, h).$$

*If the mean process,* $\{A(s) - A(t)\}_{s\in[t,\,t+h]}$, *conditional on information at time t is a predetermined function over* $[t, t+h]$, *then the conditional return covariance matrix equals the conditional expectation of the quadratic return variation process, i.e., for* $0 \le t \le t+h \le T$,

$$\mathrm{cov}(r(t+h, h) \mid \mathcal{F}_t) = E([r, r]_{t+h} - [r, r]_t \mid \mathcal{F}_t). \tag{6}$$

Under the conditions leading to equation (6), the quadratic variation is *the* critical ingredient in volatility measurement and forecasting. This follows as the quadratic variation represents the actual variability of the return innovations, and the conditional covariance matrix is the conditional expectation of this quantity. Moreover, it implies that the time $t+h$ ex-post realized quadratic variation is an unbiased estimator for the return covariance matrix conditional on information at time $t$.

Although the corollary's strong implications rely upon specific assumptions, these sufficient conditions are not as restrictive as an initial assessment may suggest, and they are satisfied for a wide set of popular models. For example, a constant mean is frequently invoked in daily or weekly return models. Equation (6) further allows for deterministic intra-period variation in the conditional mean, induced by time-of-day or other calendar effects. Of course, equation (6) also accommodates a stochastic mean process as long as it remains a function, over the interval $[t, t+h]$, of variables in the time $t$ information set. Specification (6) does, however, preclude feedback effects from the random intra-period evolution of the system to the instantaneous mean. Although such feedback effects may be present in high-frequency returns, they are likely trivial in magnitude over daily or weekly frequencies, as we argue subsequently. It is also worth stressing that (6) is compatible with the existence of an asymmetric return-volatility relation (sometimes called a leverage effect), which arises from a correlation between the return innovations, measured as deviations from the conditional mean, and the innovations to the volatility process. In other words, the leverage effect is separate from a contemporaneous correlation between the return innovations and the instantaneous mean return. Furthermore, as emphasized above, equation (6) does allow for the return innovations over $[t-h, t]$ to impact the conditional mean over $[t, t+h]$ and onwards, so that the intra-period evolution of the system may still impact the future expected returns. In fact, this is how potential interaction between risk and return is captured in discrete-time stochastic volatility or ARCH models with leverage effects.

In contrast to equation (6), the first expression in Corollary 1 involving $\Gamma_A$ explicitly accommodates continually evolving random variation in the conditional mean process, although the random mean variation must be independent of the return innovations. Even with this feature present, the quadratic variation is likely an order of magnitude larger than the mean variation, and hence the former remains the critical determinant of the return volatility over shorter horizons. This observation follows from the fact that over horizons of length $h$, with $h$ small, the variance of the mean return is of order $h^2$, while the quadratic variation is of order $h$. It is, of course, an empirical question whether these results are a good guide for volatility measurement at relevant frequencies.[12] To illustrate the implications at a daily horizon, consider an asset return with standard deviation of 1% daily, or 15.8% annually, and a (large) mean return of 0.1%, or about 25% annually. The squared mean return is still only one-hundredth of the variance. The expected daily variation of the mean return is obviously smaller yet, unless the required daily return is assumed to behave truly erratically within the day. In fact, we would generally expect the within-day variance of the expected

[12] Merton (1982) provides a similar intuitive account of the continuous record *h-asymptotics*. These limiting results are also closely related to the theory rationalizing the quadratic variation formulas in Proposition 2 and Theorem 1.

daily return to be much smaller than the expected daily return itself. Hence, the daily return fluctuations induced by within-day variations in the mean return are almost certainly trivial. For a weekly horizon, similar calculations suggest that the identical conclusion applies.

The general case, covered by Theorem 1, allows for direct intra-period interaction between the return innovations and the instantaneous mean. This occurs, for example, when there is a leverage effect, or asymmetry, by which the volatility impacts the contemporaneous mean drift. In this—for some assets—empirically relevant case, a string of negative within-period return innovations will be associated with an increase in return volatility, which in turn raises the risk premium and the return drift. Relative to the corollary, the theorem involves the additional $\Gamma_{AM}$ terms. Nonetheless, the intuition discussed above remains intact. It is readily established that the $ik$th component of these terms may be bounded, $\{\Gamma_{AM}(t+h,h)\}_{i,k} \leq \{\mathrm{var}(A_i(t+h)-A_i(t) \mid \mathcal{F}_t)\}^{1/2}\{\mathrm{var}(M_k(t+h)-M_k(t) \mid \mathcal{F}_t)\}^{1/2}$, where the latter terms are of order $h$ and $h^{1/2}$ respectively, so the $\Gamma_{AM}$ terms are at most of order $h^{3/2}$, which again is dominated by the corresponding quadratic variation of order $h$. Moreover, this upper bound is quite conservative, because it allows for a correlation of unity, whereas typical correlations estimated from daily or weekly returns are much lower, de facto implying that the quadratic variation process is the main driving force behind the corresponding return volatility.

We now turn towards a more ambitious goal. Because the above results carry implications for the measurement and modeling of return volatility, it is natural to ask whether we can also infer something about the appropriate specification of the return generating process that builds on the realized volatility measures. Obviously, at the present level of generality, requiring only square integrability and absence of arbitrage, we cannot derive specific distributional results. Nonetheless, we may obtain a useful benchmark under somewhat more restrictive conditions, including a continuous price process, i.e., no jumps or, $\Delta M \equiv 0$. We first recall the martingale representation theorem.[13]

Proposition 3: *For any n-dimensional square-integrable arbitrage-free logarithmic price process, p, with continuous sample path and a full rank of the associated $n \times n$ quadratic variation process, $[r,r]_t$, we have a.s. (P) for all $0 \leq t \leq T$,*

$$r(t+h,h) = p(t+h) - p(t) = \int_0^h \mu_{t+s}\,ds + \int_0^h \sigma_{t+s}\,dW(s), \tag{7}$$

*where $\mu_s$ denotes an integrable predictable $n \times 1$ dimensional vector, $\sigma_s = (\sigma_{(i,j),s})_{i,j=1,\dots,n}$ is an $n \times n$ matrix, $W(s)$ is an $n \times 1$ dimensional standard Brownian motion, integration of a matrix or vector with respect to a scalar denotes component-wise integration, so that*

$$\int_0^h \mu_{t+s}\,ds = \left(\int_0^h \mu_{1,t+s}\,ds, \dots, \int_0^h \mu_{n,t+s}\,ds\right)',$$

[13] See, for example, Karatzas and Shreve (1991, Chapter 3).

*and integration of a matrix with respect to a vector denotes component-wise integration of the associated vector, so that*

$$(8)\qquad \int_0^h \sigma_{t+s}\, dW(s) = \left( \int_0^h \sum_{j=1,\ldots,n} \sigma_{(i,j),t+s}\, dW_j(s), \ldots, \int_0^h \sum_{j=1,\ldots,n} \sigma_{(n,j),t+s}\, dW_j(s) \right)'.$$

*Moreover, we have*

$$(9)\qquad P\left[ \int_0^h (\sigma_{(i,j),t+s})^2\, ds < \infty \right] = 1, \quad 1 \le i, j \le n.$$

*Finally, letting $\Omega_s = \sigma_s \sigma_s'$, the increments to the quadratic return variation process take the form*

$$(10)\qquad [r,r]_{t+h} - [r,r]_t = \int_0^h \Omega_{t+s}\, ds.$$

The requirement that the $n \times n$ matrix $[r,r]_t$ is of full rank for all $t$, implies that no asset is redundant at any time, so that no individual asset return can be spanned by a portfolio created by the remaining assets. This condition is not restrictive; if it fails, a parallel representation may be achieved on an extended probability space.[14]

We are now in position to state a distributional result that inspires our empirical modeling of the full return generating process in Section 6 below. It extends the results recently discussed by Barndorff-Nielsen and Shephard (2002a) by allowing for a more general specification of the conditional mean process and by accommodating a multivariate setting. It should be noted that for volatility forecasting, as discussed in Sections 5 and 6.1 below, we do not need the auxiliary assumptions invoked here.

THEOREM 2: *For any n-dimensional square-integrable arbitrage-free price process with continuous sample paths satisfying Proposition* 3, *and thus representation* (7), *with conditional mean and volatility processes $\mu_s$ and $\sigma_s$ independent of the innovation process $W(s)$ over $[t, t+h]$, we have*

$$(11)\qquad r(t+h,h) \mid \sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,h]} \sim N\left( \int_0^h \mu_{t+s}\, ds, \int_0^h \Omega_{t+s}\, ds \right),$$

*where $\sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,h]}$ denotes the $\sigma$-field generated by $(\mu_{t+s}, \sigma_{t+s})_{s\in[0,h]}$.*

[14] See Karatzas and Shreve (1991, Section 3.4).

PROOF: Clearly, $r(t+h, h) - \int_0^h \mu_{t+s}\, ds = \int_0^h \sigma_{t+s}\, dW(s)$ and $E\big(\int_0^h \sigma_{t+s} dW(s) \mid \sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,h]}\big) = 0$. We proceed by establishing the normality of $\int_0^h \sigma_{t+s}\, dW(s)$ conditional on the volatility path $\{\sigma_{t+s}\}_{s\in[0,h]}$. The integral is $n$-dimensional, and we define

$$\int_0^h \sigma_{t+s}\, dW(s) = \left(\int_0^h (\sigma_{(1),t+s})'\, dW(s), \ldots, \int_0^h (\sigma_{(n),t+s})'\, dW(s)\right)',$$

where $\sigma_{(i),s} = (\sigma_{(i,1),s}, \ldots, \sigma_{(i,n),s})'$, so that $\int_0^h (\sigma_{(i),t+s})'\, dW(s)$ denotes the $i$th element of the $n \times 1$ vector in equation (8). The vector is multivariate normal if and only if any (nonzero) linear combination of the elements are univariate normal. Each element of the vector represents a sum of independent stochastic integrals, as detailed in equation (8). Any nonzero linear combination of this $n$-dimensional vector will thus produce another linear combination of the same $n$ independent stochastic integrals. Moreover, the linear combination will be nonzero given the full rank condition of Proposition 3. It will therefore be normally distributed if each constituent component of the original vector in equation (8) is normally distributed conditional on the volatility path. A typical element of the sums in equation (8), representing the $j$th volatility factor loading of asset $i$ over $[t, t+h]$, takes the form, $I_{i,j}(t+h, h) \equiv \int_0^h \sigma_{(i,j),t+s}\, dW_j(s)$, for $1 \leq i, j \leq n$. Obviously, $I_{i,j}(t) \equiv I_{i,j}(t,t)$ is a continuous local martingale, and then by the "change of time" result (see, e.g., Protter (1992, Chapter II, Theorem 41)), it follows that $I_{i,j}(t) = B([I_{i,j}, I_{i,j}]_t)$, where $B(t)$ denotes a standard univariate Brownian motion. In addition,

$$\begin{aligned} I_{i,j}(t+h, h) &= I_{i,j}(t+h) - I_{i,j}(t) \\ &= B([I_{i,j}, I_{i,j}]_{t+h}) - B([I_{i,j}, I_{i,j}]_t), \end{aligned}$$

and this increment to the Brownian motion is distributed $N(0, [I_{i,j}, I_{i,j}]_{t+h} - [I_{i,j}, I_{i,j}]_t)$. Finally, the quadratic variation governing the variance of the Gaussian distribution is readily determined to be $[I_{i,j}, I_{i,j}]_{t+h} - [I_{i,j}, I_{i,j}]_t = \int_0^h (\sigma_{(i,j),t+s})^2\, ds$ (see, e.g., Protter (1992, Chapter II.6)), which is finite by equation (9) of Proposition 3. Conditional on the ex-post realization of the volatility path, the quadratic variation is given (measurable), and the conditional normality of $I_{i,j}(t+h, h)$ follows. Because both the mean and the volatility paths are independent of the return innovations over $[t, t+h]$, the mean is readily determined from the first line of the proof. This verifies the conditional normality asserted in equation (11). The only remaining issue is to identify the conditional return covariance matrix. For the $ik$th element of the matrix we have

$$
\begin{aligned}
&\operatorname{cov}\left[\int_0^h (\sigma_{(i),\,t+s})'\, dW(s), \int_0^h (\sigma_{(k),\,t+s})'\, dW(s) \mid \sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,\,h]}\right] \\
&\quad = E\left[\sum_{j=1,\,\dots,\,n} \int_0^h \sigma_{(i,\,j),\,t+s}\, dW_j(s)\right. \\
&\qquad\qquad \left. \cdot \sum_{j=1,\,\dots,\,n} \int_0^h \sigma_{(k,\,j),\,t+s}\, dW_j(s) \mid \sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,\,h]}\right] \\
&\quad = \sum_{j=1,\dots,n} E\left[\int_0^h \sigma_{(i,\,j),\,t+s}\sigma_{(k,\,j),\,t+s}\, ds \mid \sigma\{\mu_{t+s}, \sigma_{t+s}\}_{s\in[0,\,h]}\right] \\
&\quad = \sum_{j=1,\dots,n} \int_0^h \sigma_{(i,\,j),\,t+s}\sigma_{(k,\,j),\,t+s}\, ds \\
&\quad = \int_0^h (\sigma_{(i),\,t+s})'\sigma_{(k),\,t+s}\, ds \\
&\quad = \left(\int_0^h \sigma_{t+s}(\sigma_{t+s})'\, ds\right)_{ik} \\
&\quad = \left(\int_0^h \Omega_{t+s}\, ds\right)_{ik}.
\end{aligned}
$$

This confirms that each element of the conditional return covariance matrix equals the corresponding element of the variance term in equation (11). *Q.E.D.*

Notice that the distributional characterization in Theorem 2 is conditional on the ex-post sample-path realization of $(\mu_s, \sigma_s)_{s\in[t,\,t+h]}$. The theorem may thus appear to be of little practical relevance, because such realizations typically are not observable. However, Proposition 2 and equation (10) suggest that we may construct approximate measures of the realized quadratic variation, and hence of the conditional return variance, directly from high-frequency return observations. In addition, as discussed previously, for daily or weekly returns, the conditional mean variation is negligible relative to the return volatility. Consequently, ignoring the time variation of the conditional mean, it follows by Theorem 2 that the distribution of the daily returns, say, is determined by a normal mixture with the daily realized quadratic return variation governing the mixture.

Given the auxiliary assumptions invoked in Theorem 2, the normal mixture distribution is strictly only applicable if the price process has continuous sample paths and the volatility and mean processes are independent of the within-period return innovations. The latter implies a conditionally symmetric return distribution. This raises two main concerns. First, some recent evidence suggests the possibility of discrete jumps in asset prices, rendering sample paths discontinuous.[15] But these studies also find that jumps are infrequent and have a jump size distribution about which there is little consensus. Second, for some asset classes there is evidence of leverage effects that may indicate a correlation between concurrent return and volatility innovations. However, as argued above, such contemporaneous correlation effects are likely to be unimportant quantitatively at the daily or weekly horizon. Indeed, Theorem 2 allows for the more critical impact leading from the current return innovations to the volatility in subsequent periods (beyond time $t+h$), corresponding to the effect captured in the related discrete-time ARCH and stochastic volatility literature. We thus retain the normal mixture distribution as a natural starting point for our empirical work. However, if the return-volatility asymmetry is important and the forecast horizon, $h$, relatively long, say monthly or quarterly, then one may expect the empirical return distribution to display asymmetries that are incompatible with the symmetric return distribution (conditional on time $t$ information) implied by Theorem 2. One simple diagnostic is to check if the realized volatility-standardized returns over the relevant horizon fail to be normally distributed, as this will speak to the importance of incorporating jumps and/or contemporaneous return innovation-volatility interactions into the modeling framework.

In summary, the arbitrage-free setting imposes a semi-martingale structure that leads directly to the representation in Proposition 1 and the associated quadratic variation in Proposition 2. In addition, property (i) and equation (3) in Proposition 2 suggest a practical way to approximate the quadratic variation. Theorem 1 and the associated Corollary 1 reveal the intimate relation between the quadratic variation and the return volatility process. For the continuous sample path case, we further obtain the representation in equation (7), and the quadratic variation reduces by equation (10) to $\int_0^h \Omega_{t+s}\,ds$, which is often referred to as the *integrated volatility*. Theorem 2 consequently strengthens Theorem 1 by showing that the realized quadratic variation is not only a useful estimator of the ex-ante conditional volatility, but also, under auxiliary assumptions, identical to the realized integrated return volatility over the relevant horizon. Moreover, the theorem delivers a reference distribution for appropriately standardized returns. Taken as a whole, the results provide a general framework for integration of high-frequency intraday data into the measurement, modeling, and forecasting of daily and lower frequency return volatility and return distributions, tasks to which we now turn.

## 3. MEASURING REALIZED EXCHANGE RATE VOLATILITY

Practical implementation of the procedures suggested by the theory in Section 2 must confront the fact that no financial market provides a frictionless trading environment with continuous price recording. Consequently, the notion of quadratic return variation is an abstraction that, strictly speaking, cannot be

[15] See, for example, Andersen, Benzoni, and Lund (2002), Bates (2000), Bakshi, Cao, and Chen (1997), Pan (2002), and Eraker, Johannes, and Polson (2002).

observed. Nevertheless, the continuous-time arbitrage-free framework directly motivates the creation of our return series and associated volatility measures from high-frequency data. We do not claim that this provides exact counterparts to the (nonexisting) corresponding continuous-time quantities. Instead, we use the theory to guide and inform collection of the data, transformation of the data into volatility measures, and selection of the models used to construct conditional return volatility and density forecasts, after which we assess the usefulness of the theory through the lens of predictive accuracy.

### 3.1. *Data*

Our empirical analysis focuses on the spot exchange rates for the U.S. dollar, the Deutschemark, and the Japanese yen.[16] The raw data consist of all interbank DM/\$ and ¥/\$ bid/ask quotes displayed on the Reuters FXFX screen during the sample period, December 1, 1986 through June 30, 1999.[17] These quotes are merely indicative (that is, nonbinding) and subject to various market microstructure "frictions," including strategic quote positioning and standardization of the size of the quoted bid/ask spread. Such features are generally immaterial when analyzing longer horizon returns, but they may distort the statistical properties of the underlying "equilibrium" high-frequency intraday returns.[18] The sampling frequency at which such considerations become a concern is intimately related to market activity. For our exchange rate series, preliminary analysis based on the methods of ABDL (2000b) suggests that the use of equally-spaced thirty-minute returns strikes a satisfactory balance between the accuracy of the continuous-record asymptotics underlying the construction of our realized volatility measures on the one hand, and the confounding influences from market microstructure frictions on the other.[19]

The actual construction of the returns follows Müller et al. (1990) and Dacorogna et al. (1993). First, we calculate thirty-minute prices from the linearly interpolated logarithmic average of the bid and ask quotes for the two ticks immediately before and after the thirty-minute time stamps throughout the global 24-hour trading day. Second, we obtain thirty-minute returns as the first difference of the logarithmic prices.[20] In order to avoid modeling specific weekend

[16] Before the advent of the Euro, the dollar, Deutschemark and yen were the most actively traded currencies in the foreign exchange market, with the DM/\$ and ¥/\$ accounting for nearly fifty percent of the daily trading volume, according to a 1996 survey by the Bank for International Settlements.

[17] The data comprise several million quotes kindly supplied by Olsen & Associates. Average daily quotes number approximately 4,500 for the Deutschemark and 2,000 for the Yen.

[18] See Bai, Russell, and Tiao (2000) and Zumbach, Corsi, and Trapletti (2002) for discussion and quantification of various aspects of microstructure bias in the context of realized volatility.

[19] An alternative approach would be to utilize all of the observations by explicitly modeling the high-frequency market microstructure. That approach, however, is much more complicated and subject to numerous pitfalls of its own.

[20] We follow the standard convention of the interbank market by measuring the exchange rates and computing the corresponding rates of return from the prices of \$1 expressed in terms of DM and ¥, i.e., DM/\$ and ¥/\$. Similarly, we express the cross rate as the price of one DM in terms of ¥, i.e., ¥/DM.

effects, we exclude all of the returns from Friday 21:00 GMT until Sunday 21:00 GMT. Similarly, to avoid complicating the inference by the decidedly slower trading activity during certain holiday periods, we delete a number of other inactive days from the sample. We are left with a bivariate series of thirty-minute DM/\$ and ¥/\$ returns spanning a total of 3,045 days. In order to explicitly distinguish the empirically constructed continuously compounded discretely sampled returns and corresponding volatility measures from the theoretical counterparts in Section 2, we will refer to the former by time subscripts. Specifically, for the half-hour returns, $r_{t+\Delta,\Delta}$, $t=\Delta, 2\Delta, 3\Delta, \ldots, 3045$, where $\Delta = 1/48 \approx 0.0208$. Also, for notational simplicity we label the corresponding daily returns by a single time subscript, so that $r_{t+1} \equiv r_{t+1,1} \equiv r_{t+\Delta,\Delta} + r_{t+2\Delta,\Delta} + \cdots r_{t+1,\Delta}$ for $t = 1, 2, \ldots, 3045$. Finally, we partition the full sample period into an "in-sample" estimation period covering the 2,449 days from December 1, 1986 through December 1, 1996, and a genuine "out-of-sample" forecast evaluation period covering the 596 days from December 2, 1996 through June 30, 1999.[21]

## 3.2. *Construction of Realized Volatilities*

The preceding discussion suggests that meaningful ex-post interdaily volatility measures may be constructed by cumulating cross-products of intraday returns sampled at an appropriate frequency, such as thirty minutes. In particular, based on the bivariate vector of thirty-minute DM/\$ and ¥/\$ returns, i.e., with $n = 2$, we define the *h-day realized volatility*, for $t = 1, 2, \ldots, 3045$, $\Delta = 1/48$, by

$$V_{t,h} \equiv \sum_{j=1,\ldots,h/\Delta} r_{t-h+j\cdot\Delta,\Delta} \cdot r'_{t-h+j\cdot\Delta,\Delta} = R'_{t,h} R_{t,h}, \tag{12}$$

where the $(h/\Delta) \times n$ matrix, $R_{t,h}$, is defined by $R'_{t,h} \equiv (r_{t-h+\Delta,\Delta}, r_{t-h+2\cdot\Delta,\Delta}, \ldots, r_{t,\Delta})$. As before, we simplify the notation for the daily horizon by defining $V_t \equiv V_{t,1}$. The $V_{t,h}$ measure constitutes the empirical counterpart to the $h$-period quadratic return variation and, for the continuous sample path case, the integrated volatility. In fact, by Proposition 2, as the sampling frequency of the intraday returns increases, or $\Delta \to 0$, $V_{t,h}$ converges almost surely to the quadratic variation.

One issue frequently encountered in multivariate volatility modeling is that constraints must be imposed to guarantee positive definiteness of estimated covariance matrices. Even for relatively low-dimensional cases such as three or four assets, imposition and verification of conditions that guarantee positive definiteness can be challenging; see, for example, the treatment of multivariate GARCH processes in Engle and Kroner (1995). Interestingly, in contrast, it is straightforward to establish positive definiteness of $V_{t,h}$. The following lemma follows from the correspondence between our realized volatility measures and

[21] All of the empirical results in ABDL (2000a, 2001), which in part motivate our approach, were based on data for the in-sample period only, justifying the claim that our forecast evaluation truly is "out-of-sample."

standard unconditional sample covariance matrix estimators which, of course, are positive semi-definite by construction.

LEMMA 1: *If the columns of $R_{t,h}$ are linearly independent, then the realized covariance matrix, $V_{t,h}$, defined in equation* (12) *is positive definite*.

PROOF: It suffices to show that $a'V_{t,h}a > 0$ for all nonzero $a$. Linear independence of the columns of $R_{t,h}$ ensures that $b_{t,h} = R_{t,h}a \neq 0$, $\forall a \in \mathbb{R}^n \backslash \{0\}$, and in particular that at least one of the elements of $b_{t,h}$ is nonzero. Hence $a'V_{t,h}a = a'R'_{t,h}R_{t,h}a = b'_{t,h}b_{t,h} = \sum_{j=1,\ldots,h/\Delta}(b_{t,h})_j^2 > 0$, $\forall a \in \mathbb{R}^n \backslash \{0\}$. *Q.E.D.*

The fact that positive definiteness of the realized covariance matrix is virtually assured, even in high-dimensional settings, is encouraging. However, the lemma also points to a problem that will arise for *extremely* high-dimensional systems. The assumption of linear independence of the columns of $R_{t,h}$, although weak, will ultimately be violated as the dimension of the price vector increases relative to the sampling frequency of the intraday returns. Specifically, for $n > h/\Delta$ the rank of the $R_{t,h}$ matrix is obviously less than $n$, so $R'_t R_t = V_t$ will not have full rank and it will fail to be positive definite. Hence, although the use of $V_t$ facilitates rigorous measurement of conditional volatility in much higher dimensions than is feasible with most alternative approaches, it does not allow the dimensionality to become arbitrarily large. Concretely, the use of thirty-minute returns, corresponding to $1/\Delta = 48$ intraday observations, for construction of daily realized volatility measures, implies that positive definiteness of $V_t$ requires $n$, the number of assets, to be no larger than 48.

Because realized volatility $V_t$ is observable, we can model and forecast it using standard and relatively straightforward time-series techniques. The diagonal elements of $V_t$, say $v_{t,1}$ and $v_{t,2}$, correspond to the daily DM/$ and ¥/$ realized variances, while the off-diagonal element, say $v_{t,12}$, represents the daily realized covariance between the two rates. We could then model $\text{vech}(V_t) = (v_{t,1}, v_{t,12}, v_{t,2})'$ directly but, for reasons of symmetry, we replace the realized covariance with the realized variance of the ¥/DM cross rate which may be done, without loss of generality, in the absence of triangular arbitrage, resulting in a system of three realized volatilities.

Specifically, by absence of triangular arbitrage, the continuously compounded return on the ¥/DM cross rate must be equal to the difference between the ¥/$ and DM/$ returns, which has two key consequences. First, it implies that, even absent direct data on the ¥/DM rate, we can infer the cross rate using the DM/$ and ¥/$ data. Consequently, the realized cross-rate variance, $v_{t,3}$, may be constructed by summing the implied thirty-minute squared cross-rate returns,

$$v_{t,3} = \sum_{j=1,\ldots,1/\Delta} [(-1,1)' \cdot r_{t-1+j\cdot\Delta,\Delta}]^2. \tag{13}$$

Second, because this implies that $v_{t,3} = v_{t,1} + v_{t,2} - 2v_{t,12}$, we can infer the realized covariance from the three realized volatilities,

$$v_{t,12} = 1/2 \cdot (v_{t,1} + v_{t,2} - v_{t,3}). \tag{14}$$

TABLE I

DAILY RETURN DISTRIBUTIONS

| | Mean | Std. Dev. | Skewness | Kurtosis | $Q(20)$[d] | $Q^2(20)$[e] |
|---|---|---|---|---|---|---|
| *Returns*[a] | | | | | | |
| DM/$ | −0.007 | 0.700 | 0.003 | 5.28 | 14.13 | 140.3 |
| ¥/$ | −0.010 | 0.692 | −0.129 | 6.64 | 27.88 | 142.6 |
| Portfolio[c] | −0.008 | 0.630 | −0.046 | 5.81 | 20.30 | 111.5 |
| *Standardized Returns*[a,b] | | | | | | |
| DM/$ | −0.007 | 0.993 | 0.032 | 2.57 | 19.42 | 23.32 |
| ¥/$ | 0.007 | 0.964 | −0.053 | 2.66 | 32.13 | 24.83 |
| Portfolio[c] | −0.005 | 0.993 | 0.028 | 2.61 | 25.13 | 29.20 |

[a] The daily returns cover December 1, 1986 through December 1, 1996.
[b] The bottom panel refers to the distribution of daily returns standardized by realized volatility.
[c] Portfolio refers to returns on an equally-weighted portfolio.
[d] Ljung-Box test statistics for up to twentieth order serial correlation in returns.
[e] Ljung-Box test statistics for up to twentieth order serial correlation in squared returns.

Building on this insight, we infer the covariance from the three variances, $v_t \equiv (v_{t,1}, v_{t,2}, v_{t,3})'$, and the identity in equation (14) instead of directly modeling $\text{vech}(V_t)$.[22]

We now turn to a discussion of the pertinent empirical regularities that guide our specification of the trivariate forecasting model for the three DM/$, ¥/$, and ¥/DM volatility series.

## 4. PROPERTIES OF EXCHANGE RATE RETURNS AND REALIZED VOLATILITIES

The in-sample distributional features of the DM/$ and ¥/$ returns and the corresponding realized volatilities have been characterized previously by ABDL (2000a, 2001).[23] Here we briefly summarize those parts of the ABDL results that are relevant for the present inquiry. We also provide new results for the ¥/DM cross rate volatility and an equally-weighted portfolio that explicitly incorporate the realized covariance measure discussed above.

### 4.1. *Returns*

The statistics in the top panel of Table I refer to the two daily dollar-denominated returns, $r_{t,1}$ and $r_{t,2}$, and the equally-weighted portfolio,

[22] The no-triangular-arbitrage restrictions are, of course, not available outside the world of foreign exchange. However, these restrictions are in no way crucial to our general approach, as the realized variances and covariances could all be modeled directly. We choose to substitute out the realized covariance in terms of the cross-rate because it makes for a clean and unified presentation of the empirical work, allowing us to exploit the approximate lognormality of the realized variances (discussed below).

[23] For a prescient early contribution along these lines, see also Zhou (1996).

$1/2 \cdot (r_{t,1} + r_{t,2})$. As regards unconditional distributions, all three return series are approximately symmetric with zero mean. However, the sample kurtoses indicate more probability mass in the center and in the tails of the distribution relative to the normal, which is confirmed by the kernel density estimates shown in Figure 1. As regards conditional distributions, the Ljung-Box test statistics indicate no serial correlation in returns, but strong serial correlation in squared returns.[24] The results are entirely consistent with the extensive literature documenting fat tails and volatility clustering in asset returns, dating at least to Mandelbrot (1963) and Fama (1965).

The statistics in the bottom panel of Table I refer to the distribution of the standardized daily returns $r_{t,1} \cdot v_{t,1}^{-1/2}$ and $r_{t,2} \cdot v_{t,2}^{-1/2}$, along with the standardized daily equally-weighted portfolio returns $1/2 \cdot (r_{t,1} + r_{t,2}) \cdot (1/4 \cdot v_{t,1} + 1/4 \cdot v_{t,2} + 1/2 \cdot v_{t,12})^{-1/2}$, or equivalently by equation (14), $1/2 \cdot (r_{t,1} + r_{t,2}) \cdot (1/2 \cdot v_{t,1} + 1/2 \cdot v_{t,2} - 1/4 \cdot v_{t,3})^{-1/2}$. The standardized-return results provide striking contrasts to the raw-return results. First, the sample kurtoses indicate that the standardized returns are well approximated by a Gaussian distribution, as confirmed by the kernel density estimates in Figure 1, which clearly convey the approximate normality. Second, in contrast to the raw returns, the standardized returns display no evidence of volatility clustering.[25]

Of course, the realized volatility used for standardizing the returns is only observable ex-post. Nonetheless, the result is in stark contrast to the typical finding that, when standardizing daily returns by the one-day-ahead forecasted variance from ARCH or stochastic volatility models, the resulting distributions are invariably leptokurtic, albeit less so than for the raw returns; see, for example, Baillie and Bollerslev (1989) and Hsieh (1989). In turn, this has motivated the widespread adoption of volatility models with non-Gaussian conditional densities, as suggested by Bollerslev (1987).[26] The normality of the standardized returns in Table I and Figure 1 suggests a different approach: a fat-tailed normal mixture distribution governed by the realized volatilities, consistent with the results in Theorem 2.[27] We now turn to a discussion of the distribution of the realized volatilities.

### 4.2. *Realized Volatilities*

The statistics in the top panel of Table II summarize the distribution of the realized volatilities, $v_{t,i}^{1/2}$, for each of the three exchange rates: DM/\$, ¥/\$, and

[24] Under the null hypothesis of white noise, the reported Ljung-Box statistics are distributed as chi-squared with twenty degrees of freedom. The five percent critical value is 31.4, and the one percent critical value is 37.6.

[25] Similar results obtain for the multivariate standardization $V_t^{-1/2} r_t$, where $[\cdot]^{-1/2}$ refers to the Cholesky factor of the inverse matrix, as documented in ABDL (2000a).

[26] This same observation also underlies the ad hoc multiplication factors often employed by practioners in the construction of VaR forecasts.

[27] Note that the mixed normality result in Theorem 2 does not generally follow by a standard central limit theorem except under special conditions as delineated in the theorem.

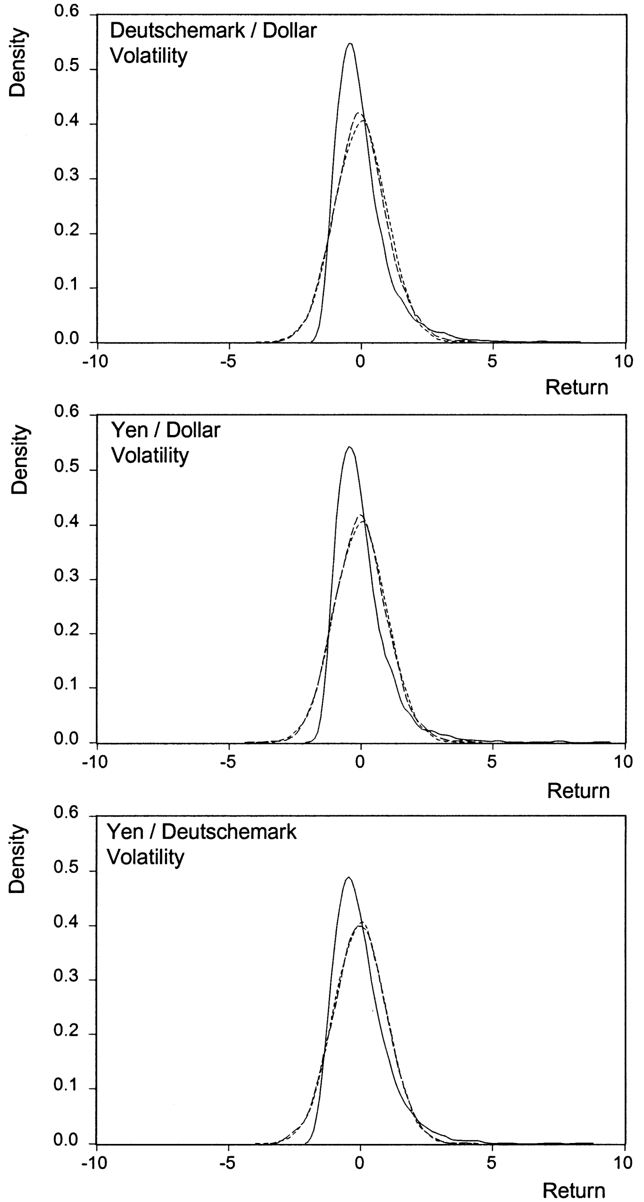


FIGURE 1.—Return distributions. The figure shows kernel estimates of the density of daily returns on the DM/$ rate, the ¥/$ rate, and an equally-weighted portfolio. The sample period extends from December 1, 1986 through December 1, 1996. The solid line is the estimated density of raw returns, standardized using its (constant) sample mean and sample standard deviation. The dashed line is the estimated density of returns standardized using its constant sample mean and time-varying realized standard deviation. The dotted line is a $N(0, 1)$ density for visual reference.

TABLE II
DAILY REALIZED VOLATILITY DISTRIBUTIONS

| | Mean | St. Dev. | Skewness | Kurtosis | $Q(20)$[d] | $d$[e] |
|---|---|---|---|---|---|---|
| *Volatility*[a, b] | | | | | | |
| DM/$ | 0.626 | 0.283 | 1.99 | 10.49 | 5249.2 | — |
| ¥/$ | 0.618 | 0.290 | 2.20 | 12.94 | 4155.1 | — |
| ¥/DM | 0.571 | 0.234 | 1.49 | 7.86 | 10074.2 | — |
| *Logarithmic Volatility*[a, c] | | | | | | |
| DM/$ | −0.554 | 0.405 | 0.251 | 3.29 | 7659.6 | 0.387 |
| ¥/$ | −0.572 | 0.419 | 0.191 | 3.44 | 5630.0 | 0.413 |
| ¥/DM | −0.637 | 0.388 | 0.058 | 3.04 | 12983.2 | 0.430 |

[a] The sample covers December 1, 1986 through December 1, 1996.
[b] The top panel refers to the distribution of realized standard deviations, $v_t^{1/2}$.
[c] The bottom panel refers to the distribution of logarithmic realized standard deviations, $1/2 \cdot \log(v_t)$.
[d] Ljung-Box test statistics for up to twentieth order serial correlation.
[e] Log-periodogram regression estimate of the fractional integration parameter, $d$, based on the $m = [T^{4/5}] = 514$ lowest-frequency periodogram ordinates. The asymptotic standard error for all of the $d$ estimates is $\pi \cdot (24 \cdot m)^{-1/2} = 0.028$.

¥/DM. All the volatilities are severely right-skewed and leptokurtic. In contrast, the skewness and kurtosis for the three logarithmic standard deviations, $y_{t,i} \equiv 1/2 \cdot \log(v_{t,i})$, shown in the bottom panel of Table II, appear remarkably Gaussian. Figure 2 confirms these impressions by presenting kernel density estimates for the logarithmic realized volatilities, which are almost indistinguishable from the normal reference densities.

The log-normality of realized volatility suggests the use of standard linear Gaussian approaches for modeling and forecasting the realized logarithmic volatilities. Moreover, combining the results for the returns in Table I, which suggest that $r_t \cdot V_t^{-1/2} \sim N(0, I)$, with the results for the realized volatilities in Table II, which suggest that $y_t \equiv (y_{t,1}, y_{t,2}, y_{t,3})' \sim N(\mu, \Omega)$, we should expect the overall return distribution (not conditioned on the realized volatility) to be well approximated by a lognormal-normal mixture.[28] Our density forecasts and VaR calculations presented below explicitly build on this insight.

Turning again to Table II, the Ljung-Box statistics indicate strong serial correlation in the realized daily volatilities, in accord with the significant Ljung-Box statistics for the squared (nonstandardized) returns in the top panel of Table I. It is noteworthy, however, that the $Q^2(20)$ statistics in Table I are orders of magnitude smaller than the $Q(20)$ statistics in Table II. This reflects the fact that, relative to the daily realized volatilities, the daily squared returns are very noisy volatility proxies, and this noise masks the strong persistence in the underlying (latent) volatility dynamics.[29]

[28] The lognormal-normal mixture distribution has previously been advocated by Clark (1973), without any of the direct empirical justification provided here.

[29] See Andersen and Bollerslev (1998), Meddahi (2002), and Andersen, Bollerslev, and Meddahi (2002) for a detailed efficiency comparison of various volatility proxies.

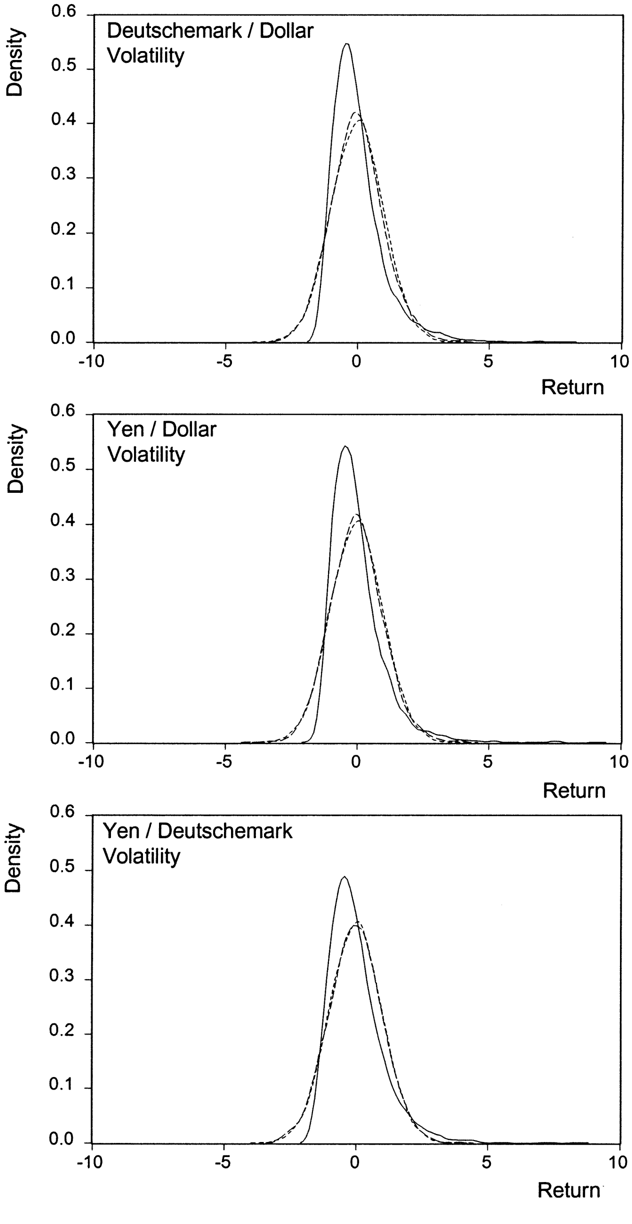


FIGURE 2.— Realized volatility distributions. The figure shows kernel estimates of the density of daily realized DM/\$, ¥/\$, and ¥/DM volatility. The sample period extends from December 1, 1986 through December 1, 1996. The solid line is the estimated density of the realized standard deviation, standardized to have zero mean and unit variance. The dashed line is the estimated density of the realized logarithmic standard deviation, standardized to have zero mean and unit variance. The dotted line is a $N(0, 1)$ density for visual reference.

Following early theoretical long-memory volatility work by Robinson (1991), many subsequent studies suggest the empirical relevance of long memory in asset return volatility, including for example Ding, Granger, and Engle (1993), Baillie, Bollerslev, and Mikkelsen (1996), and Andersen and Bollerslev (1997a). Long-memory, or fractionally-integrated, processes for the volatility also help to explain otherwise anomalous features in options data, such as volatility smiles even for long-dated options (see, for example, Renault (1997), Comte and Renault (1998), and Bollerslev and Mikkelsen (1999)). Hence in the last column of Table II we report estimates of the degree of fractional integration, obtained using the Geweke and Porter-Hudak (1983) (GPH) log-periodogram regression estimator as formally developed by Robinson (1995). The three estimates of $d$ are all significantly greater than zero and less than one half when judged by the standard error of 0.028 in the asymptotic normal distribution. Moreover, the three estimates are very close, indicative of a common degree of long-run dependence in the logarithmic volatilities. The multivariate extension of the GPH estimator developed by Robinson (1995) provides a formal framework for testing this hypothesis. On implementing Robinson's estimator we obtain a common estimate of 0.401, and the corresponding test statistic for identical values of $d$ across the three volatilities has a $p$-value of 0.510 in the asymptotic chi-square distribution with three degrees of freedom.

Figure 3 provides graphical confirmation and elaboration of the long-memory results. It displays the sample autocorrelations of the realized logarithmic volatilities out to a displacement of 70 days, or about one quarter. The slow hyperbolic autocorrelation decay symptomatic of long memory is evident, and the qualitatively identical autocorrelation values across the three volatilities supports the assertion of a common degree of fractional integration. Figure 3 also shows the sample autocorrelations of the logarithmic volatilities fractionally differenced by applying the filter $(1-L)^{0.401}$. It is evident that this single fractional differencing operator eliminates the bulk of the univariate serial dependence in each of the three realized logarithmic volatilities, although Ljung-Box portmanteau tests (not reported here) do reject the hypothesis of white noise fractionally-differenced volatilities.

It is possible that the three series are fractionally cointegrated, so that a linear combination will exhibit a degree of fractional integration less than 0.401. On heuristically testing for this by regressing each of the logarithmic volatilities on the two other logarithmic volatilities and a constant, and then estimating the degree of fractional integration in the residuals, the three estimates for $d$ are 0.356, 0.424, and 0.393, respectively, all of which are very close to the value of $d$ for the original series in Table II. Hence the realized logarithmic volatility series do not appear to be fractionally cointegrated.[30]

Meanwhile, the realized logarithmic volatility series are all strongly contemporaneously correlated. In particular, the sample correlations between $y_{t,1}$ and $y_{t,2}$

[30] Formal semiparametric frequency domain based testing procedures for fractional cointegration have recently been developed by Robinson and Marinucci (2001).

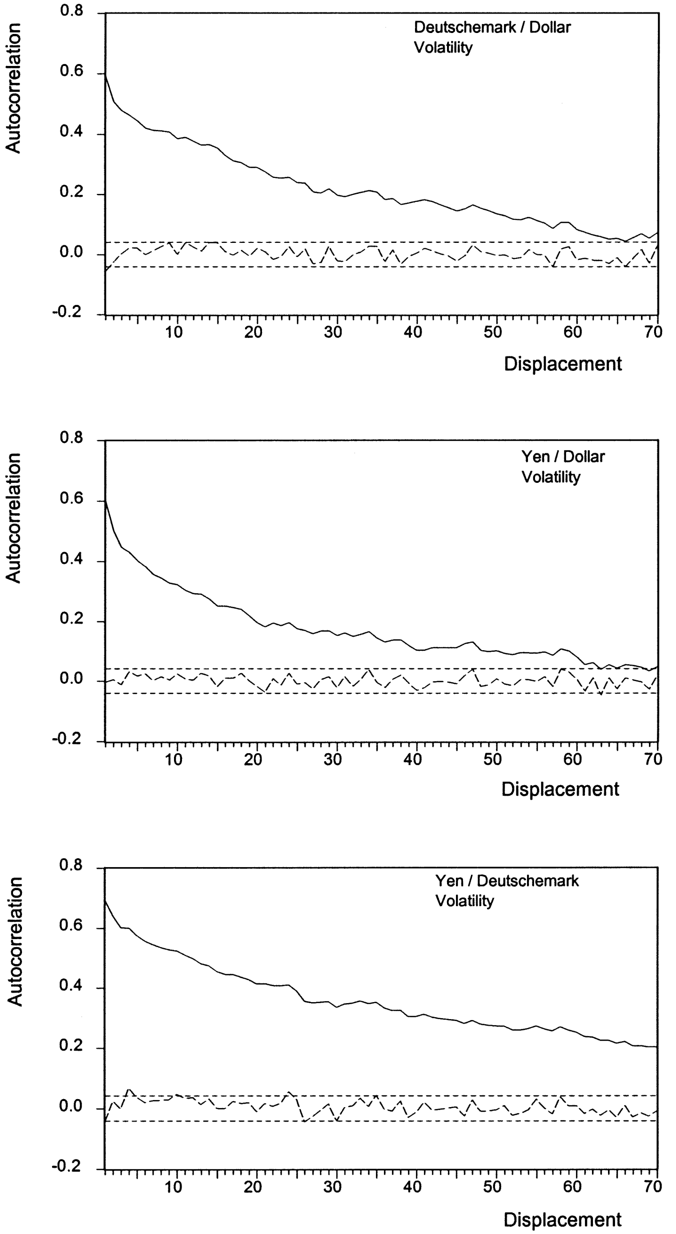


FIGURE 3.— Realized volatility autocorrelations. The figure shows the sample autocorrelation functions for daily DM/\$, ¥/\$, and ¥/DM realized volatility. The sample period extends from December 1, 1986 through December 1, 1996. The solid line gives the autocorrelation function of realized logarithmic standard deviation, while the dashed line refers to the autocorrelation function of realized logarithmic standard deviation fractionally differenced by $(1-L)^{0.401}$. The dotted lines are the Bartlett two standard error bands.

and $y_{t,3}$ are respectively 0.591 and 0.665, while the correlation between $y_{t,2}$ and $y_{t,3}$ equals 0.648. This is, of course, entirely consistent with the extant ARCH and stochastic volatility literatures. In the next section, we propose a simple multivariate model capable of accommodating both the strong dynamics and contemporaneous correlations in the realized volatilities.

## 5. A VAR FOR MODELING AND FORECASTING REALIZED VOLATILITY

The distributional features highlighted in the previous section suggest that a long-memory Gaussian VAR for the realized logarithmic volatilities should provide a good description of the volatility dynamics. We therefore consider the simple trivariate VAR (henceforth VAR-RV),

$$\Phi(L)(1-L)^d(y_t-\mu)=\epsilon_t, \tag{15}$$

where $\epsilon_t$ is a vector white noise process.[31] The model is easily estimated by applying OLS equation-by-equation. In so doing, we impose the normalization $\Phi(0)=I$, and fix the value of $d$ at the earlier-reported common estimate of 0.401. We also assume that the orders of the lag polynomials in $\Phi(L)$ are all equal to five days, or one week. This choice is somewhat arbitrary, and the model could easily be refined through a more detailed specification search explicitly allowing for zero parameter restrictions and/or different autoregressive lag lengths.[32] Additional explanatory variables, such as interest rate differentials, daily trading activity measures, lagged daily signed returns, etc., could also easily be included. However, in order to facilitate the comparison with the daily volatility models in common use, for which the mechanics of including additional explanatory variables are much more complicated and typically not entertained, we restrict our attention to the simple unrestricted VAR in equation (15).

Many of the estimated VAR coefficients (not shown) are statistically significant, and all the roots of the estimated matrix lag polynomial $\Phi(L)$ are outside the unit circle, consistent with covariance stationarity. Moreover, Ljung-Box tests for serial correlation in the VAR residuals reveal no evidence against the white noise hypothesis, indicating that the VAR has successfully accommodated all volatility dynamics not already captured by the first-stage long memory filter.

It is interesting to note that the VAR evidently does not add a great deal relative to a stacked set of univariate ARs. In particular, much of the volatility variation is explained by the univariate long-memory models (the $R^2$ values are in the neighborhood of 50%), and little of the variation of the residuals from the univariate long-memory models is explained by the VAR (the $R^2$ values

[31] Provided that all of the roots of $|\Phi(z)|=0$ lie outside the unit circle, the model is stationary, and the impulse response coefficients associated with the lag $k$ shocks are simply given by the powers in the matrix lag polynomial $\Psi(L)\equiv\Phi(L)^{-1}(1-L)^{1-d}$, say $\Psi_k$. Moreover, the cumulative impulse response coefficients, $\Psi_1+\Psi_2+\cdots+\Psi_k$, eventually dissipate at the slow hyperbolic rate of $k^{d-1}$.

[32] Both the Akaike and Schwarz information criteria select a first-order VAR. Degrees of freedom are plentiful, however, so we included a week's worth of lags to maintain conservatism.

are in the neighborhood of 2%). Effectively, the univariate one-parameter long-memory models are so successful at explaining the realized volatility dynamics that little is left for the VAR. This is also evident from the previously discussed plots of the autocorrelations in Figure 3. Nevertheless, the Ljung-Box statistics for the three univariate fractionally differenced volatility series all indicate significant serial correlation, while those for the residuals from the VAR do not. Moreover, the VAR does seem to capture some cross-rate linkages. In particular, Granger causality tests reveal some evidence of slight predictive enhancement from including lags of the logarithmic DM/$ and ¥/$ volatility in the realized logarithmic ¥/DM volatility equation.

It is natural to conjecture that the VAR-RV based realized volatility forecasts will outperform those from traditional daily ARCH and related volatility models. Our forecasts are based on explicitly-modeled long-memory dynamics, which seem to be a crucial feature of the data. Long-memory may, of course, also be incorporated in otherwise standard ARCH models, as proposed by Robinson (1991) and Baillie, Bollerslev, and Mikkelsen (1996). As such, the genuinely distinctive feature of our approach is instead that it offers simple, yet effective, incorporation of information contained in the high-frequency data. This should enable the realized volatilities and their forecasts to adapt more quickly to changes in the level of the underlying latent volatility. In the next section, we explore these conjectures in detail.

## 6. EVALUATING AND COMPARING ALTERNATIVE VOLATILITY FORECASTS

Volatility forecasts play a central role in the financial decision making process. In this section we assess the performance of the realized volatility forecasts generated from our simple VAR-RV model. For initial illustration, we plot, in Figure 4, the daily realized DM/$, ¥/$, and ¥/DM standard deviations, along with the corresponding one-day-ahead VAR-RV forecasts for the out-of-sample period, December 2, 1996, through June 30, 1999. It appears that the VAR-RV does a good job of capturing both the low-frequency and the high-frequency movements in the realized volatilities. We next proceed to a more thorough statistical evaluation of the forecasts along with a comparison to several alternative volatility forecasting procedures currently in widespread use.

### 6.1. *Forecast Evaluation*

Many methods have been proposed for modeling and forecasting financial market volatility, and we compare our VAR-RV forecasts to those of several competitors, at both one-day and ten-day horizons.[33]

[33] The multi-step forecast horizons also provide a link to the literature on temporal aggregation of ARCH and stochastic volatility models, notably Drost and Nijman (1993), Drost and Werker (1996), and Meddahi and Renault (2002). In contrast to the parametric volatility models analyzed in these studies, the realized volatility approach affords a relatively simple solution to the temporal aggregation problem.

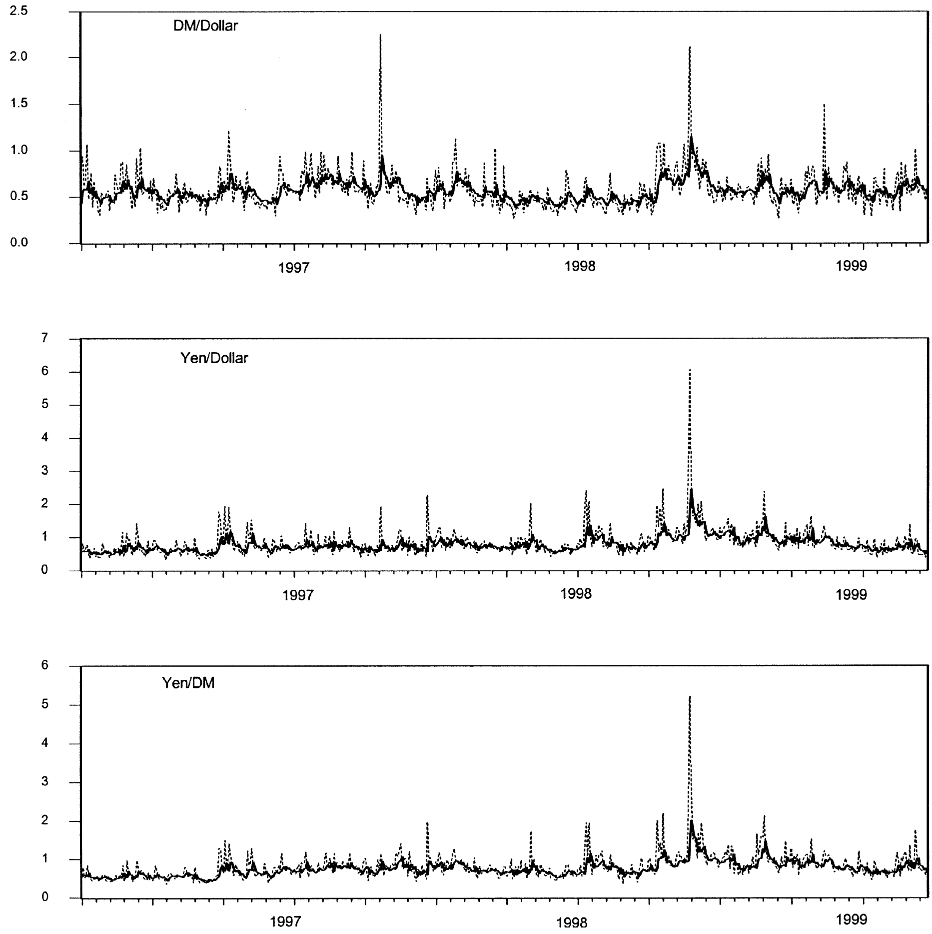


FIGURE 4.—Realized volatility and out-of-sample VAR-RV forecasts. The figure shows time series of daily realized volatility for DM/\$, ¥/\$, and ¥/DM, along with one-day-ahead VAR-RV forecasts. The plot spans the out-of-sample period from December 2, 1996 through June 30, 1999. The dotted line is realized volatility, while the solid line gives the corresponding one-day-ahead VAR-RV forecast from a long-memory vector autoregression for the daily realized volatility. See the main text for details.

First, we compare the VAR-RV forecasts to those obtained from a fifth-order VAR for the long-memory filtered daily logarithmic absolute returns (henceforth VAR-ABS). This makes for an interesting comparison, as the model structures are identical in all respects except for the volatility proxy: one uses daily realized volatility, while the other uses daily absolute returns.

Second, we compare the VAR-RV forecasts to those obtained from fifth-order univariate autoregressions for the long-memory filtered daily realized volatilities (henceforth AR-RV). This lets us assess our earlier finding from a forecasting perspective, that the multivariate interaction across the realized volatilities is minimal, in which case the forecasting performance of the VAR-RV and AR-RV models should be comparable.

Third, we compare the VAR-RV forecasts to those generated by the most widespread procedure in academic applications, the GARCH model of Engle (1982) and Bollerslev (1986), with GARCH(1, 1) constituting the leading case. As with the VAR model discussed in the previous section, we base the GARCH(1, 1) model estimates on the 2,449 daily in-sample returns from December 1, 1996, through December 1, 1996. Consistent with previous results reported in the literature, the quasi-maximum likelihood parameter estimates indicate a strong degree of volatility persistence, with the autoregressive roots for each of the three rates equal to 0.986, 0.968, and 0.990, respectively.

Fourth, we compare the VAR-RV forecasts to those of the most widespread procedure used by practitioners, J. P. Morgan's (1997) RiskMetrics. We calculate the RiskMetrics daily variances and covariances as exponentially weighted averages of the cross products of daily returns, using a smoothing factor of $\lambda = 0.94$. This corresponds to an IGARCH(1, 1) filter in which the intercept is fixed at zero and the moving average coefficient in the ARIMA(0, 1, 1) representation for the squared returns equals $-0.94$.

Fifth, we compare the VAR-RV forecasts to those of a variant of the GARCH model that incorporates long memory, the daily FIEGARCH(1, $d$, 0) model of Bollerslev and Mikkelsen (1996). The FIEGARCH model is a variant of the FIGARCH model of Baillie, Bollerslev, and Mikkelsen (1996), which, while retaining a long-memory component, employs a different volatility structure that enforces stationarity and coheres naturally with our modeling of the logarithmic volatility in the VAR-RV model.

Finally, we compare the VAR-RV volatility forecasts to those produced from a high-frequency FIEGARCH model fit to the "deseasonalized" and "filtered" half-hour returns. The deseasonalization is motivated by the fact that the intraday volatilities contain strong "seasonal" components associated with the opening and closing hours of exchanges worldwide. As noted by Andersen and Bollerslev (1997b), Martens (2001), and Martens, Chang, and Taylor (2002) among others, these intraday patterns severely corrupt the estimation of traditional volatility models based on the raw high-frequency returns. For simplicity, we estimate the intraday patterns by simply averaging the individual squared returns in the various intra-day intervals, resulting in the "seasonal" factors

$$s_i^2 = \frac{1}{T}\sum_{t=1}^{T} r_{it}^2 \qquad (i = 1, \ldots, 48),$$

where $r_{it}$ denotes the return in the $i$th interval on day $t$, based upon which we construct the "seasonally adjusted" high frequency returns,

$$\tilde{r}_{it} = \frac{r_{it}}{s_i} \qquad (i = 1, \ldots, 48; t = 1, \ldots, T).$$

Most high-frequency asset returns also display significant own serial correlation. These dependencies are generally directly attributable to various market microstructure frictions (see, for example, the discussion in Andersen and

Bollerslev (1997b) and Bai, Russell, and Tiao (2000)). To reduce the impact of this "spurious" serial correlation we therefore also apply a simple first-order autoregressive "filter" to the high-frequency returns before the estimation of the FIEGARCH model.

No universally acceptable loss function exists for the ex-post evaluation and comparison of nonlinear model forecasts, and in the context of volatility modeling, several statistical procedures have been used for assessing the quality of competing forecasts (see, for example, the discussion in Andersen, Bollerslev, and Lange (1999) and Christoffersen and Diebold (1999)). Following Andersen and Bollerslev (1998), and in the tradition of Mincer and Zarnowitz (1969) and Chong and Hendry (1986), we evaluate the alternative volatility forecasts by projecting the realized volatilities on a constant and the various model forecasts.

For the one-day-ahead in-sample and out-of-sample forecasts reported in Tables III.A and III.B, the forecast evaluation regressions take the form[34]

$$\left(\{v_{t+1}\}_i\right)^{1/2} = b_0 + b_1 \cdot \left(\{v_{t+1|t,VAR\text{-}RV}\}_i\right)^{1/2} + b_2 \cdot \left(\{v_{t+1|t,Model}\}_i\right)^{1/2} + u_{t+1,i}. \tag{16}$$

The results are striking. For the in-sample regressions including just one volatility forecast, the regression $R^2$ is always the highest for the VAR-RV model, and for almost none of the VAR-RV forecasts can we reject the hypothesis that $b_0 = 0$ and $b_1 = 1$ using the corresponding $t$ tests.[35] In contrast, we reject the hypothesis that $b_0 = 0$ and/or $b_2 = 1$ for most of the VAR-ABS, AR-RV, GARCH, RiskMetrics, daily FIEGARCH, and intraday FIEGARCH in-sample forecasts. Moreover, on including both the VAR-RV and an alternate forecast in the same regression, the coefficient estimates for $b_1$ and $b_2$ are generally close to unity and near zero, respectively. Finally, inclusion of the alternative forecasts improves the $R^2$'s very little relative to those based solely on the VAR-RV forecasts.[36]

Things are only slightly less compelling for the one-day-ahead out-of-sample forecasts, shown in Table III.B. Although (in the single forecast regressions) the VAR-RV model has a higher $R^2$ than most of the alternative forecasting methods,

[34] Following standard practice in the literature, we focus on forecasts for the standard deviation, $v_t^{1/2}$. Of course, the transformed forecasts are not formally unbiased, so we also experimented with a first order Taylor series expansion of the square root and exponential functions to adjust for this bias. The resulting regression estimates and $R^2$'s were almost identical to the ones reported in Table III. Similarly, the regressions for the realized variances, $v_t$, and logarithmic standard deviations, $y_t \equiv 1/2 \cdot \log(v_t)$, produced very similar results to the ones reported here. Detailed tables appear in the supplemental Appendix to this paper, available at *www.ssc.upenn.edu/~diebold*.

[35] The reported (conventional) heteroskedasticity robust standard errors from the forecast evaluation regressions will generally be downward biased as they fail to incorporate the parameter estimation error uncertainty in the different volatility forecasting models; see West and McCracken (1998).

[36] These results are consistent with Engle (2000), who reports that the inclusion of the lagged daily realized variance in the conditional variance equation of a GARCH(1, 1) model for the daily DM/$ returns analyzed here renders the coefficient associated with the lagged daily squared returns insignificant. The HARCH model in Müller et al. (1997) also highlights the importance of intraday returns in modeling daily volatility. In contrast, the results in Taylor and Xu (1997), based on a limited one-year sample, suggest that the lagged daily realized variance offers little incremental explanatory power over a univariate GARCH(1, 1) model.

TABLE III.A

FORECAST EVALUATION—IN-SAMPLE, ONE-DAY-AHEAD

| | $b_0$[a] | | $b_1$[a] | | $b_2$[a] | | $R^2$ |
|---|---|---|---|---|---|---|---|
| *DM/$*[b] | | | | | | | |
| VAR-RV | −0.008 | (0.020) | 1.065 | (0.035) | — | | 0.355 |
| VAR-ABS | 0.229 | (0.019) | — | | 1.206 | (0.061) | 0.161 |
| AR-RV | −0.012 | (0.020) | — | | 1.072 | (0.035) | 0.351 |
| Daily GARCH | −0.071 | (0.026) | — | | 1.012 | (0.040) | 0.265 |
| Daily RiskMetrics | 0.109 | (0.019) | — | | 0.766 | (0.030) | 0.265 |
| Daily FIEGARCH | −0.042 | (0.025) | — | | 0.961 | (0.038) | 0.256 |
| Intraday FIEGARCH deseason/filter | −0.172 | (0.029) | — | | 1.234 | (0.045) | 0.324 |
| VAR-RV + VAR-ABS | −0.020 | (0.020) | 1.022 | (0.047) | 0.112 | (0.073) | 0.356 |
| VAR-RV + AR-RV | −0.007 | (0.020) | 1.197 | (0.310) | −0.134 | (0.314) | 0.355 |
| VAR-RV + Daily GARCH | −0.048 | (0.024) | 0.944 | (0.060) | 0.162 | (0.062) | 0.357 |
| VAR-RV + Daily RiskMetrics | −0.019 | (0.020) | 0.943 | (0.060) | 0.124 | (0.047) | 0.357 |
| VAR-RV + Daily FIEGARCH | −0.045 | (0.023) | 0.952 | (0.062) | 0.150 | (0.061) | 0.357 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.076 | (0.032) | 0.811 | (0.074) | 0.339 | (0.101) | 0.359 |
| *¥/$*[b] | | | | | | | |
| VAR-RV | −0.008 | (0.021) | 1.064 | (0.038) | — | | 0.380 |
| VAR-ABS | 0.258 | (0.018) | — | | 1.179 | (0.065) | 0.157 |
| AR-RV | −0.007 | (0.021) | — | | 1.064 | (0.038) | 0.374 |
| Daily GARCH | −0.094 | (0.032) | — | | 1.050 | (0.049) | 0.294 |
| Daily RiskMetrics | 0.148 | (0.020) | — | | 0.714 | (0.033) | 0.263 |
| Daily FIEGARCH | −0.087 | (0.027) | — | | 1.054 | (0.043) | 0.315 |
| Intraday FIEGARCH deseason/filter | −0.242 | (0.029) | — | | 1.347 | (0.047) | 0.373 |
| VAR-RV + VAR-ABS | −0.020 | (0.021) | 1.021 | (0.046) | 0.122 | (0.066) | 0.381 |
| VAR-RV + AR-RV | −0.006 | (0.021) | 1.337 | (0.360) | −0.277 | (0.360) | 0.380 |
| VAR-RV + Daily GARCH | −0.072 | (0.030) | 0.883 | (0.060) | 0.252 | (0.076) | 0.386 |
| VAR-RV + Daily RiskMetrics | −0.018 | (0.021) | 0.953 | (0.054) | 0.114 | (0.045) | 0.383 |
| VAR-RV + Daily FIEGARCH | −0.089 | (0.026) | 0.808 | (0.050) | 0.346 | (0.057) | 0.392 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.139 | (0.036) | 0.615 | (0.101) | 0.618 | (0.133) | 0.391 |
| *¥/DM*[b] | | | | | | | |
| VAR-RV | −0.023 | (0.013) | 1.080 | (0.025) | — | | 0.509 |
| VAR-ABS | 0.189 | (0.014) | — | | 1.336 | (0.051) | 0.303 |
| AR-RV | −0.020 | (0.013) | — | | 1.074 | (0.025) | 0.499 |
| Daily GARCH | 0.102 | (0.015) | — | | 0.821 | (0.027) | 0.394 |
| Daily RiskMetrics | 0.156 | (0.012) | — | | 0.742 | (0.023) | 0.374 |
| Daily FIEGARCH | 0.034 | (0.016) | — | | 0.941 | (0.029) | 0.404 |
| Intraday FIEGARCH deseason/filter | −0.178 | (0.018) | — | | 1.291 | (0.033) | 0.471 |
| VAR-RV + VAR-ABS | −0.029 | (0.013) | 1.012 | (0.037) | 0.148 | (0.060) | 0.510 |
| VAR-RV + AR-RV | −0.022 | (0.013) | 1.207 | (0.206) | −0.129 | (0.203) | 0.509 |
| VAR-RV + Daily GARCH | −0.030 | (0.013) | 0.938 | (0.047) | 0.148 | (0.043) | 0.513 |
| VAR-RV + Daily RiskMetrics | −0.023 | (0.013) | 0.986 | (0.046) | 0.091 | (0.037) | 0.511 |
| VAR-RV + Daily FIEGARCH | −0.042 | (0.014) | 0.918 | (0.048) | 0.189 | (0.047) | 0.514 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.075 | (0.021) | 0.851 | (0.072) | 0.305 | (0.094) | 0.512 |

[a] OLS estimates from regressions of realized volatility on a constant and forecasts from different models with heteroskedasticity robust standard errors in parentheses.

[b] The forecast evaluation period is December 1, 1987 through December 1, 1996, for a total of 2,223 daily observations. All model parameter estimates are based on data from December 1, 1986 through December 1, 1996.

TABLE III.B

FORECAST EVALUATION—OUT-OF-SAMPLE, ONE-DAY-AHEAD

| | $b_0$[a] | | $b_1$[a] | | $b_2$[a] | | $R^2$ |
|---|---|---|---|---|---|---|---|
| *DM/$*[b] | | | | | | | |
| VAR-RV | 0.021 | (0.049) | 0.987 | (0.092) | — | | 0.249 |
| VAR-ABS | 0.439 | (0.028) | — | | 0.450 | (0.089) | 0.028 |
| AR-RV | 0.007 | (0.049) | — | | 1.022 | (0.093) | 0.246 |
| Daily GARCH | 0.051 | (0.063) | — | | 0.854 | (0.105) | 0.096 |
| Daily RiskMetrics | 0.219 | (0.042) | — | | 0.618 | (0.075) | 0.097 |
| Daily FIEGARCH | 0.305 | (0.052) | — | | 0.436 | (0.083) | 0.037 |
| Intraday FIEGARCH deseason/filter | −0.069 | (0.060) | — | | 1.012 | (0.099) | 0.266 |
| VAR-RV + VAR-ABS | 0.035 | (0.046) | 1.018 | (0.107) | −0.106 | (0.103) | 0.250 |
| VAR-RV + AR-RV | 0.015 | (0.047) | 0.764 | (0.532) | 0.235 | (0.525) | 0.249 |
| VAR-RV + Daily GARCH | 0.015 | (0.060) | 0.980 | (0.134) | 0.016 | (0.156) | 0.249 |
| VAR-RV + Daily RiskMetrics | 0.017 | (0.046) | 0.979 | (0.133) | 0.014 | (0.112) | 0.249 |
| VAR-RV + Daily FIEGARCH | 0.088 | (0.050) | 1.067 | (0.119) | −0.181 | (0.111) | 0.254 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.073 | (0.059) | 0.403 | (0.185) | 0.662 | (0.197) | 0.275 |
| *¥/$*[b] | | | | | | | |
| VAR-RV | −0.006 | (0.110) | 1.085 | (0.151) | — | | 0.329 |
| VAR-ABS | 0.349 | (0.086) | — | | 1.256 | (0.241) | 0.115 |
| AR-RV | −0.054 | (0.114) | — | | 1.157 | (0.158) | 0.338 |
| Daily GARCH | −0.002 | (0.147) | — | | 1.020 | (0.187) | 0.297 |
| Daily RiskMetrics | 0.164 | (0.108) | — | | 0.767 | (0.131) | 0.266 |
| Daily FIEGARCH | −0.289 | (0.193) | — | | 1.336 | (0.236) | 0.373 |
| Intraday FIEGARCH deseason/filter | −0.394 | (0.189) | — | | 1.647 | (0.263) | 0.380 |
| VAR-RV + VAR-ABS | −0.038 | (0.111) | 1.037 | (0.159) | 0.179 | (0.132) | 0.331 |
| VAR-RV + AR-RV | −0.074 | (0.112) | −0.692 | (0.566) | 1.881 | (0.576) | 0.340 |
| VAR-RV + Daily GARCH | −0.081 | (0.144) | 0.733 | (0.121) | 0.424 | (0.247) | 0.346 |
| VAR-RV + Daily RiskMetrics | −0.022 | (0.115) | 0.859 | (0.113) | 0.219 | (0.123) | 0.336 |
| VAR-RV + Daily FIEGARCH | −0.269 | (0.213) | 0.364 | (0.245) | 0.977 | (0.469) | 0.383 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.389 | (0.240) | 0.028 | (0.288) | 1.610 | (0.614) | 0.380 |
| *¥/DM*[b] | | | | | | | |
| VAR-RV | −0.047 | (0.101) | 1.146 | (0.143) | — | | 0.355 |
| VAR-ABS | 0.405 | (0.062) | — | | 1.063 | (0.175) | 0.119 |
| AR-RV | −0.037 | (0.098) | — | | 1.129 | (0.140) | 0.344 |
| Daily GARCH | 0.243 | (0.092) | — | | 0.692 | (0.119) | 0.300 |
| Daily RiskMetrics | 0.248 | (0.084) | — | | 0.668 | (0.107) | 0.286 |
| Daily FIEGARCH | 0.101 | (0.105) | — | | 0.918 | (0.144) | 0.263 |
| Intraday FIEGARCH deseason/filter | −0.231 | (0.150) | — | | 1.455 | (0.217) | 0.404 |
| VAR-RV + VAR-ABS | −0.039 | (0.097) | 1.176 | (0.167) | −0.080 | (0.146) | 0.355 |
| VAR-RV + AR-RV | −0.043 | (0.096) | 1.891 | (0.956) | −0.750 | (0.862) | 0.356 |
| VAR-RV + Daily GARCH | −0.020 | (0.086) | 0.873 | (0.137) | 0.215 | (0.159) | 0.364 |
| VAR-RV + Daily RiskMetrics | −0.029 | (0.093) | 0.941 | (0.130) | 0.160 | (0.117) | 0.360 |
| VAR-RV + Daily FIEGARCH | −0.061 | (0.109) | 1.033 | (0.132) | 0.127 | (0.128) | 0.356 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.226 | (0.157) | 0.229 | (0.228) | 1.210 | (0.509) | 0.407 |

[a] OLS estimates from regressions of realized volatility on a constant and forecasts from different models with heteroskedasticity robust standard errors in parentheses.

[b] The forecast evaluation period is December 2, 1996 through June 30, 1999, for a total of 596 daily observations. All model parameter estimates are based on data from December 1, 1986 through December 1, 1996.

it is edged out by the intraday FIEGARCH predictions, which exhibit somewhat higher $R^2$'s.[37] However, while we cannot reject the hypothesis that $b_0 = 0$ and $b_1 = 1$ for the VAR-RV forecasts, for some of the scenarios the hypothesis that $b_0 = 0$ and $b_2 = 1$ is rejected for the intraday FIEGARCH model based on the conventional standard errors and traditional levels of significance.

Turning to the ten-day-ahead forecasts, the VAR-RV results are still good. The evaluation regressions are

$$\left(\left\{\sum_{j=1,\ldots,10} v_{t+j}\right\}_i\right)^{1/2} = b_0 + b_1 \cdot \left(\left\{\sum_{j=1,\ldots,10} v_{t+j|t,VAR\text{-}RV}\right\}_i\right)^{1/2} + b_2 \cdot \left(\left\{\sum_{j=1,\ldots,10} v_{t+j|t,Model}\right\}_i\right)^{1/2} + u_{t+10,i}. \tag{17}$$

Most of the in-sample and out-of-sample results in Tables III.C and III.D, respectively, favor the VAR-RV forecasts.[38] In almost every (single-regression) case the VAR-RV forecast exhibits a higher $R^2$ than the alternative methods. As with the one-day-ahead regressions discussed above, the estimates for $b_1$ are generally not significantly different from unity, while very few of the estimates for $b_0$ and $b_2$ in the multiple regressions including both the VAR-RV forecasts and the other forecasts are statistically significantly different from zero. These results are especially noteworthy insofar as several previous studies have found it difficult to outperform simple daily GARCH(1, 1) based exchange rate volatility forecasts using more complicated univariate or multivariate models (for example, Hansen and Lunde (2001) and Sheedy (1998)), or ARCH models estimated directly from high-frequency data (for example, Beltratti and Morana (1999)).

In closing this subsection, we note that the good forecasting performance of the VAR-RV model appears robust to a much larger set of models and scenarios than those reported here. The supplemental Appendix contains extensive results for many other volatility models, as well as scenarios involving alternative volatility measures and exclusion of the "once-in-a-generation" anomalous yen movement on October 7–9, 1998 (see, Cai, Cheung, Lee, and Melvin (2000)). Overall, across the many scenarios examined, the VAR-RV forecasts are usually the most accurate and consistently among the very best. Of course, in a few instances, one or another competitor may perform comparatively well, but those exceptions simply prove the rule: the VAR-RV based forecasts are *systematically* best or nearly best in all cases.

[37] In a related context, Bollerslev and Wright (2001) find the forecasts from a simple univariate AR model for the realized DM/$ volatility to be only slightly inferior to the forecasts from a much more complicated frequency domain procedure applied directly to the underlying high-frequency squared returns.

[38] To account for the overlap in the multi-day forecasts, we use robust standard errors, calculated using an unweighted covariance matrix estimator allowing for up to ninth order serial correlation in $u_{t+10,i}$.

TABLE III.C

FORECAST EVALUATION—IN-SAMPLE, TEN-DAYS-AHEAD

| | $b_0$[a] | | $b_1$[a] | | $b_2$[a] | | $R^2$ |
|---|---|---|---|---|---|---|---|
| *DM/$*[b] | | | | | | | |
| VAR-RV | −0.038 | (0.136) | 1.131 | (0.076) | — | | 0.438 |
| VAR-ABS | 0.648 | (0.147) | — | | 1.372 | (0.149) | 0.290 |
| AR-RV | −0.070 | (0.141) | — | | 1.149 | (0.079) | 0.436 |
| Daily GARCH | 0.027 | (0.190) | — | | 0.934 | (0.091) | 0.380 |
| Daily RiskMetrics | 0.708 | (0.134) | — | | 0.637 | (0.066) | 0.343 |
| Daily FIEGARCH | 0.185 | (0.165) | — | | 0.857 | (0.079) | 0.379 |
| Intraday FIEGARCH deseason/filter | −5.019 | (0.599) | — | | 3.306 | (0.281) | 0.313 |
| VAR-RV + VAR-ABS | −0.052 | (0.138) | 1.034 | (0.109) | 0.187 | (0.159) | 0.440 |
| VAR-RV + AR-RV | −0.045 | (0.136) | 0.935 | (0.708) | 0.199 | (0.709) | 0.438 |
| VAR-RV + Daily GARCH | −0.147 | (0.151) | 0.860 | (0.144) | 0.280 | (0.132) | 0.447 |
| VAR-RV + Daily RiskMetrics | −0.015 | (0.141) | 1.053 | (0.153) | 0.057 | (0.098) | 0.439 |
| VAR-RV + Daily FIEGARCH | −0.085 | (0.141) | 0.891 | (0.148) | 0.224 | (0.124) | 0.444 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.722 | (0.562) | 1.037 | (0.105) | 0.401 | (0.321) | 0.440 |
| *¥/$*[b] | | | | | | | |
| VAR-RV | 0.008 | (0.153) | 1.108 | (0.087) | — | | 0.399 |
| VAR-ABS | 0.821 | (0.126) | — | | 1.271 | (0.140) | 0.232 |
| AR-RV | 0.002 | (0.157) | — | | 1.112 | (0.090) | 0.396 |
| Daily GARCH | −0.050 | (0.209) | — | | 0.969 | (0.101) | 0.367 |
| Daily RiskMetrics | 0.862 | (0.135) | — | | 0.566 | (0.069) | 0.303 |
| Daily FIEGARCH | 0.335 | (0.187) | — | | 0.808 | (0.093) | 0.280 |
| Intraday FIEGARCH deseason/filter | −5.405 | (0.603) | — | | 3.501 | (0.288) | 0.317 |
| VAR-RV + VAR-ABS | −0.004 | (0.154) | 1.038 | (0.113) | 0.146 | (0.146) | 0.400 |
| VAR-RV + AR-RV | 0.010 | (0.156) | 1.240 | (0.742) | −0.133 | (0.765) | 0.399 |
| VAR-RV + Daily GARCH | −0.200 | (0.175) | 0.724 | (0.141) | 0.423 | (0.140) | 0.421 |
| VAR-RV + Daily RiskMetrics | 0.043 | (0.154) | 1.018 | (0.145) | 0.063 | (0.091) | 0.400 |
| VAR-RV + Daily FIEGARCH | −0.001 | (0.160) | 1.086 | (0.140) | 0.024 | (0.117) | 0.399 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −1.236 | (0.564) | 0.935 | (0.136) | 0.734 | (0.350) | 0.403 |
| *¥/DM*[b] | | | | | | | |
| VAR-RV | −0.177 | (0.101) | 1.185 | (0.062) | — | | 0.591 |
| VAR-ABS | 0.523 | (0.101) | — | | 1.486 | (0.111) | 0.440 |
| AR-RV | −0.175 | (0.102) | — | | 1.183 | (0.062) | 0.589 |
| Daily GARCH | 0.427 | (0.103) | — | | 0.789 | (0.057) | 0.520 |
| Daily RiskMetrics | 0.686 | (0.089) | — | | 0.667 | (0.049) | 0.469 |
| Daily FIEGARCH | 0.284 | (0.099) | — | | 0.876 | (0.056) | 0.522 |
| Intraday FIEGARCH deseason/filter | −5.387 | (0.451) | — | | 3.777 | (0.235) | 0.460 |
| VAR-RV + VAR-ABS | −0.163 | (0.104) | 1.054 | (0.114) | 0.231 | (0.145) | 0.595 |
| VAR-RV + AR-RV | −0.176 | (0.101) | 1.454 | (0.748) | −0.270 | (0.735) | 0.591 |
| VAR-RV + Daily GARCH | −0.114 | (0.104) | 0.873 | (0.121) | 0.260 | (0.079) | 0.606 |
| VAR-RV + Daily RiskMetrics | −0.144 | (0.124) | 1.109 | (0.146) | 0.055 | (0.084) | 0.592 |
| VAR-RV + Daily FIEGARCH | −0.151 | (0.100) | 0.890 | (0.128) | 0.266 | (0.092) | 0.603 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.764 | (0.415) | 1.092 | (0.086) | 0.388 | (0.265) | 0.592 |

[a] OLS estimates from regressions of realized volatility on a constant and forecasts from different models with heteroskedasticity robust standard errors, calculated using an unweighted covariance matrix estimator allowing for up to ninth order serial correlation in the error term, in parentheses.

[b] The forecast evaluation period is December 1, 1987 through December 1, 1996, for a total of 2,223 daily observations. All model parameter estimates are based on data from December 1, 1986 through December 1, 1996.

TABLE III.D

FORECAST EVALUATION—OUT-OF-SAMPLE, TEN-DAYS-AHEAD

| | $b_0$[a] | $b_1$[a] | $b_2$[a] | $R^2$ |
|---|---|---|---|---|
| *DM/\$*[b] | | | | |
| VAR-RV | 0.417 (0.329) | 0.827 (0.194) | — | 0.197 |
| VAR-ABS | 1.652 (0.243) | — | 0.233 (0.224) | 0.011 |
| AR-RV | 0.375 (0.358) | — | 0.857 (0.211) | 0.195 |
| Daily GARCH | 0.780 (0.397) | — | 0.559 (0.205) | 0.082 |
| Daily RiskMetrics | 1.304 (0.265) | — | 0.312 (0.144) | 0.056 |
| Daily FIEGARCH | 1.537 (0.350) | — | 0.172 (0.166) | 0.013 |
| Intraday FIEGARCH deseason/filter | −2.069 (0.971) | — | 1.848 (0.459) | 0.152 |
| VAR-RV + VAR-ABS | 0.470 (0.302) | 1.083 (0.256) | −0.527 (0.291) | 0.234 |
| VAR-RV + AR-RV | 0.412 (0.363) | 0.769 (0.808) | 0.062 (0.904) | 0.197 |
| VAR-RV + Daily GARCH | 0.443 (0.340) | 0.847 (0.292) | −0.031 (0.266) | 0.198 |
| VAR-RV + Daily RiskMetrics | 0.441 (0.318) | 0.966 (0.274) | −0.147 (0.173) | 0.204 |
| VAR-RV + Daily FIEGARCH | 0.657 (0.306) | 1.140 (0.263) | −0.402 (0.212) | 0.241 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.614 (0.641) | 0.642 (0.217) | 0.636 (0.345) | 0.206 |
| *¥/\$*[b] | | | | |
| VAR-RV | 0.188 (0.530) | 1.114 (0.250) | — | 0.278 |
| VAR-ABS | 0.957 (0.572) | — | 1.514 (0.534) | 0.183 |
| AR-RV | 0.090 (0.571) | — | 1.164 (0.270) | 0.279 |
| Daily GARCH | 0.706 (0.403) | — | 0.810 (0.175) | 0.263 |
| Daily RiskMetrics | 1.329 (0.322) | — | 0.526 (0.133) | 0.211 |
| Daily FIEGARCH | 0.375 (0.574) | — | 0.909 (0.236) | 0.218 |
| Intraday FIEGARCH deseason/filter | −5.223 (1.285) | — | 3.621 (0.606) | 0.228 |
| VAR-RV + VAR-ABS | 0.069 (0.569) | 0.929 (0.199) | 0.456 (0.301) | 0.287 |
| VAR-RV + AR-RV | 0.071 (0.637) | −0.250 (1.082) | 1.423 (1.343) | 0.280 |
| VAR-RV + Daily GARCH | 0.195 (0.502) | 0.686 (0.253) | 0.387 (0.196) | 0.297 |
| VAR-RV + Daily RiskMetrics | 0.203 (0.550) | 1.091 (0.322) | 0.015 (0.104) | 0.278 |
| VAR-RV + Daily FIEGARCH | 0.207 (0.551) | 1.160 (0.320) | −0.047 (0.237) | 0.278 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −1.466 (0.824) | 0.885 (0.261) | 0.989 (0.446) | 0.283 |
| *¥/DM*[b] | | | | |
| VAR-RV | 0.155 (0.438) | 1.119 (0.214) | — | 0.293 |
| VAR-ABS | 1.023 (0.527) | — | 1.398 (0.477) | 0.215 |
| AR-RV | 0.178 (0.420) | — | 1.106 (0.206) | 0.293 |
| Daily GARCH | 1.173 (0.324) | — | 0.578 (0.135) | 0.328 |
| Daily RiskMetrics | 1.258 (0.329) | — | 0.529 (0.136) | 0.299 |
| Daily FIEGARCH | 0.927 (0.368) | — | 0.714 (0.157) | 0.244 |
| Intraday FIEGARCH deseason/filter | −3.196 (0.737) | — | 2.887 (0.372) | 0.225 |
| VAR-RV + VAR-ABS | 0.143 (0.448) | 0.910 (0.206) | 0.408 (0.577) | 0.301 |
| VAR-RV + AR-RV | 0.159 (0.473) | 0.845 (2.888) | 0.271 (2.695) | 0.293 |
| VAR-RV + Daily GARCH | 0.667 (0.367) | 0.426 (0.353) | 0.404 (0.252) | 0.341 |
| VAR-RV + Daily RiskMetrics | 0.614 (0.289) | 0.557 (0.287) | 0.301 (0.227) | 0.316 |
| VAR-RV + Daily FIEGARCH | 0.205 (0.398) | 0.929 (0.279) | 0.154 (0.239) | 0.296 |
| VAR-RV + Intraday FIEGARCH deseason/filter | −0.918 (0.558) | 0.917 (0.216) | 0.751 (0.307) | 0.299 |

[a] OLS estimates from regressions of realized volatility on a constant and forecasts from different models with heteroskedasticity robust standard errors, calculated using an unweighted covariance matrix estimator allowing for up to ninth order serial correlation in the error term, in parentheses.

[b] The forecast evaluation period is December 2, 1996 through June 30, 1999, for a total of 596 daily observations. All model parameter estimates are based on data from December 1, 1986 through December 1, 1996.

### 6.2. *On the Superiority of VAR-RV Forecasts*

Why does the VAR-RV produce superior forecasts? We have identified the quadratic variation and its empirical counterpart, the realized volatility, as the key objects of interest for volatility measurement, and we consequently assess our various volatility forecast relative to this measure. It is perhaps not surprising that models built directly for the realized volatility produce forecasts superior to those obtained from less direct methods, a conjecture consistent with the literature on forecasting under the relevant loss function, such as Christoffersen and Diebold (1997).

There is a more direct reason for the superior performance of the VAR-RV forecasts, however. The essence of forecasting is quantification of the mapping from the past and present into the future. Hence, quite generally, superior estimates of present conditions translate into superior forecasts of the future. Realized volatility excels in this dimension: it provides a relatively precise and quickly-adapting estimate of current volatility, because it exploits valuable intraday information. Standard models based on daily data such as GARCH and RiskMetrics rely on long and slowly decaying weighted moving averages of past squared returns and therefore adapt only gradually to volatility movements. Suppose, for example, that the true volatility has been low for many days, $t = 1, \ldots, T-1$, so that both realized and GARCH volatilities are presently low as well. Now suppose that the true volatility increases sharply on day $T$ and that the effect is highly persistent as is typical. Realized volatility for day $T$, which makes effective use of the day-$T$ information, will increase sharply as well, as is appropriate. GARCH or RiskMetrics volatility, in contrast, will not change at all on day $T$, as they depend only on squared returns from days $T-1, T-2, \ldots,$ and they will increase only gradually on subsequent days, as they approximate volatility via a long and slowly decaying exponentially weighted moving average.

Figure 5 confirms the above assertions graphically. We display the realized standard deviations for DM/$ returns, ¥/$ returns, and ¥/DM returns, along with the corresponding one-day-ahead GARCH forecasts for the out-of-sample period, December 2, 1996, through June 30, 1999. The GARCH forecasts appear to track the low-frequency variation adequately, matching the broad temporal movements in the volatilities, but they track much less well at higher frequencies. Note the striking contrast with Figure 4 which, as discussed earlier, reveals a close coherence between the daily realized volatilities and the VAR-RV forecasts at high as well as low frequencies.

We provide a more detailed illustration of the comparative superiority of the VAR-RV forecasts in Figure 6, which depicts four illustrative DM/$ episodes of thirty-five days each.[39] First, for days one through twenty-five (the non-shaded region) we show the daily realized volatility together with the one-day-ahead forecasts made on the previous day using the VAR-RV and GARCH models. The

[39] The actual start dates for the forecasts in each of the four panels are 11/16/1988, 3/6/1989, 5/13/1991, and 8/28/1998, respectively.

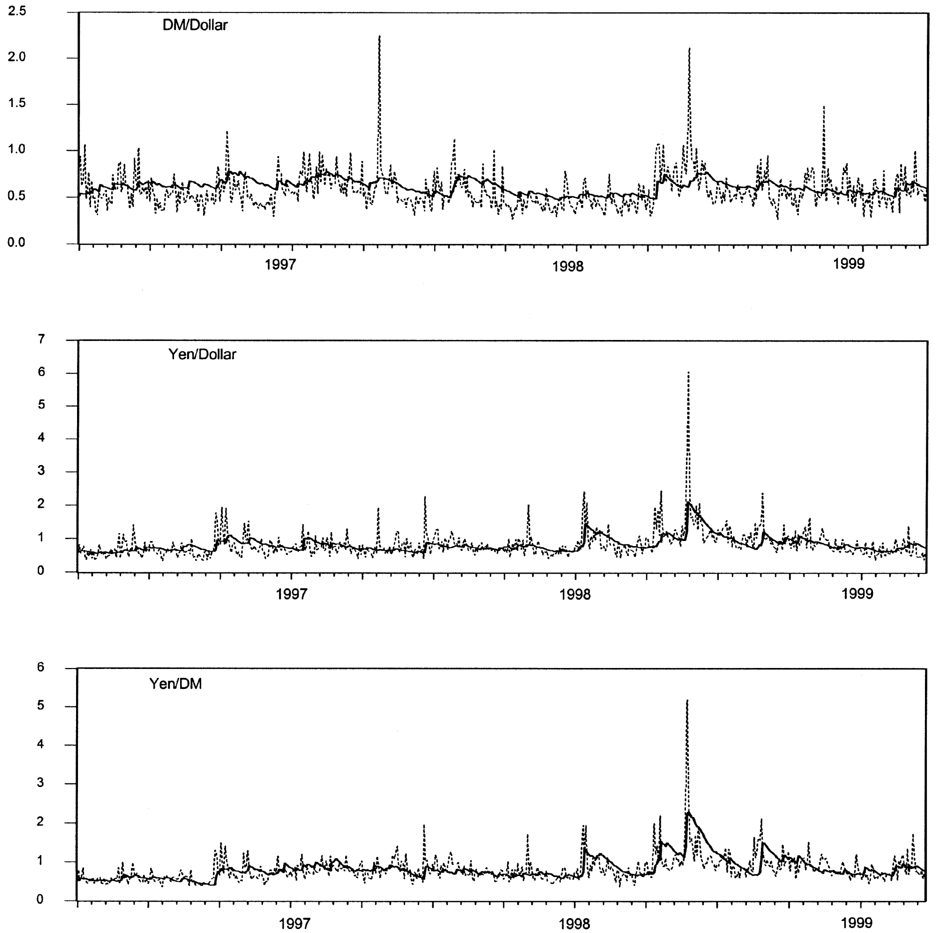


FIGURE 5.—Realized volatility and out-of-sample GARCH forecasts. The figure shows time series of daily realized volatility for DM/\$, ¥/\$, and ¥/DM, along with one-day-ahead GARCH(1, 1) forecasts. The plot spans the out-of-sample period, running from December 2, 1996 through June 30, 1999. The dotted line is realized volatility, while the solid line gives the corresponding one-day-ahead GARCH forecast. See the main text for details.

accuracy of the VAR-RV forecasts is striking, as is the *in*accuracy of the GARCH forecasts, and their inability to adapt to high-frequency movements. Second, for days twenty-six through thirty-five (the shaded region), we continue to display the daily realized volatility, but we show one- through ten-day-ahead VAR-RV and GARCH forecasts based on information available on day twenty-five. Hence the forecasts for day twenty-six are one-day-ahead, the forecasts for day twenty-seven are two-day-ahead, and so on. Examination of these multi-step trajectories makes clear the forecasting benefit of having a superior estimate of current volatility: in each case the VAR-RV forecasts "take off" from a good estimate

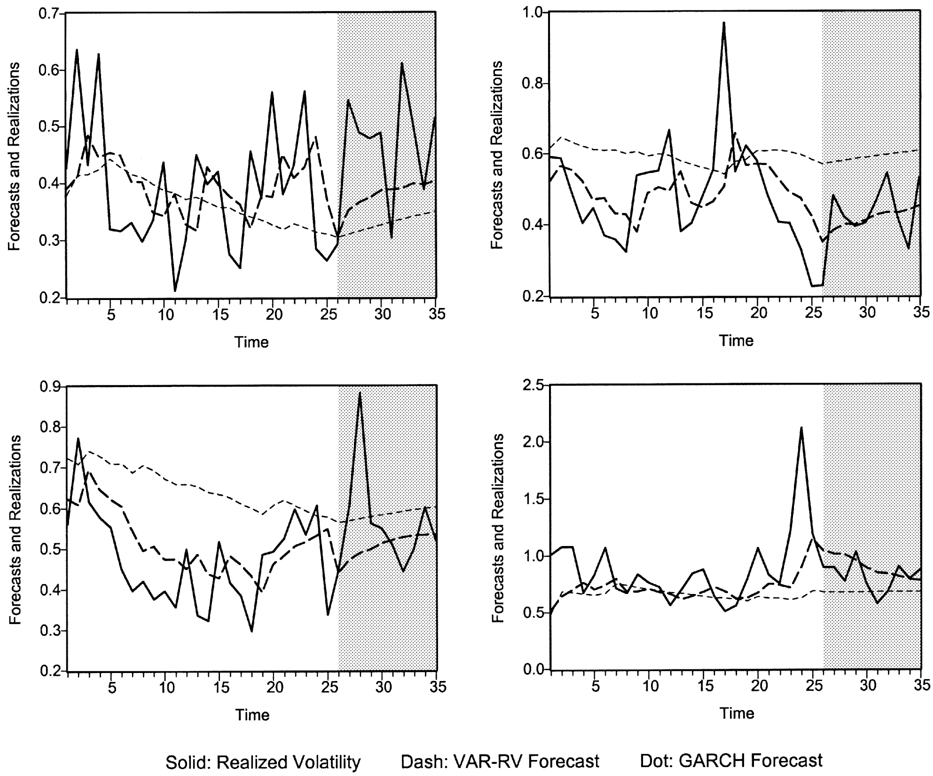


FIGURE 6.—Realized volatility and two out-of-sample forecasts. The figure shows four out-of-sample episodes of thirty-five days each. For each of the first twenty-five days, we show the daily realized volatility together with the one-day-ahead forecasts made on the previous day using the VAR-RV and GARCH(1, 1) models. Then, for days twenty-six through thirty-five (shaded), we continue to show daily realized volatility, but we show multi-step VAR-RV and GARCH forecasts based on information available on day twenty-five. Hence the forecasts for day twenty-six are one-day-ahead, the forecasts for day twenty-seven are two-day-ahead, and so on. See the main text for details.

of current (day twenty-five) volatility, and then revert to the unconditional mean of approximately 0.6 in a fashion consistent with historical experience. In contrast, the lack of an equally accurate assessment of current volatility within the GARCH setting results in markedly inferior longer-run volatility forecasts.

The above findings do not reflect a failure of the GARCH model per se, but rather the efficacy of exploiting volatility measures based on intraday return observations. In fact, in a related empirical study Andersen and Bollerslev (1998) find that a daily GARCH(1, 1) model explains about as much of the future variation in daily exchange rate volatility as is theoretically feasible if the model were true in population. The point is that the noise in daily squared returns necessarily renders the measurements of the current volatility innovation imprecise, independent of the correct model for the daily volatility.

### 6.3. *On the Role of Smoothing in the Construction of Realized Volatility Forecasts*

Although the realized volatility is less noisy than, for example, the daily squared or absolute returns, it nevertheless contains measurement error, as emphasized in recent work by Andreou and Ghysels (2002), Bai, Russell, and Tiao (2000), Barndorff-Nielsen and Shephard (2002a, b), and Meddahi (2002) among others; see also the discussion in Andersen, Bollerslev, and Diebold (2002). Importantly, the approach to volatility modeling and forecasting that we advocate here remains appropriate in the presence of such measurement errors.

Intuitively, because the realized volatility is subject to measurement error, it seems desirable that fitted and forecasted realized volatilities should—one way or another—involve *smoothing*, to reduce the effects of the error. Indeed, this is the basic message of much recent work on realized volatility measurement. For example, Andreou and Ghysels (2002) recommend smoothing in a nonparametric fashion, while Barndorff-Nielsen and Shephard (2002a) work with a specific stochastic volatility model, which allows them to quantify the distribution of the measurement error in the realized volatility proxy, and then to fit and forecast the corresponding latent integrated volatility (quadratic variation) using an optimal nonlinear smoother and filter based on a state-space representation of the model. As such, both the fitted and forecasted volatilities are ultimately smoothed functions of the history of the daily realized volatilities. Meanwhile, the measurement errors in the (unsmoothed) realized volatilities are (approximately) serially uncorrelated under quite general conditions. This justifies our fitting of a reduced form model directly to the realized volatility without explicitly accounting for the measurement error in the dependent variable. Of course, our approach also involves smoothing, if only implicitly, as both the fitted and forecasted volatilities become smoothed functions of the history of the daily realized volatilities.

This direct approach to volatility modeling and forecasting is further corroborated by the recent theoretical results in Andersen, Bollerslev, and Meddahi (2002) for the general class of eigenfunction stochastic volatility models. In particular, population forecasts of the (latent) integrated volatility formed by projecting the volatility on the history of the realized volatilities are almost as accurate as forecasts formed by projecting on the (unattainable) history of integrated volatilities. This implies, among other things, that the measurement error component is largely irrelevant for forecasting: "feasible" forecasts based on the history of the realized volatilities are approximately as accurate as "infeasible" forecasts based on the history of the integrated volatilities, or on any other empirical smoothed volatility measures, including for example those of Andreou and Ghysels (2002). Moreover, the theoretical difference between the projection on current realized volatility versus current and past realized volatilities is small, suggesting that parsimonious ARMA-type models fit directly to the realized volatility—precisely as advocated in this paper—should perform well in practical forecasting situations.

### 6.4. *Density Forecasts and Quantile Calculations: a VAR for VaR*

In light of the good performance of the VAR-RV volatility forecasts, we now turn to an investigation of the corresponding return density forecasts and VaR calculations, in which we explicitly incorporate the theoretically motivated distributional features of the realized volatilities and standardized returns highlighted in Section 4.

Measuring and forecasting portfolio Value-at-Risk, or VaR, and fluctuations in VaR due to changing market conditions and/or portfolio shares, is an important part of modern financial risk management; see, e.g., Gouriéroux, Laurent, and Scaillet (2000).[40] Our results suggest that accurate return density forecasts and associated VaR estimates may be obtained from a long-memory VAR for realized volatility, coupled with the assumption of normally distributed standardized returns. We assess this conjecture using the methods of Diebold, Gunther, and Tay (1998). The basic idea is that a good density forecast should satisfy two criteria. First, the nominal $\rho$ percent VaR should be exceeded only $\rho$ percent of the time, for all $\rho$, which we call correct unconditional calibration. Second, a violation of nominal $\rho_1$ percent VaR today should convey no information as to whether nominal $\rho_2$ percent VaR will be violated tomorrow, for all $\rho_1$ and $\rho_2$. If a density forecast satisfies the two criteria, we say that it is correctly conditionally calibrated. More formally, suppose that the daily returns, $r_t$, are generated from the series of one-day-ahead conditional densities, $f(r_t \mid \mathcal{F}_{t-1})$, where $\mathcal{F}_{t-1}$ denotes the full information set available at time $t-1$. If the series of one-day-ahead conditional density forecasts, $f_{t|t-1}(r_t)$, coincides with $f(r_t \mid \mathcal{F}_{t-1})$, it then follows under weak conditions that the sequence of probability integral transforms of $r_t$ with respect to $f_{t|t-1}(\cdot)$ should be iid uniformly distributed on the unit interval. That is, $\{z_t\}$ is distributed as iid $U(0,1)$, where we define the probability integral transform, $z_t$, as the cumulative density function corresponding to $f_{t|t-1}(\cdot)$ evaluated at $r_t$; i.e., $z_t = \int_{-\infty}^{r_t} f_{t|t-1}(u)\,du$. Hence the adequacy of the VAR-RV based volatility forecast and the lognormal-normal mixture distribution may be assessed by checking whether the corresponding distribution of $\{z_t\}$ is iid $U(0,1)$.

To this end, we report in Table IV the percentage of the realized DM/\$, ¥/\$, and equally-weighted portfolio returns that are less than various quantiles forecast by the long-memory lognormal-normal mixture model. The close correspondence between the percentages in each column and the implied quantiles by the model, both in-sample and out-of-sample, is striking. It is evident that the VAR-RV lognormal-normal mixture model affords a very close fit for all of the relevant VaRs.

Although uniformity of the $z_t$ sequences is necessary for adequacy of the density forecasts, it is not sufficient. The $z_t$'s must be independent as well, to guarantee, for example, that violation of a particular quantile forecast on day $t$ conveys

[40] Portfolio VaR at confidence level $\rho$ percent and horizon $k$ is simply the $\rho$th percentile of the $k$-step-ahead portfolio return density forecast. For an overview, see Duffie and Pan (1997). When calculating VaR at confidence level $\rho$ and horizon $k$, the appropriate values of $\rho$ and $k$ are generally situation-specific, although the Basel Committee on Banking Supervision has advocated the systematic use of five- and one-percent VaR, and one- and ten-day horizons.

TABLE IV

DISTRIBUTIONS OF ONE-DAY-AHEAD PROBABILITY INTEGRAL TRANSFORMS

| Quantile[a] | 1% | 5% | 10% | 90% | 95% | 99% |
|---|---|---|---|---|---|---|
| *In-Sample*[b] | | | | | | |
| DM/$ | 0.016 | 0.059 | 0.105 | 0.895 | 0.943 | 0.987 |
| ¥/$ | 0.016 | 0.061 | 0.103 | 0.901 | 0.951 | 0.990 |
| Portfolio[d] | 0.010 | 0.052 | 0.091 | 0.912 | 0.958 | 0.990 |
| *Out-of-Sample*[c] | | | | | | |
| DM/$ | 0.005 | 0.045 | 0.092 | 0.893 | 0.941 | 0.990 |
| ¥/$ | 0.019 | 0.055 | 0.099 | 0.884 | 0.956 | 0.993 |
| Portfolio[d] | 0.010 | 0.042 | 0.079 | 0.909 | 0.968 | 1.000 |

[a] Selected points of the empirical cumulative distribution function for the probability integral transform of returns with respect to the density forecasts from our VAR-RV long-memory lognormal-normal mixture model.
[b] The in-sample forecast evaluation period is December 1, 1987 through December 1, 1996.
[c] The out-of-sample period is December 2, 1996 through June 30, 1999.
[d] Portfolio is an equally-weighted DM/$, ¥/$ portfolio.

no information regarding its likely violation on day $t+1$. In general, dependence in $z_t$ would indicate that dynamics have been inadequately modeled and captured by the forecasts. To assess independence, Figure 7 therefore plots the sample autocorrelation functions for $(z_t - \bar{z})$ and $(z_t - \bar{z})^2$ corresponding to the one-day ahead out-of-sample density forecasts for the DM/$, the ¥/$, and the equally-weighted portfolio returns. All told, there is no evidence of serial correlation in any of the six series, indicating that the model's density forecasts are not only correctly unconditionally calibrated, but also correctly conditionally calibrated. This apparent lack of dependence is confirmed by formal Ljung-Box portmanteau tests for the joint significance of the depicted autocorrelations and also carries over to the in-sample period, as further detailed in the supplemental Appendix.

Finally, it is worth noting that although the lognormal distribution is very convenient from an empirical modeling perspective, it is not closed under temporal aggregation. Consequently, the lognormal-normal mixture distribution is formally horizon specific. However, as shown in the Appendix, almost identical results obtain both in- and out-of-sample for the one- and ten-day horizons using an inverse Gaussian-normal mixture distribution.[41] This latter distribution is, of course, closed under temporal aggregation.

## 7. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Guided by a general theory for continuous-time arbitrage-free price processes, we develop a framework for the direct modeling and forecasting of realized volatility and correlation. Our methods are simple to implement empirically, even in multivariate situations. We illustrate the idea in the context of the foreign

[41] This parallels the findings related to the unconditional distribution of the realized volatility in Barndorff-Nielsen and Shephard (2002a), who find the inverse Gaussian and the lognormal distributions to be virtually indistinguishable empirically.

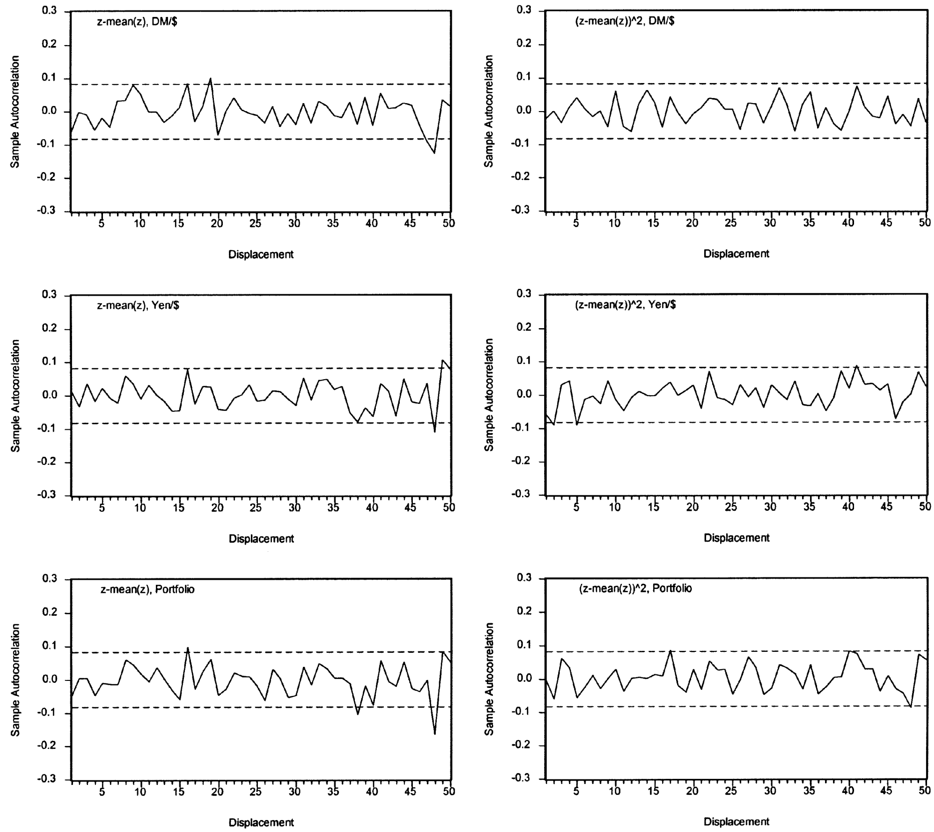


FIGURE 7.— Dependence structure of probability integral transforms, out-of-sample one-day-ahead density forecasts. The figure graphs the sample autocorrelation functions of $(z_t - \bar{z})$ and $(z_t - \bar{z})^2$, where $z_t$ denotes the probability integral transform of returns with respect to the one-day-ahead density forecasts from our long-memory lognormal-normal mixture model; see the main text for details. The out-of-sample period is December 2, 1996 through June 30, 1999. The three subplots correspond to DM/$ returns, ¥/$ returns, and equally-weighted portfolio returns. The dashed lines are Bartlett two standard error bands.

exchange market, specifying and estimating a long-memory Gaussian VAR for a set of daily logarithmic realized volatilities. The model produces very successful volatility forecasts, generally dominating those from conventional GARCH and related approaches. It also generates well-calibrated density forecasts and associated quantile, or VaR, estimates for our multivariate foreign exchange application. Numerous interesting directions for future research remain.

First, the realized volatility measures used in this paper do not distinguish between variability originating from continuous price movements or jumps. However, as discussed in Section 2, the dynamic impact may differ across the two sources of variability. Hence, it is possible that improved volatility forecasting models may be constructed by explicitly modeling the jump component, if present. Recent results in Maheu and McCurdy (2002) based on Markov switching models also suggest that explicitly accounting for nonlinear features in the realized volatility may result in even better volatility forecasts.

Second, although the lognormal-normal and inverse Gaussian-normal mixture distributions both work very well in the present context, the predictive distribution could be refined and adapted to more challenging environments using the numerical simulation methods of Geweke (1989), the Cornish-Fisher expansion of Baillie and Bollerslev (1992), or the recalibration methods of Diebold, Hahn, and Tay (1999). This would, for example, be required to deal with conditionally asymmetric return distributions over longer horizons arising from a significant correlation between volatility and return innovations.

Third, although we focused on using density forecasts to obtain VaR estimates (quantiles), the same density forecasts could of course be used to calculate other objects of interest in financial risk management. Examples include the probability of loss exceeding a specified threshold (shortfall probabilities), and the expected loss conditional upon loss exceeding a pre-specified threshold (expected shortfall), as discussed for example in Heath, Delbaen, Eber, and Artzner (1999) and Basak and Shapiro (2001).

Fourth, our approach to exchange rate density forecasting could be extended to other classes of financial assets. Although the structure of our proposed modeling framework builds directly on the empirical regularities for the foreign exchange markets documented in Section 4, the empirical features characterizing other asset markets appear remarkably similar, as shown for example by Andersen, Bollerslev, Diebold, and Ebens (2001) for U.S. equities.

Fifth, and perhaps most importantly, volatility forecasts figure prominently in many practical financial decisions extending well beyond risk management into spot and derivative asset pricing (see, for example, Bollerslev and Mikkelsen (1999)) and portfolio allocation (see, for example, Busse (1999) and Fleming, Kirby, and Ostdiek (2001, 2002)). It will be of interest to explore the gains afforded by the simple volatility modeling and forecasting procedures developed here, particularly in high-dimensional settings, and to compare the results to those arising from more standard recent multivariate volatility modeling procedures such as Engle (2002) and Tse and Tsui (2002). In this regard, a couple of issues merit particular attention. One critical task is to develop realized volatility forecasting models that are parameterized in ways that guarantee positive definiteness of forecasted covariance matrices within high-dimensional settings. Because the in-sample realized covariance matrix is positive definite under quite general conditions, one approach would be to model the Cholesky factors rather than the realized covariance matrix itself. The corresponding forecasts for the Cholesky factors are then readily transformed into forecasts for the future variances and covariances by simple matrix multiplications. More precisely, under the conditions of Lemma 1 it follows that $V_t$ is positive definite, $t = 1, 2, \ldots, T$. Hence, there exists a corresponding unique sequence of lower triangular Cholesky factors, $P_t$, such that $V_t = P_t P_t'$, $t = 1, 2, \ldots, T$. The resulting data vector, $\text{vech}(P_t)$, $t = 1, 2, \ldots, T$, may be modeled directly and used to produce a forecast, say $\text{vech}(P_{T+h|T})$, which may in turn be converted back into a forecast of $V_{T+h|T} \equiv P_{T+h|T} P'_{T+h|T}$.

Finally, and also of particular relevance in high-dimensional situations, allowing for factor structure in the modeling and forecasting of realized volatility may prove useful, as factor structure is central to both empirical and theoretical financial economics. Previous research on factor volatility models has typically relied on complex procedures involving a latent volatility factor, as for example in Diebold and Nerlove (1989), Engle, Ng, and Rothschild (1990), King, Sentana, and Wadhwani (1994), and Meddahi and Renault (2002). In contrast, factor analysis of realized volatility should be relatively straightforward, even in high-dimensional environments. Moreover, the identification of explicit volatility factors, and associated market-wide variables that underlie the systematic volatility movements, may help to provide an important step towards a better understanding of "the economics of volatility."

*Dept. of Finance, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A., and NBER; t-andersen@kellogg.northwestern.edu,*

*Dept. of Economics, Duke University, Durham, NC 27708, U.S.A., and NBER; boller@econ.duke.edu,*

*Dept. of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, U.S.A., and NBER; fdiebold@sas.upenn.edu,*

*and*

*Charles River Associates, Inc., 170 South Main St., Suite 500, Salt Lake City, UT 84101, U.S.A.; plabys@crai.com*

## REFERENCES

ANDERSEN, T. G., L. BENZONI, AND J. LUND (2002): "Estimating Jump-Diffusions for Equity Returns," *Journal of Finance*, 57, 1239–1284.

ANDERSEN, T. G., AND T. BOLLERSLEV (1997a): "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975–1005.

——— (1997b), "Intraday Seasonality and Volatility Persistence in Foreign Exchange and Equity Markets," *Journal of Empirical Finance*, 4, 115–158.

——— (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885–905.

ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2002): "Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics*, ed. by L. P. Hansen and Y. A-Sahalia. Amsterdam: North-Holland, forthcoming.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND H. EBENS (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2000a): "Exchange Rate Returns Standardized by Realized Volatility Are (Nearly) Gaussian," *Multinational Finance Journal*, 4, 159–179.

——— (2000b): "Market Microstructure Effects and the Estimation of Integrated Volatility," Northwestern University, Duke University and University of Pennsylvania, Work in Progress.

——— (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

ANDERSEN, T. G., T. BOLLERSLEV, AND S. LANGE (1999): "Forecasting Financial Market Volatility: Sample Frequency vis-a-vis Forecast Horizon," *Journal of Empirical Finance*, 6, 457–477.

ANDERSEN, T. G., T. BOLLERSLEV, AND N. MEDDAHI (2002): "Analytic Evaluation of Volatility Forecasts," Northwestern University, Duke University, and University of Montreal, Unpublished Manuscript.

ANDREOU, E., AND E. GHYSELS (2002): "Rolling-Sample Volatility Estimators: Some New Theoretical, Simulation and Empirical Results," *Journal of Business and Economic Statistics*, 20, 363–376.

BACK, K. (1991): "Asset Prices for General Processes," *Journal of Mathematical Economics*, 20, 317–395.

BAI, X., J. R. RUSSELL, AND G. C. TIAO (2000): "Beyond Merton's Utopia: Effects of Non-Normality and Dependence on the Precision of Variance Estimates Using High-Frequency Financial Data," Manuscript, Graduate School of Business, University of Chicago.

BAILLIE, R. T., AND T. BOLLERSLEV (1989): "The Message in Daily Exchange Rates: A Conditional Variance Tale," *Journal of Business and Economic Statistics*, 7, 297–305.

——— (1992): "Prediction in Dynamic Models with Time-Dependent Conditional Variances," *Journal of Econometrics*, 52, 91–113.

BAILLIE, R. T., T. BOLLERSLEV, AND H. O. MIKKELSEN (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3–30.

BAKSHI, G., C. CAO, AND Z. CHEN (1997): "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049.

BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2001): "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," (with discussion), *Journal of the Royal Statistical Society, Series B*, 63, 167–241.

——— (2002a): "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, Part 2, 253–280.

——— (2002b): "Estimating Quadratic Variation Using Realised Variance," *Journal of Applied Econometrics*, 17, 457–477.

BASAK, S., AND A. SHAPIRO (1999): "Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices," *Review of Financial Studies*, 14, 371–405.

BATES, D. (2000): "Post-'87 Crash Fears in S&P 500 Futures Options," *Journal of Econometrics*, 94, 181–238.

BELTRATTI, A., AND C. MORANA (1999): "Computing Value-at-Risk with High-Frequency Data," *Journal of Empirical Finance*, 6, 431–455.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

——— (1987): "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, 542–547.

BOLLERSLEV, T., R. F. ENGLE, AND D. B. NELSON (1994): "ARCH Models," in *Handbook of Econometrics, Volume IV*, ed. by. R. F. Engle and D. McFadden. Amsterdam: North-Holland, 2959–3038.

BOLLERSLEV, T., AND H. O. MIKKELSEN (1996): "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 151–184.

——— (1999): "Long-Term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 92, 75–99.

BOLLERSLEV, T., AND J. H. WRIGHT (2001): "High-Frequency Data, Frequency Domain Inference, and Volatility Forecasting," *Review of Economics and Statistics*, 83, 596–602.

BUSSE, J. A. (1999): "Volatility Timing in Mutual Funds: Evidence from Daily Returns," *Review of Financial Studies*, 12, 1009–1041.

CAI, J., Y. L. CHEUNG, R. S. K. LEE, AND M. MELVIN (2001): "Once-in-a-Generation' Yen Volatility in 1998: Fundamentals, Intervention and Order Flow," *Journal of International Money and Finance*, 20, 327–347.

CHONG, Y. Y., AND D. F. HENDRY (1986): "Econometric Evaluation of Linear Macroeconomic Models," *Review of Economic Studies*, 53, 671–690.

CHRISTOFFERSEN, P. F., AND F. X. DIEBOLD (1997): "Optimal Prediction Under Asymmetric Loss," *Econometric Theory*, 13, 808–817.

——— (1999): "How Relevant is Volatility Forecasting for Financial Risk Management," *Review of Economics and Statistics*, 82, 12–22.

CLARK, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–155.

COMTE, F., AND E. RENAULT (1998): "Long Memory in Continuous Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291–323.

DACOROGNA, M. M., U. A. MÜLLER, R. J. NAGLER, R. B. OLSEN, AND O. V. PICTET (1993): "A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market," *Journal of International Money and Finance*, 12, 413–438.

DIEBOLD, F. X., T. GUNTHER, AND A. S. TAY (1998): "Evaluating Density Forecasts, with Applications to Financial Risk Management," *International Economic Review*, 39, 863–883.

DIEBOLD, F. X., J. HAHN, AND A. S. TAY (1999): "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange," *Review of Economics and Statistics*, 81, 661–673.

DIEBOLD, F. X., AND M. NERLOVE (1989): "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, 4, 1–22.

DING, Z., C. W. J. GRANGER, AND R. F. ENGLE (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83–106.

DROST, F. C., AND T. E. NIJMAN (1993): "Temporal Aggregation of GARCH Processes," *Econometrica*, 61, 909–927.

DROST, F. C., AND B. J. M. WERKER (1996): "Closing the GARCH Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31–57.

DUFFIE, D., AND J. PAN (1997): "An Overview of Value-at-Risk," *Journal of Derivatives*, 4, 9–49.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987–1008.

——— (2000): "Discussion of 'The Distribution of Exchange Rate Volatility,' by Andersen, Bollerslev, Diebold and Labys," Western Finance Association Meetings, Sun Valley, Idaho.

——— (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics*, 20, 339–350.

ENGLE, R. F., AND K. KRONER (1995): "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, 122–150.

ENGLE, R. F., V. NG, AND M. ROTHSCHILD (1990): "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, 213–238.

ERAKER, B., M. JOHANNES, AND N. POLSON (2002): "Evidence for and the Impact of Jumps in Volatility and Returns," *Journal of Finance*, forthcoming.

FAMA, E. F. (1965): "The Behavior of Stock Market Prices," *Journal of Business*, 38, 34–105.

FLEMING, J., C. KIRBY, AND B. OSTDIEK (2001): "The Economic Value of Volatility Timing," *Journal of Finance*, 56, 329–352.

——— (2002): "The Economic Value of Volatility Timing Using Realized Volatility," *Journal of Financial Economics*, forthcoming.

FRENCH, K. R., G. W. SCHWERT, AND R. F. STAMBAUGH (1987): "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–29.

GEWEKE, J. (1989): "Exact Predictive Densities for Linear Models with ARCH Disturbances," *Journal of Econometrics*, 40, 63–86.

GEWEKE, J., AND S. PORTER-HUDAK (1983): "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

GHYSELS, E., A. HARVEY, AND E. RENAULT (1996): "Stochastic Volatility," in *Handbook of Statistics, Vol. 14, Statistical Methods in Finance*, ed. by G. S. Maddala. Amsterdam: North-Holland, 119–191.

GOURIÉROUX, C., J. P., LAURENT, AND O. SCAILLET (2000): "Sensitivity Analysis of Values at Risk," *Journal of Empirical Finance*, 7, 225–245.

HANSEN, P. R., AND A. LUNDE (2001): "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1, 1)?" Unpublished Manuscript, Department of Economics, Brown University.

HEATH, D., F. DELBAEN, J. M. EBER, AND P. ARTZNER (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.

HSIEH, D. A. (1989): "Modeling Heteroskedasticity in Daily Foreign Exchange Rates," *Journal of Business and Economic Statistics*, 7, 307–317.

——— (1991): "Chaos and Nonlinear Dynamics: Application to Financial Markets," *Journal of Finance*, 46, 1839–1877.

J. P. MORGAN (1997): *RiskMetrics Technical Documents* (Fourth Edition). New York: J. P. Morgan Co.

KARATZAS, I., AND S. E. SHREVE (1991): *Brownian Motion and Stochastic Calculus* (Second Edition). New York: Springer-Verlag.

KING, M., E. SENTANA, AND S. WADHWANI (1994): "Volatility and Links Between National Stock Markets," *Econometrica*, 62, 901–933.

MANDELBROT, B. (1963): "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394–419.

MAHEU, J. M., AND T. H. MCCURDY (2002): "Nonlinear Features of Realized FX Volatility," *Review of Economics and Statistics*, 84, 668–681.

MARTENS, M. (2001): "Forecasting Daily Exchange Rate Volatility Using Intraday Returns," *Journal of International Money and Finance*, 20, 1–23.

MARTENS, M., Y. C. CHANG, AND S. J. TAYLOR (2002): "A Comparison of Seasonal Adjustment Methods when Forecasting Intraday Volatility," *Journal of Financial Research*, 25, 283–299.

MEDDAHI, N. (2002): "A Theoretical Comparison Between Integrated and Realized Volatilities," *Journal of Applied Econometrics*, 17, 479–508.

MEDDAHI, N., AND E. RENAULT (2002): "Temporal Aggregation of Volatility Models," *Journal of Econometrics*, forthcoming.

MERTON, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

——— (1982): "On the Mathematics and Economics Assumptions of Continuous-Time Models," in *Financial Economics: Essays in Honor of Paul Cootner*, ed. by W. F. Sharpe and C. M. Cootner. Englewood Cliffs, New Jersey: Prentice-Hall.

MINCER, J., AND V. ZARNOWITZ (1969): "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations*, ed. by J. Mincer. New York: National Bureau of Economic Research.

MÜLLER, U. A., M. M. DACOROGNA, R. B. OLSEN, O. V. PICTET, M. SCHWARZ, AND C. MORGENEGG (1990): "Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law, and Intraday Analysis," *Journal of Banking and Finance*, 14, 1189–1208.

MÜLLER, U. A., M. M. DACOROGNA, R. D. DAVE, R. B. OLSEN, O. V. PICTET, AND J. E. WEIZSÄCKER (1997): "Volatilities of Different Time Resolution—Analyzing the Dynamics of Market Components," *Journal of Empirical Finance*, 4, 213–239.

NELSON, D. B., AND D. FOSTER (1995): "Filtering and Forecasting with Misspecified ARCH Models II: Making the Right Forecast with the Wrong Model," *Journal of Econometrics*, 67, 303–335.

PAN, J. (2002): "The Jump Risk Premium Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

PROTTER, P. (1992): *Stochastic Integration and Differential Equations: A New Approach* (Second Edition). New York: Springer Verlag.

RENAULT, E. (1997): "Econometric Models of Option Pricing Errors," in *Advances in Economics and Econometrics: Theory and Applications,* ed. by D. M. Kreps and K. F. Wallis. Cambridge: Cambridge University Press, 223–278.

ROBINSON, P. M. (1991): "Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression," *Journal of Econometrics*, 47, 67–84.

——— (1995): "Log-Periodogram Regression of Time Series with Long-Range Dependence," *Annals of Statistics*, 23, 1048–1072.

ROBINSON, P. M., AND D. MARINUCCI (2001): "Semiparametric Fractional Cointegration Analysis," *Journal of Econometrics*, 105, 225–247.

SCHWERT, G. W. (1989): "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance*, 44, 1115–1153.

SHEEDY, E. (1998): "Correlation in Currency Markets: A Risk Adjusted Perspective," *Journal of International Financial Markets, Institutions and Money*, 8, 59–82.

TAYLOR, S. J. (1986): *Modelling Financial Time Series*. Chichester: John Wiley and Sons.

TAYLOR, S. J., AND X. XU (1997): "The Incremental Volatility Information in One Million Foreign Exchange Quotations," *Journal of Empirical Finance*, 4, 317–340.

TSE, Y. K., AND A. K. C. TSUI (2002): "A Multivariate GARCH Model with Time-Varying Correlations," *Journal of Business and Economic Statistics*, 20, 351–362.

WEST, K. D., AND M. W. MCCRACKEN (1998): "Regression-Based Tests of Predictive Ability," *International Economic Review*, 39, 817–840.

ZHOU, B. (1996): "High-Frequency Data and Volatility in Foreign Exchange Rates," *Journal of Business and Economic Statistics*, 14, 45–52.

ZUMBACH, G., F. CORSI, AND A. TRAPLETTI (2002): "Efficient Estimation of Volatility using High-Frequency Data," Manuscript, Olsen & Associates, Research Institute for Applied Economics, Zürich, Switzerland.