Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance

A. Ronald Gallant; Chien-Te Hsu; George Tauchen

USING DAILY RANGE DATA TO CALIBRATE VOLATILITY DIFFUSIONS AND EXTRACT THE FORWARD INTEGRATED VARIANCE

A. Ronald Gallant, Chien-Te Hsu, and George Tauchen*

Abstract—A common model for security price dynamics is the continuous-time stochastic volatility model. For this model, Hull and White (1987) show that the price of a derivative claim is the conditional expectation of the Black-Scholes price with the forward integrated variance replacing the Black-Scholes variance. Implementing the Hull and White characterization requires both estimates of the price dynamics and the conditional distribution of the forward integrated variance given observed variables. Using daily data on close-to-close price movement and the daily range, we find that standard models do not fit the data very well and that a more general three-factor model does better, as it mimics the long-memory feature of financial volatility. We develop techniques for estimating the conditional distribution of the forward integrated variance given observed variables.

I. Introduction

This paper has two objectives: The first is to extend and implement methods for estimating diffusion models of securities prices with unobserved stochastic volatility. Consistent with several recent studies, we find that a two-factor model can generally account for the dynamics of observations on the daily close-to-close price movement for modest time spans (ten to twenty years). One of the factors is the price process itself and the second is its local volatility. However, we also find that the two-factor model cannot account for either the marginal dynamics of the daily range, viewed as a single series, nor for the joint bivariate dynamics of the daily close-to-close price movement series and daily range series taken together. We implement a three-factor model in which local price volatility is represented as the sum of two separate volatility factors, each governed by its own diffusion process. We find that the three-factor model does a better job of fitting the data, and, in particular, mimics the long-memory type behavior of financial volatility (Ding et al., 1993; Baillie et al., 1996; Bollerslev & Mikkelsen, 1996). The three-factor model is Markov with two unobserved state variables. Among other things, the Markov property makes the model relatively easy to simulate as it circumvents all computational difficulties associated with fractional integration calculations (Bollerslev & Mikkelsen, 1996; Gallant et al., 1997).

The second objective is to develop methods for extracting from observable data the conditional mean of—and, more generally, the conditional distribution of—the integrated local price variance over an arbitrary horizon of finite length. This objective is motivated by Hull and White (1987), who show that, under certain assumptions, the price of a deriva-

tive claim can be expressed as the expected value of the Black-Scholes price evaluated at the integrated volatility. The methods are extensions of techniques developed in Gallant and Tauchen (1998), who generated an estimator of the discrete-time conditional density of the observed security price process given its past as implied by a diffusion model. Here we estimate the conditional distribution of the unobserved forward integrated volatility process given the observed history. In our application, the observed history consists of lags of two series: the daily close-to-close log price movement and high/low range series. The techniques are completely general, however, and do not require the range information. They could be applied, for example, with just the price series alone, and sampled either more finely or more coarsely than the daily frequency.

II. Setup

Let p_t denote the log of a financial price series evolving in continuous time. We shall describe and subsequently estimate a class of stochastic differential equation models for the price series. Our models are stochastic volatility diffusion models. A basic form is

$$dp_t = a_{1t}dt + \sigma_{pt}dW_{1t}$$
$$d\log(\sigma_{pt}) = a_{2t}dt + b_{2t}dW_{2t} \tag{1}$$

where the first equation describes the evolution of the price process and the second describes that of the volatility process. In the above, a_{1t} is the local drift of the price, σ_{pt} is the local volatility, while a_{2t} and b_{2t} are the local drift and diffusion of the log volatility process, and W_{1t} and W_{2t} are independent Brownian motions. The model (1) is a diffusion stochastic volatility model of the type studied by Nelson (1992), Andersen and Lund (1997), and many others. In what follows, we let $\{p_t\}_{t \in [0,T]}$ denote a realization from model (1), we let $\{p_t\}_{t=0,1,\ldots,n}$ denote a discrete, equispaced sampling of length n of the diffusion, and we let $\{\Delta p_t\}_{t=1,\ldots,n}$, where $\Delta p_t = p_t - p_{t-1}$ denote the price movement series.

With rare exception, one observes only the discrete price process $\{p_t\}_{t=0,1,\ldots,n}$ and not the volatility, σ_{pt}, so the model (1) contains an unobserved Markov state variable. For reasons discussed in Hansen and Scheinkman (1995), Tauchen (1997), and Gallant and Tauchen (1997a), models with unobserved state variables are well adapted to estimation using simulation-based techniques such as Indirect Inference (Gourieroux et al., 1993), Simulated Method of Moments (Ingram & Lee, 1991; Duffie & Singleton, 1993), or Efficient Method of Moments (EMM) (Gallant & Tauchen, 1997a, 1998). For example, Engle and Lee (1997) estimate specifications of model (1) using the EMM technique

(Gallant & Tauchen, 1997a) on daily data on S&P returns, the DM/$ exchange rate, and weekly Treasury Bill yields. Gallant and Long (1997) do so as well for daily DM/$ data, while Andersen and Lund (1997) and Gallant and Tauchen (1998) do so using weekly Treasury Bill yields.

In estimation using discretely sampled data $\{p_t\}_{t=0,1,\ldots,n}$, information on the level of volatility σ_{pt}, and the trajectory of volatility $\{\sigma_{ps}\}_{s\in[t-1,t]}$ over the sampling interval, comes indirectly from the magnitude of the unexpected time price movement $|\Delta p_t - \mathcal{E}_{t-1}(\Delta p_t)|$. Typically, the magnitude of the unexpected price movement is a very noisy indicator of the trajectory of volatility over the sampling interval (Andersen & Bollerslev, 1998).

Financial economists have long known that the daily range of the price series contains extra information about the course of volatility over the day. Within a constant volatility framework, Garman and Klass (1980) and Parkinson (1980) show that use of the range can improve volatility estimates by as much as a factor of eight over the standard estimate. Beckers (1983) and Hsieh (1991) present related results and strong empirical documentation on the efficiency improvement.

This paper adopts the stochastic volatility framework of model (1). In this context, Andersen and Bollerslev (1998) present simulation evidence that there is extra information about the integrated daily variance $\int_{t-1}^t \sigma_{ps}^2\, ds$ embodied in the daily range. In our notation, the range is

$$d_t = \max_{t-1 < \tau \le t} \{p_\tau - p_{t-1}\} - \min_{t-1 < \tau \le t} \{p_\tau - p_{t-1}\},$$
$$t = 1, 2, \ldots.$$

Using results on subordination and the running maximum of Brownian motion in Karatzas and Shreve (1991), one can show that

$$d_t \stackrel{\mathscr{D}}{=} (H_1 + H_2)\left(\int_{t-1}^t \sigma_{ps}^2\, ds\right)^{1/2} \tag{2}$$

where H_1 and H_2 are negatively correlated half-Gaussian random variables. Thus, the range is seen to be an indicator variable for the integrated daily variance.

We extend previous work dealing with daily range data by actually estimating and testing continuous-time models in a stochastic volatility setting. We estimate various specifications of model (1) using price movement data $\{\Delta p_t\}_{t=1,\ldots,n}$ and range data $\{d_t\}_{t=1,\ldots,n}$, each taken separately. We also estimate specifications of model (1) and some generalizations using the joint price movement and range series $\{\Delta p_t, d_t\}_{t=1,\ldots,n}$. We find that indeed the range series contains extra information about the appropriate specification of the diffusion model and the trajectory of volatility.

For notational reasons, it proves convenient to consider an underlying state vector U_t and specify its dynamics. We view p_t and its drift and diffusion as functions of the state vector.

In the version we call the two-factor model, U_t is 2×1 and is assumed to evolve as

Two factor:

$$dU_{1t} = \alpha_{10}dt + \beta_{10} \exp(U_{2t})dW_{1t}$$
$$dU_{2t} = \alpha_{22}U_{2t}dt + (\beta_{20} + \beta_{22}U_{2t})dW_{2t} \tag{3}$$

where W_{1t} and W_{2t} are independent Brownian motions. We take

$$p_t = U_{1t}$$
$$\sigma_{pt} = \beta_{10} \exp(U_{2t}).$$

The first equation of (3) governs the evolution of the price series. It allows for a drift component $\alpha_{10}dt$, although we typically expect the drift to be very small (although above the instantaneous risk-free rate) as financial price series display very modest mean reversion and act more as nonstationary integrated processes. The component U_{2t} reflects stochastic volatility in the evolution of the price process, and we refer to the second component U_{2t} of the state as the volatility component, or just volatility for short. The absence of the intercept in the drift component U_{2t} reflects a normalization for identification. The specification (3) is the diffusion analogue of an exponential GARCH (E-GARCH) studied by Nelson (1992), Drost and Nijman (1993), and Drost and Werker (1996), with the generalization that the local volatility of U_{2t} is itself state dependent.

As will be seen, the preceding setup does a reasonably good job of fitting the observed price movement series $\{\Delta p_t\}_{t=1,2,\ldots}$. However, it does a very poor job of fitting the dynamics of the range series $\{d_t\}_{t=1,2,\ldots}$ and the joint dynamics of the two series together $\{\Delta p_t, d_t\}_{t=1,2,\ldots}$. We find that introduction of a second stochastic volatility factor—thus bringing the total number of factors to three—improves the fit. The most general model we consider is

Three Factor:

$$dU_{1t} = \alpha_{10}dt + \beta_{10} \exp(U_{2t} + U_{3t})dW_{1t}$$
$$dU_{2t} = \alpha_{22}U_{2t}dt + (\beta_{20} + \beta_{22}U_{2t})dW_{2t} \tag{4}$$
$$dU_{3t} = \alpha_{33}U_{3t}dt + (\beta_{30} + \beta_{33}U_{3t})dW_{3t}.$$

The EMM estimator we employ is a Simulated Method of Moments (SMM) estimator. Such estimators work by using very long simulations from (3) or (4) to compute predicted moments of observed discrete-time data. Let $\{\tilde{p}_t\}_{t=0,1,\ldots,n}$ denote the observed log price series. As is typical of nearly all financial levels series, this series is an I(1) process, so we use first differences $\{\tilde{p}_t - \tilde{p}_{t-1}\}_{t=1,2,\ldots,n}$. Nonstationarity of the levels series creates some mild complications in terms of the simulation strategy. What we do, in effect, is simulate the daily price movement process

$$p_s - p_t \qquad s \in [t, t+1)$$

for integer t while preserving the trajectories of U_{2t} and U_{3t} across the days. We use a weak order 2 scheme (Kloeden & Platen, 1992) to generate simulated numerical approximations to the stochastic differential equations.

With the continuous-time $\{p_s\}$ process thus generated, we take

$$\Delta p_t = p_t - p_{t-1}$$
$$d_t = \max_{t-1<s\le t} \{p_s - p_{t-1}\} - \min_{t-1<s\le t} \{p_s - p_{t-1}\} \qquad (5)$$

for $t = 1, 2, \ldots$. The extremes are taken over the half-open interval $(t - 1, t]$ to be consistent with the way data are recorded; in the theory, the continuous extension is used if needed to make the extremes well defined.

III.   Efficient Method of Moments

A.   Details

Let $\{y_t\}_{t=-\infty}^{\infty}$, $y_t \in \Re^M$, be a discrete stationary time series. In this paper, $\{y_t\}$ is either $\{\Delta p_t\}$, $\{d_t\}$, or $\{\Delta p_t, d_t\}$. When, as here, $\{y_t\}$ comes from a discretely sampled SDE system, then the SDE specification implicitly determines the density $p(y_{t-L}, \ldots, y_t | \rho)$ of a contiguous stretch of length $L + 1$ from $\{y_t\}$, where $\rho \in \Re^{p_\rho}$ is a vector of unknown system parameters, namely, the α's and β's of either (3) or (4). The fundamental problem that blocks straightforward application of standard statistical methods is that an analytic expression for $p(y_{t-L}, \ldots, y_0 | \rho)$ is not available so that, e.g., maximum-likelihood or Bayesian estimation is infeasible. However, by using simulation, an expectation of the form

$$\mathscr{E}_\rho(g) = \int \cdots \int g(y_{-L}, \ldots, y_0)$$
$$\times p(y_{-L}, \ldots, y_0 | \rho) dy_{-L} \cdots dy_0$$

can be computed for given ρ. That is, for given ρ, one can generate a simulation $\{\hat{y}_t\}_{t=-L}^{N}$ from the system and put

$$\mathscr{E}_\rho(g) = \frac{1}{N} \sum_{t=0}^{N} g(\hat{y}_{t-L}, \ldots, \hat{y}_t),$$

with N large enough that Monte Carlo error is negligible.

Gallant and Tauchen (1996a) proposed a minimum *chi*-squared estimator for ρ in this situation, which they termed the *efficient method of moments* (EMM) estimator. Being minimum *chi*-squared, the optimized *chi*-square criterion can be used to test system adequacy. The moment equations that enter the minimum *chi*-squared criterion of the EMM estimator are obtained from the score vector $(\partial/\partial\theta) \log f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ of an auxiliary model $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ that is termed the *score generator*. Gallant and Long (1997) show that, if the score generator is the SNP density $f_K(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ described below, then the efficiency of the EMM estimator can be made as close to that of maximum likelihood as desired by taking K large enough.

The first step in computing the EMM estimator $\hat{\rho}_n$ is to use the score generator

$$f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta) \qquad \theta \in \Theta \qquad (6)$$

to summarize the data $\{\tilde{y}_t\}_{t=-L}^{n}$ by computing the quasi-maximum-likelihood estimate

$$\tilde{\theta}_n = \operatorname*{argmax}_{\theta\in\Theta} \frac{1}{n} \sum_{t=0}^{n} \log [f(\tilde{y}_t | \tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \theta)],$$

and the corresponding estimate of the information matrix

$$\tilde{\mathscr{I}}_n = \frac{1}{n} \sum_{t=0}^{n} \left[\frac{\partial}{\partial\theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right] \qquad (7)$$
$$\times \left[\frac{\partial}{\partial\theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right]',$$

where $x_{t-1} = (y_{t-L}, \ldots, y_{t-1})$. The estimator (7) presumes the score generator (6) provides an adequate statistical approximation to the transition density of the data, so that $\{(\partial/\partial\theta) \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n)\}$ is essentially serially uncorrelated. If (6) is not adequate, then one of the more complicated expressions for $\mathscr{I}_n$ set forth in Gallant and Tauchen (1996a) must be used, although the EMM estimator is still consistent and asymptotically normal.

Define

$$m(\rho, \theta) = \mathscr{E}_\rho \left\{ \frac{\partial}{\partial\theta} \log [f(y_0 | y_{-L}, \ldots, y_{-1}, \theta)] \right\}$$

which is computed by averaging over a long simulation

$$m(\rho, \theta) \doteq \frac{1}{N} \sum_{t=0}^{N} \frac{\partial}{\partial\theta} \log [f(\tilde{y}_t | \tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \theta)].$$

The EMM estimator is

$$\hat{\rho}_n = \operatorname*{argmin}_{\rho\in\Re^{p_\rho}} m'(\rho, \tilde{\theta}_n)(\tilde{\mathscr{I}}_n)^{-1} m(\rho, \tilde{\theta}_n).$$

The estimator is consistent and asymptotically normally distributed with asymptotic distribution given in Gallant and Tauchen (1996a). Under the null hypothesis that $p(y_{-L}, \ldots, y_0 | \rho)$ is the correct model, n times the minimized value of the objective function is asymptotically *chi*-squared on $p_\theta - p_\rho$ degrees of freedom.

B.   Projection

One way to ensure that the selected score generator is a good approximation is to use a sieve estimator based on a nested sequence of finite dimensional parametric models

$f_K(y|x, \theta)$, $K = 1, 2, \ldots$. As noted above, the quasi-maximum-likelihood estimate $\hat{\theta}_n$ of the parameter vector θ in $f_K(y|x, \theta)$ summarizes the data; the larger is K, the more detailed the summary. This data reduction is termed the *projection step* of the EMM estimation procedure. The specific sieve that we employ is a modified Hermite series expansion termed *SNP*. The basic idea is that, because the square root of a density is an L_2 function and the Hermite functions are dense in L_2, the square of a Hermite expansion can closely approximate any density. Theoretical details are in Gallant and Nychka (1987).

Recall that, under our assumptions, the stationary distribution of a contiguous subsequence $y_{t-L}, \ldots, y_t$ has a density $p(y_{-L}, \ldots, y_0)$ defined over $\Re^l$, where $l = M(L + 1)$. Put $y = y_0$, $x = x_{-1} = (y_{-L}, \ldots, y_{-1})$, and write the stationary, marginal, and conditional densities as $p(x, y) = p(y_{-L}, \ldots, y_0)$, $p(x) = \int p(y_{-L}, \ldots, y_0) \, dy_0$, and $p(y|x) = p(x, y)/p(x)$, respectively. Let $\{\tilde{y}_t\}_{t=1}^n$ denote the realization from the process $\{y_t\}_{t=-\infty}^\infty$ that is available for analysis.

A standard method of describing a conditional density $f(y|x, \theta)$ is to set forth a location function μ_x and a scale function R_x that reduce the process $\{y_t\}_{t=-\infty}^\infty$ to an innovation process $\{z_t\}_{t=-\infty}^\infty$ via the transformation

$$z_t = R_{x_{t-1}}^{-1}(y_t - \mu_{x_{t-1}}).$$

The description is completed by setting forth a conditional density $h(z|x)$ for the innovation process. We follow this recipe in describing the sequence of SNP densities $f_K(y|x, \theta)$, where $K = 1, 2, \ldots$

The location function μ_x is given by

$$\mu_{x_{t-1}} = b_0 + Bx_{t-1}. \tag{8}$$

It is presumed to depend on $L_u \leq L$ lags which is accomplished by putting leading columns of B to zero as required.

The scale function R_x is given by

$$\text{vech} \, (R_{x_{t-1}}) = \rho_0 + P|e_{t-1}^*| \tag{9}$$

where vech (R) denotes a vector of length $M(M + 1)/2$ containing the elements of the upper triangle of R, $e_{t-1}^* = [(y_{t-L_r} - \mu_{x_{t-L_r-1}}), \ldots, (y_{t-1} - \mu_{x_{t-2}})]$, and $|\cdot|$ denotes elementwise absolute value with the corner at zero smoothed slightly to permit differentiation. The scale function depends on L_r lagged (unnormalized) innovations $(y_t - \mu_{x_{t-1}})$ and $L_r + L_u \leq L$ lagged y_t in total.

For α with nonnegative integer elements, let $z^\alpha = z_1^{\alpha_1} \cdots z_M^{\alpha_M}$ and $|\alpha| = \sum_{k=1}^M \alpha_k$; similarly for x^β. Consider the density

$$h_K(z|x) = \frac{[P_K(z, x)]^2 \phi(z)}{\int [P_K(u, x)]^2 \phi(u) \, du} \tag{10}$$

formed from the polynomial

$$P_K(z, x) = \sum_{\alpha=0}^{K_z} \left(\sum_{\beta=0}^{K_x} a_{\beta\alpha} x^\beta \right) z^\alpha$$

where $\phi(z) = (2\pi)^{-M/2} e^{-z'z/2}$. In equation (10), the term $P_K(z, x)\sqrt{\phi(z)}$ is a Hermite function of degree K_z in z whose coefficients are polynomials of degree K_x in x, which, as noted above, approximates the square root of a density and therefore enters (10) as the square. The shape of the innovation density $h_K(z_t|x_{t-1})$ varies with x_{t-1} which permits $h_K(z_t|x_{t-1})$ to exhibit general conditional shape heterogeneity. By putting selected elements of the matrix $A = [a_{\beta\alpha}]$ to zero, $P_K(z, x)$ can be made to depend on only $L_p \leq L$ lags from x. In applications where M is large, the coefficients $a_{\beta\alpha}$ corresponding to monomials z^α that represent high-order interactions can be set to zero with little effect on the adequacy of approximations. Let $I_z = 0$ indicate that no interaction coefficients are set to zero, $I_z = 1$ indicate that coefficients corresponding to interactions z^α of order larger than $K_z - 1$ are set to zero, and so on; similarly for x^β and I_x.

The change of variables $y_t = R_{x_{t-1}} z_t + \mu_{x_{t-1}}$ to obtain the density

$$f_K(y_t|x_{t-1}, \theta)$$
$$= \frac{\{P_K[R_{x_{t-1}}^{-1}(y_t - \mu_{x_{t-1}}), x_{t-1}]\}^2 \phi[R_{x_{t-1}}^{-1}(y_t - \mu_{x_{t-1}})]}{|\det (R_{x_{t-1}})| \int [P_K(u, x_{t-1})]^2 \phi(u) \, du} \tag{11}$$

completes the description of the SNP density. The vector θ contains the coefficients $A = [a_{\beta\alpha}]$ of the Hermite function, the coefficients $[b_0, B]$ of the location function, and the coefficients $[\rho_0, P]$ of the scale function. To achieve identification, the coefficient $a_{0,0}$ is set to 1. The tuning parameters are L_u, L_r, L_p, K_z, I_z, K_x, and I_x, which determine K and the dimension p_K of θ.

Some characteristics of $f_K(y_t|x_{t-1}, \theta)$ may be noted. If K_z, K_x, and L_r are put to 0, then $f_K(y_t|x_{t-1}, \theta)$ defines a Gaussian vector autoregression. If K_x and L_r are put to 0, then $f_K(y_t|x_{t-1}, \theta)$ defines a non-Gaussian vector autoregression model with homogeneous innovations. If K_z and K_x are put to 0, then $f_K(y_t|x_{t-1}, \theta)$ defines a Gaussian ARCH model. If K_x is put to 0, then $f_K(y_t|x_{t-1}, \theta)$ defines a non-Gaussian ARCH model with homogeneous innovations. If $K_z > 0$, $K_x > 0$, $L_p > 0$, $L_u > 0$, and $L_r > 0$, then $f_K(y_t|x_{t-1}, \theta)$ defines general nonlinear process with heterogeneous innovations.

How best to select the tuning parameters L_u, L_r, L_p, K_z, I_z, K_x, and I_x is an open question. A strategy found to work well is to move upward along an expansion path using the BIC criterion

$$\text{BIC} = s_n(\tilde{\theta}) + (1/2)(p_K/n) \log (n),$$

$$s_n(\theta) = -\frac{1}{n} \sum_{t=0}^n \log [f_K(\tilde{y}_t|\tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \theta)],$$

FIGURE 1—TOP: IBM DAILY LOG PRICE MOVEMENTS, $\{\Delta p_T\}$; BOTTOM: DAILY HIGH/LOW SPREAD $\{d_t\}$.
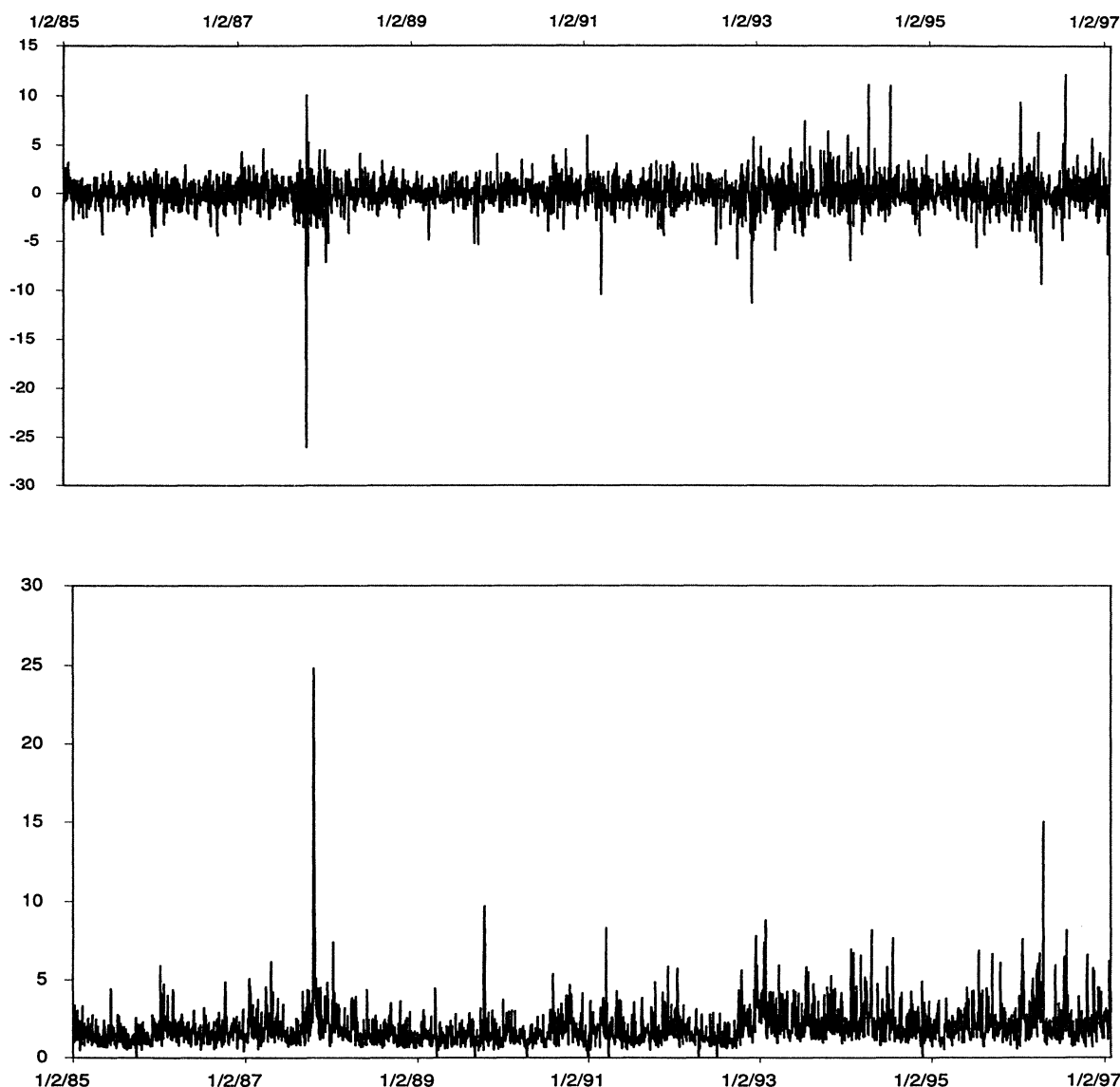
(Schwarz, 1978) to guide the search, models with small values of BIC being preferred.

The expansion path has a tree structure. Rather than examining the full tree, the strategy is to expand first in L_u with $L_r = L_p = K_z = K_x = 0$ until BIC turns upward. Next, expand L_r with $L_p = K_z = K_x = 0$, then expand K_z with $K_x = 0$, and lastly L_p and K_x. It is useful to expand in K_z, L_p, and K_x at a few intermediate values of L_r because it sometimes happens that the smallest value of BIC lies elsewhere within the tree.

It is essential that $f_K(y_t|x_{t-1}, \tilde{\theta}_n)$ not represent an explosive process, as discussed in detail by Tauchen (1998). Following Gallant and Tauchen (1997b), we use a transformation of each component of x_{t-1} by a logarithmic spline that substantially enhances numerical stability and has a negligible effect on either evaluation of $f_K(y_t|x_{t-1}, \tilde{\theta}_n)$ over the data or on the value computed for $\tilde{\theta}_n$.

IV. Data

We estimate the diffusion models described above using daily data on IBM from January 1, 1985, to January 31, 1997, for 3,152 observations before differencing. The top panel of figure 1 shows the close-close series $\{\Delta p_t\}$ series; the bottom panel shows the daily high-low range series $\{d_t\}$. We concentrate on a single series because we expect users to employ the methodology for pricing put/call options on individual stocks, which are widely traded. Furthermore, the so-called "leverage effect" (Nelson, 1991)—which is an asymmetric relationship between price movements and volatility—is much less pronounced in individual stock price series than it is in broad stock price index series. Kim and Kon (1994) present direct evidence on the very weak leverage effect for most of the component stocks of the DOW. Likewise, Tauchen et al. (1996) find no evidence for a

leverage effect in four stock series, one of which is IBM for a time period that overlaps with the data here. The absence of a leverage effect is important because the Hull/White formulae presume independence between the driving Brownian motions of the stock series and its volatility.

V. Estimation of the Score Generator (Projection Step)

To implement the EMM estimator, we require a score generator that adequately describes the dynamics of the observed discrete time series. For reasons discussed in Section III.B, we use the SNP model. To determine an appropriate SNP specification, we follow the protocol described in detail in Gallant and Tauchen (1998). We use the BIC to move upwards along the upward expansion path until an initial tentative model is identified. Because BIC is known to be conservative and therefore sometimes select an overly parsimonious model, this model is then subjected to a battery of diagnostics to see if further expansion is warranted. The diagnostic on the conditional mean is a regression of each of the standardized residuals

$$\hat{z}_t = \text{diag} \, [\hat{\Sigma}_{t-1}(y_t)]^{-1/2}[y_t - \hat{\mu}_{t-1}(y_t)]$$

on a constant and the unique elements of $\{y_{t-k}, y_{t-k} \otimes y_{t-k}, y_{t-k} \otimes y_{t-k} \otimes y_{t-k}\}_{k=1}^5$, where diag $[\hat{\Sigma}_{t-1}(y_t)]$ is the diagonal elements from the estimated conditional variance and $\hat{\mu}_{t-1}(y_t)$ is the estimated conditional mean, both of which conditional on x_{t-1}. The diagnostic on the conditional variance is taken from the same regression, except that the dependent variable is the squared standardized residuals. This test provides power against general nonlinear misspecification of either the conditional mean or the conditional variance function.

For the univariate close-to-close data, $y_t = \Delta p_t$, the initial candidate models selected by the BIC criterion is $L_u = 1$, $L_r = 6$, $L_p = 0$, $K_z = 4$, $K_x = 0$. For this SNP specification, the adjusted R^2's in the diagnostic regressions in mean as well as variance are less than 1%, so no further expansion is made.

For the univariate range series, $y_t = d_t$, the initial candidate model specification is $L_u = 6, L_r = 6, L_p = 0, K_z = 9, K_x = 0$. All these SNP specifications pass the diagnostic tests with the adjusted R^2's less than 1%, so again we do not expand the model further.

In the case of the bivariate series $y_t = (\Delta p_t, d_t)$, the initial candidate SNP specification selected by the BIC criterion is $L_u = 5, L_r = 6, L_p = 0, K_z = 9, I_z = 4, K_x = 0$. These SNP specifications also pass the diagnostic tests with the adjusted R^2's less than 1%.

In the following, we use these data-determined score generators to implement EMM estimation of various stochastic volatility diffusion models of financial price movements.

VI. Estimation of the SDEs

We employ the EMM method to estimate specifications of the stochastic volatility models using the univariate close-to-

TABLE 1.—OBJECTIVE FUNCTION VALUES: TWO-FACTOR MODEL

	α_{10}	α_{22}	β_{10}	β_{20}	β_{22}	N	$\chi^2(\hat{\rho})$	df	p-value
close to close, $\{\Delta p_t\}$									
	*	*	*	*		50k	23.156	9	0.006
	*	*	*	*		75k	23.013	9	0.006
	*	*	*	*	*	50k	19.640	8	0.012
	*	*	*	*	*	75k	18.415	8	0.018
spread, $\{d_t\}$									
	*	*	*	*		50k	387.406	19	0.000
	*	*	*	*		75k	392.206	19	0.000
	*	*	*	*	*	50k	44.138	18	5.5e-4
	*	*	*	*	*	75k	44.860	18	4.3e-4
bivariate, $\{\Delta p_t, d_t\}$									
	*	*	*	*	*	50k	147.935	60	2.2e-9
	*	*	*	*	*	75k	145.754	60	4.3e-9

Note: * denotes a freely estimated parameter.

close, the univariate range, and the bivariate score generators. In the computations, we scale time t such that the interval $(t, t + 1]$ is one day and the simulation step size is $\Delta = 1/24$. Throughout we report the results based on simulation size $N = 50,000$ and $N = 75,000$.

Table 1 shows the value of the EMM objective function under various specifications of the two-factor model (3). As seen from the top portion of the table, the specification with $\beta_{22} = 0$ fails to fit the scores defined by the score generator fitted to the $\{\Delta p_t\}$ series alone. Relaxing the restriction $\beta_{22} = 0$ provides an acceptable fit. Table 2 shows the parameter estimates, Wald-type standard deviations, and 95% confidence intervals obtained by inverting criterion difference test based on the concentrated (profile) objective function using an approximation procedure described in Gallant and Tauchen (1997a, 1997c). Such confidence intervals inherit the invariance properties of criterion difference test, and, unlike intervals based on the Wald test, they can be quite asymmetric when the objective function surface is quite asymmetric in that particular parameter. As seen from the top portion of table 2, the point estimates of the two-factor model fitted to the $\{\Delta p_t\}$ series appear reasonable, with all parameters save α_{10} quite statistically significant. The finding that a two-factor stochastic volatility diffusion model can adequately describe the marginal dynamics of a price movement series $\{\Delta p_t\}$ alone is consistent with the findings of Melino and Turnbull (1990), Engle and Lee (1996), Gallant and Long (1997), among others.

However, as seen from the middle and bottom portions of table 1, the two-factor model has considerable problems accounting for either the marginal dynamics of the range series $\{d_t\}$ or the joint dynamics of $\{\Delta p_t, d_t\}$. Also, in table 2, the point estimates are different depending upon whether the score generator for the marginal $\{d_t\}$ series or the joint $\{\Delta p_t, d_t\}$ series is used. Taken all together, the evidence in tables 1 and 2 suggest that the success in fitting the marginal dynamics of the price movement series is misleading, and, in fact, the two-factor model misses important aspects of the price dynamics.

TABLE 2.—PARAMETER ESTIMATES: TWO-FACTOR MODELS

	α_{10}	α_{22}	β_{10}	β_{20}	β_{22}
close to close, $\{\Delta p_t\}$					
	0.017	−0.156	1.258	0.181	0.180
S Dev	0.022	0.052	0.039	0.037	0.031
95% lower	−0.019	−0.239	1.185	0.099	0.115
95% upper	0.061	−0.066	1.339	0.239	0.246
spread, $\{d_t\}$					
	1.108	−0.100	0.557	0.883	−0.773
S Dev	0.068	0.096	0.480	0.869	0.258
95% lower	1.108	−0.132	0.557	0.732	−0.773
95% upper	1.146	−0.100	0.649	0.883	−0.727
bivariate, $\{\Delta p_t, d_t\}$					
	−0.002	−0.164	1.237	0.171	0.085
S Dev	0.016	0.021	0.012	0.012	0.010
95% lower	−0.027	−0.227	1.214	0.156	0.067
95% upper	0.024	−0.138	1.260	0.202	0.102

Notes: (1) Simulation size is 75,000.
(2) S. Dev. denotes Wald standard deviation.
(3) 95% lower and 95% upper denote boundaries of 95% confidence interval obtained by inverting the criterion difference test.

Table 3 shows the values of the objective function for the three-factor model (4) fitted to the scores for the bivariate $\{\Delta p_t, d_t\}$ series; Table 4 shows the parameter estimates. The three-factor model is formally rejected at conventional significance levels, so perhaps one should view the fitted model as a calibrated model. One should also keep in mind the recent Monte Carlo evidence (Chumacero, 1997; Andersen et al., 1998; Zhou, 1998) that EMM generally provides quite reliable parameter estimates but, like other minimum chi-squared estimators, the omnibus χ^2 test can be biased towards rejection.

The parameter estimates shown in table 4 are reasonable and generally statistically significant. The most-interesting parameter estimates are those of α_{22} and α_{33}. The two components in the volatility process have very different dynamics: The factor U_{2t} is strongly mean reverting while the factor U_{3t} is extremely persistent. The existence of long-run and short-run volatility components has been proposed by Andersen and Bollerslev (1997), Ding and Granger (1996), Engle and Lee (1993), and Liesenfeld

(1997). This characteristic of financial data volatility shows up directly in our estimation of the three-factor diffusion model.

Figure 2 shows the autocorrelation function of $|\Delta \tilde{p}_t - \Delta \tilde{p}_t^*|$ for the observed series $\{\Delta \tilde{p}_t\}_{t=2}^{3152}$, where $\Delta \tilde{p}_t^*$ is the predicted value from an AR(1) regression. The contribution of $\Delta \tilde{p}_t^*$ is essentially negligible apart from the constant, but it is used to ensure appropriate centering of the price movement series before taking the absolute value. The autocorrelation function of the absolute centered price movement damps sharply at the first few lags and then very slowly drifts towards zero at the longer lags. This pattern is entirely consistent with that of the figures in Ding et al. (1993), who attribute the shape to long memory in volatility.

The top panel of figure 3 shows the exact analogue of figure 2 as predicted by the two-factor model at the point estimates. Specifically, the top panel shows the autocorrelation function of $|\Delta \tilde{p}_t - \Delta \tilde{p}_t^*|$ computed on a simulated series $\{\Delta \tilde{p}_t\}_{t=2}^{75000}$. Evidently, the two-factor model cannot mimic the characteristic shape of the autocorrelation function of the absolute centered price movement. Likewise, the bottom panel of figure 3 shows, along with a reference GARCH, the autocorrelation function of end-of-day volatility evaluated on the simulated trajectory of length 75,000. Evidently, the two-factor model is more GARCH-like in appearance and does not capture the appearance of long memory.

Figure 4 contains plots of the same quantities as in figure 3 except the computations are done on a simulated trajectory of length 75,000 from the three-factor model. The three-factor model can mimic the appearance of long memory. We emphasize that the three-factor model is Markov, not long-memory model, and it is considerably easier to manipulate and simulate. Nonetheless, its ability to mimic long memory is striking.

Figure 5 indicates why the three-factor model with two stochastic AR(1) components mimics so much better the appearance of long memory in volatility. The top panel shows the spectral density of a fractionally integrated

TABLE 3.—OBJECTIVE FUNCTION: THREE-FACTOR MODELS FITTED TO IBM BIVARIATE SCORE GENERATOR

	α_{10}	α_{22}	α_{33}	β_{10}	β_{20}	β_{22}	β_{30}	β_{33}	N	$\chi^2 (\hat{\rho})$	df	p-value
bivariate, $\{\Delta p_t, d_t\}$												
	*	*	*	*	*	*	*	*	50k	124.725	57	5.78e-7
	*	*	*	*	*	*	*	*	75k	126.454	57	3.52e-7

Note: * denotes a freely estimated parameter.

TABLE 4.—PARAMETER ESTIMATES: THREE-FACTOR MODEL

	α_{10}	α_{22}	α_{33}	β_{10}	β_{20}	β_{22}	β_{30}	β_{33}
bivariate $\{\Delta p_t, d_t\}$								
	−0.007	−0.231	−0.268e-3	1.383	0.184	0.142	0.0076	0.026
S Dev	0.019	0.024	0.275e-3	0.133	0.009	0.035	0.0021	0.006
95% lower	−0.010	−0.233	−0.269e-3	1.382	0.183	0.134	0.0076	0.026
95% upper	−0.001	−0.228	−0.259e-3	1.387	0.185	0.148	0.0077	0.027

Notes: (1) Simulation size is 75,000.
(2) S. Dev. denotes Wald standard deviation.
(3) 95% lower and 95% upper denote boundaries of 95% confidence interval obtained by inverting the criterion difference test.

FIGURE 2—SAMPLE AUTOCORRELATION FUNCTION OF CENTERED ABSOLUTE PRICE MOVEMENTS, $|\Delta\bar{p} - \Delta\bar{p}^*|$.
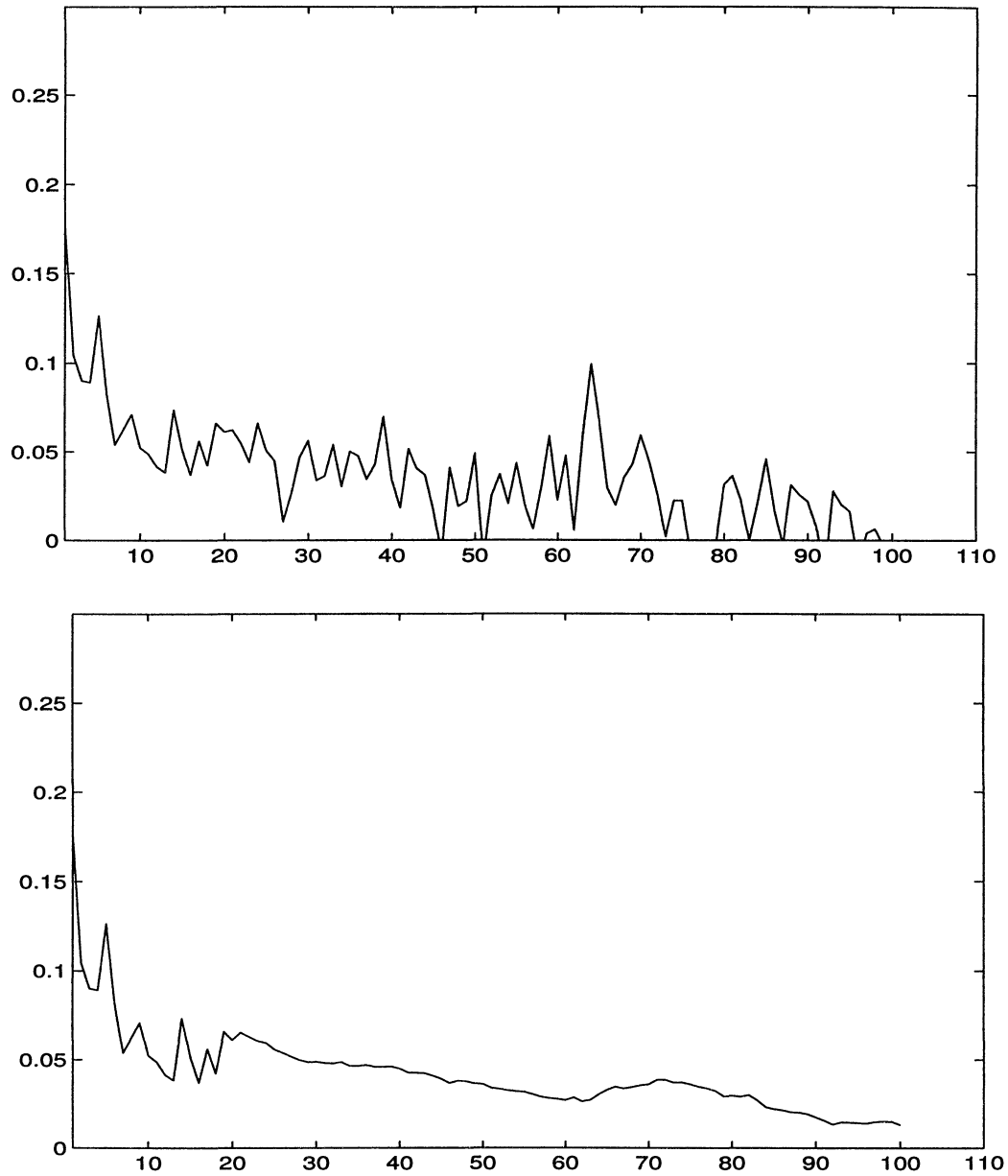
Top panel is the raw autocorrelegram function; bottom panel is with smoothing at longer lag lengths.

process with parameter $d = 0.40$ along with the best least-squares approximation from an AR(1) model. The AR(1) approximation is obtained via

$$\min_{a_1,\sigma_1} \sum_{j=1}^{100} \left\{ \left[\frac{1}{1 - \cos(\lambda_j)} \right]^{0.40} - \frac{\sigma_1^2}{1 + a_1^2 - 2a_1 \cos(\lambda_j)} \right\}^2$$

where $\lambda_j = (j/100)\pi$. Evidently, the AR(1) cannot match very well the spectral density of the fractionally integrated process over the entire domain. However, the lower panel shows the best least-squares approximation of the spectral density of the sum of two AR(1) processes to that of the

fractionally integrated process. The best approximation is obtained via solving

$$\min_{a_1,\sigma_1,a_2,\sigma_2} \sum_{j=1}^{100} \left\{ \left[\left| \frac{1}{1 - \cos(\lambda_j)} \right| \right]^{0.40} \right.$$
$$\left. - \frac{\sigma_1^2}{1 + a_1^2 - 2a_1 \cos(\lambda_j)} - \frac{\sigma_2^2}{1 + a_1^2 - 2a_1 \cos(\lambda_j)} \right\}^2.$$

As seen from the bottom panel of figure 5, the approximation is superb. One can even view the situation the other way around: Arguably, the truth is a two-component volatility

FIGURE 3—AUTOCORRELATION FUNCTIONS AT TWO-FACTOR MODEL.
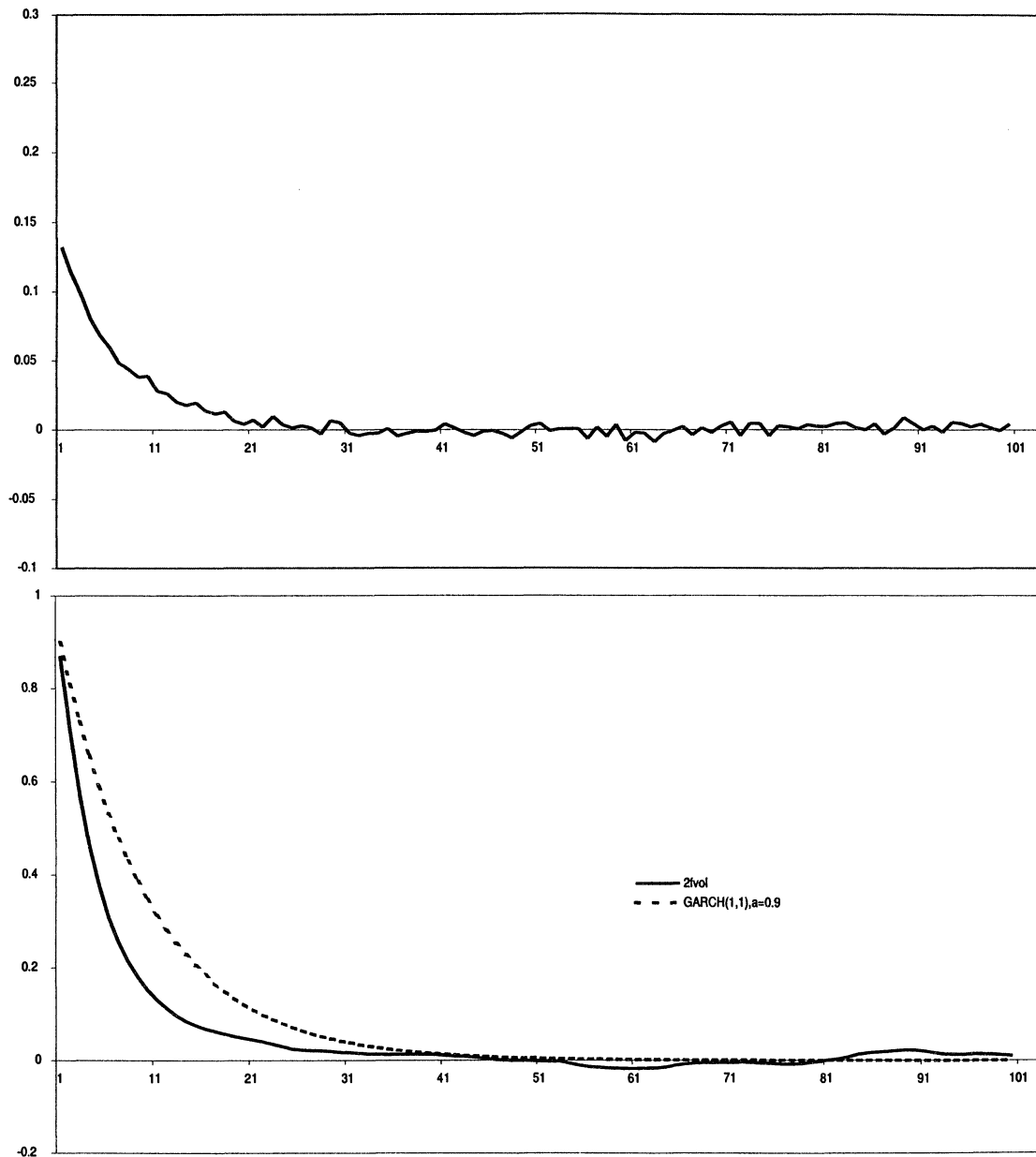
Top: Autocorrelation function of $|\Delta p - \overline{\Delta p}|$ implied by the two-factor model; bottom: autocorrelation function of end-of-day volatility implied by the two-factor model with a reference GARCH.

model that is well approximated by a fractionally integrated process.

VII.   Extracting the Integrated Volatility

A.   Background

Consider the price, π_t, at time t of a derivative asset that depends upon the value of a security price $S_{t+J} = e^{p_{t+J}}$, $J > 0$, and assume the dynamics of the process p are

$$dp_t = a_t dt + \sigma_{pt} dW_t, \tag{12}$$

where σ_{pt} represent stochastic volatility and W_t is Brownian motion. Assume the risk-neutral dynamics are

$$dp_t = \left(r - \frac{1}{2}\sigma_{pt}^2\right)dt + \sigma_{pt}dW_t \tag{13}$$

where r is the interest rate, assumed to be constant. Hull and White (1987) show that, under the assumption of no-risk premium on the volatility factor, the price π_t of the claim is given as

$$\pi_t = \mathscr{E}_t[\mathscr{B}(S_t, Jv_{t,t+J}, r, J)] \tag{14}$$

FIGURE 4—AUTOCORRELATION FINDINGS AT THREE-FACTOR MODEL
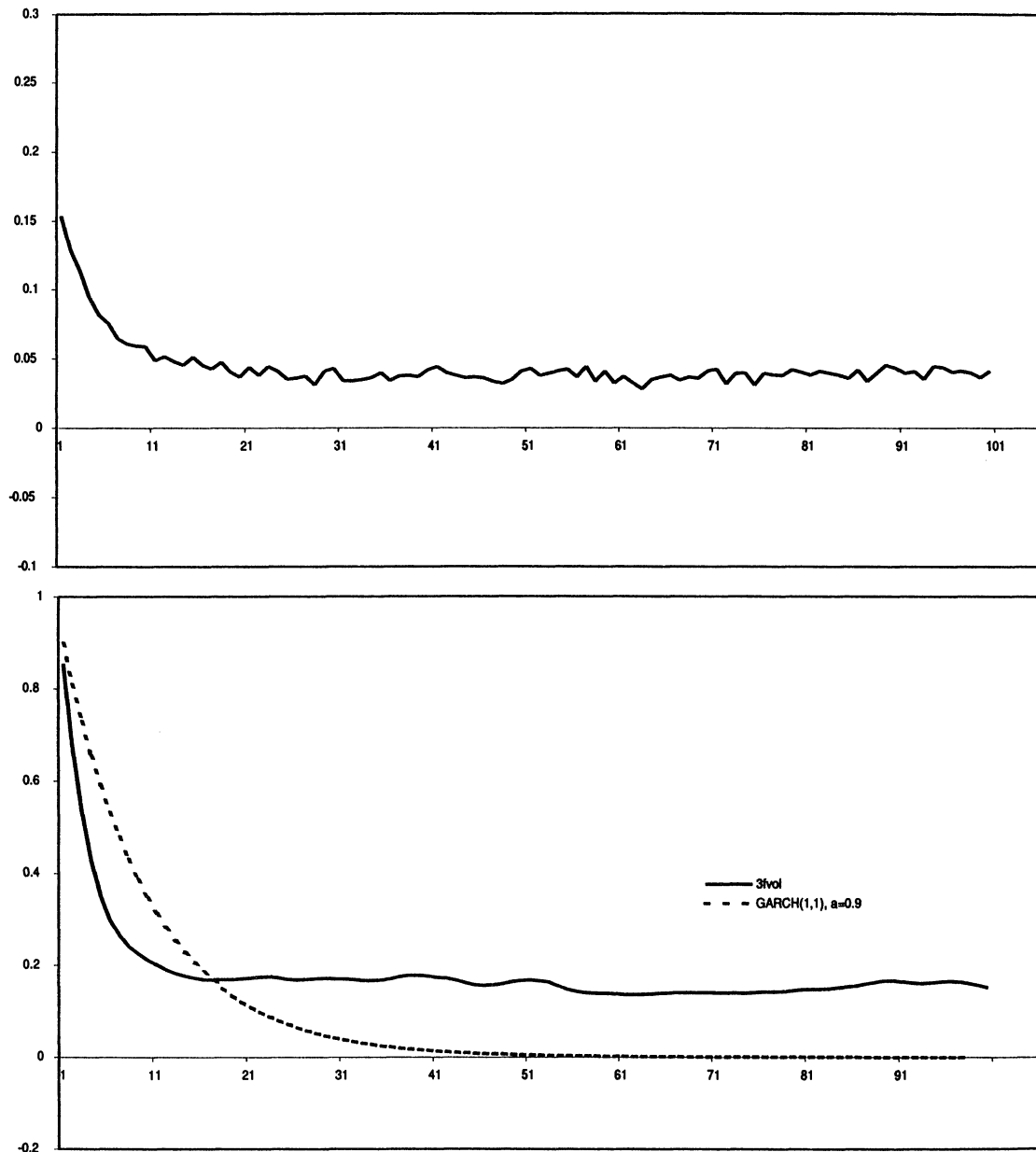
Top: Autocorrelation function of $|\Delta p - \Delta \bar{p}|$ implied by the three-factor model; bottom: autocorrelation function of end-of-day volatility implied by the three-factor model with a reference GARCH.

where $\mathcal{E}_t(\cdot)$ is the conditional expectations operator, $\mathcal{B}(\cdot)$ is the Black-Scholes price of the derivative claim, and

$$v_{t,t+J} = \frac{1}{J} \int_0^J \sigma^2_{p,t+s} \, ds \qquad (15)$$

is the mean integrated local variance over the life of the derivative. Note that $v_{t,t+J}$ is, in general, a nontrivial random variable at time t which is integrated out in equation (14).

Let $\mathcal{F}_t$ denote the information available to the analyst at time t. To be specific, in our empirical work, $\mathcal{F}_t$ is $\{p_{t-j}, d_{t-j}\}_{j=0}^{t-1}$. We extend the reprojection technique devel-

oped in Gallant and Tauchen (1998) in order to estimate $p(v_{t,t+J}|\mathcal{F}_t)$, the conditional density of $v_{t,t+J}$ given $\mathcal{F}_t$. This density can be used to compute the empirical analogue of equation (14):

$$\pi_t^* = \mathcal{E}[\mathcal{B}(S_t, Jv_{t,t+J}, r, J)|\mathcal{F}_t]$$
$$= \int \mathcal{B}(S_t, Jv_{t,t+J}, r, J) p(v_{t,t+J}|\mathcal{F}_t) \, dv_{t,t+J}. \qquad (16)$$

There is a subtle distinction between equation (14) and (16), as the former expectation is taken with respect to a larger information set. Among other things, this larger information set includes the unobservable stochastic local variance

FIGURE 5—SPECTRAL DENSITY APPROXIMATIONS

One Component Approximation to Long Memory with d=0.40

Two Component Approximation to Long Memory with d=0.40
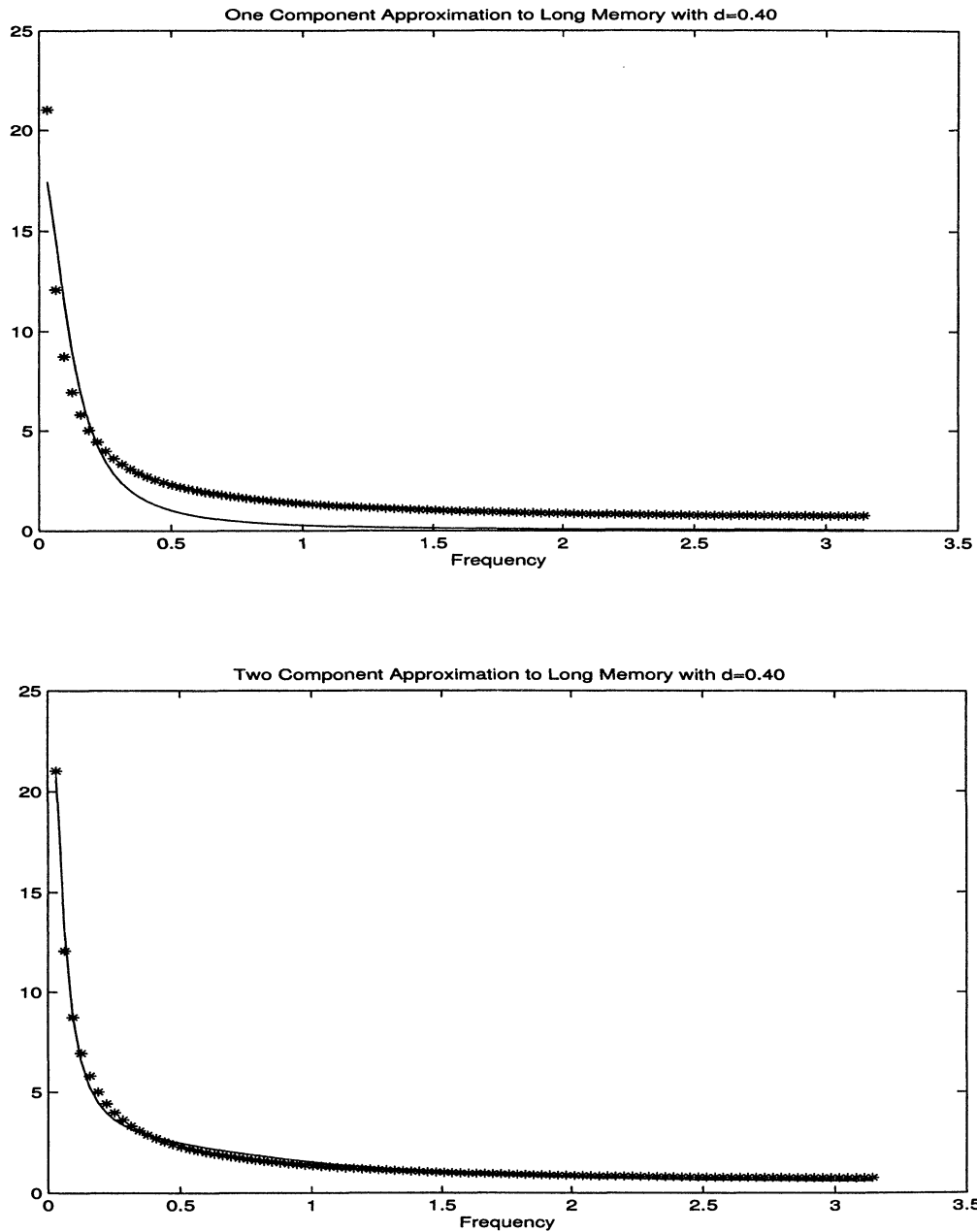
The top panel shows the best approximation of the spectral density of an AR(1) process (solid) to that of a fractionally integrated process (*) with parameter $d = 0.40$; the bottom panel shows the best approximation of the spectral density of the sum of two AR(1) processes (solid) to that of a fractionally integrated process (*) with parameter $d = 0.40$.

process σ_t^2. The econometric estimator of $p(v_{t,t+J}|\mathscr{F}_t)$ developed here is potentially useful for many purposes, including, for example, forecasting the course of the local variance

$$v_{t,t+J}^* = \mathscr{E}(v_{t,t+J}|\mathscr{F}_t)$$
$$= \int v_{t,t+J} p(v_{t,t+J}|\mathscr{F}_t) \, dv_{t,t+J} \qquad (17)$$

or the pricing calculations in equation (16).

As a simpler alternative yet, one might consider estimating

$$v_{t,t+J}^{**} = \mathscr{P}(v_{t,t+J}|\mathscr{F}_t) \qquad (18)$$

where $\mathscr{P}(\cdot|\cdot)$ denotes the linear projection operator, i.e., the linear regression operator. The linear projection of $v_{t,t+J}$ onto $\mathscr{F}_t$ provides an easy-to-compute forecast of the course of variance over the life of the derivative. Also, it might be used

in a pricing calculation like

$$\pi_t^{**} = \mathscr{B}(S_t, Jv_{t,t+J}^*, r, J) \tag{19}$$

as a practical compromise to either equation (14) or (16).

B. The Conditional Distribution of the Integrated Local Variance

The task is to estimate the conditional distribution of the mean integrated forward variance $v_{t,t+J} = (1/J) \int_0^J \sigma_{p,t+s}^2 \, ds$ conditional on $\{\Delta p_{t-j}, d_{t-j}\}_{j=0,1,\ldots}$. Given that we are using a simulation estimator, this is conceptually straightforward. The idea is to generate a long simulation from the estimated SDE model, nonparametrically estimate the conditional distribution on that simulation, and then evaluate the conditional distribution at the observed data.

The mechanics of the calculation are this: Let $\{\Delta \hat{p}_t, \hat{d}_t\}_{t=1}^N$ denote a realized simulated trajectory of $\Delta p_t, d_t$, and let $\{\hat{\sigma}_{p,t}^2\}_{t\in[0,N]}$ denote the corresponding continuous record of $\sigma_{p,t}^2$. The simulation is generated using the EMM estimates of the diffusion. These quantities are readily available from the simulation process used to implement the EMM estimator, except that the continuous record $\{\hat{\sigma}_{p,t}^2\}_{t\in[0,N]}$ is actually approximated by a very fine weak-2 discretization. Define

$$\hat{v}_{t,t+J} = \frac{1}{J} \int_0^J \hat{\sigma}_{p,t+s}^2 \, ds, \qquad t = 1, 2, \ldots, N - J,$$

which is the J-step ahead mean integrated forward variance, and let

$$\hat{x}_t = (\Delta \hat{p}_{t-L}, \hat{d}_{t-L}, \ldots, \Delta \hat{p}_t, \hat{d}_t).$$

The idea is to determine the conditional density $p(v_{t,t+J}|\mathscr{F}_t)$ by nonparametrically estimating within the simulation the conditional density of $p(\hat{v}_{t,t+J}|\hat{x}_t)$. To avoid clutter, at times we drop the t subscript and write $p(v_J|x)$ for the conditional density. To determine $p(v_J|x)$, we adapt the SNP density of Gallant and Tauchen (1989) as described above. The SNP estimate density estimate is $f_K(w|x, \hat{\theta})$, where $w = \log(v_J)$,

$$\hat{\theta} = \underset{\theta}{\mathrm{argmax}} \sum_{t=L+1}^{N-J} \log[f_K(\hat{w}_{t,t+J}|\hat{x}_t, \theta)],$$

$\hat{w}_{t,t+J} = \log(\hat{v}_{t,t+J})$, and $f_K(\cdot|\cdot, \theta)$ is defined in equation (11). As for the tuning parameters, we set $L_r = 0$, because, in this context, the explanatory variables are not lags of the dependent variable, and we set the others generously large in an effort to pin down the density with high numerical accuracy. (N is huge; we use $N = 75,000$.) Put

$$\hat{p}(v_J|x) = f_K[\log(v_J)|x, \hat{\theta}] \frac{1}{v_J}. \tag{20}$$

Under conditions given in Gallant and Tauchen (1989, 1998) and Gallant and Long (1997), $\hat{p}(v_J|x)$ is a consistent estimate of the conditional density of the mean integrated variance. Apart from pricing calculations such as equation (16), the conditional density (20) can be used directly for extracting from the observed data the latent mean integrated variance via the computation

$$\int v_J \hat{p}(v_J|\tilde{x}_t) dv_J, \qquad t = L + 1, \ldots, n, \tag{21}$$

where, it will be recalled that $\tilde{x}_t$ denotes the observed data value. To summarize, simulated data are used to determine $\hat{\theta}$ and thereby the functional form of $\hat{p}(v_J|x)$. Then, $\hat{p}(v_J|\tilde{x}_t)$ gives us the predictive density of $v_{t,t+J}$ at observation point t.

Figure 6 shows this variance-extraction calculation using the estimated two-factor model. The three panels show time series plots of the square roots of the conditional means of the integrated forward variance $(\int v_J \hat{p}(v_J|\tilde{x}_t) \, dv_J)^{1/2}$, for J corresponding to one day ahead, one month ahead, and one year ahead, with each integrated variance annualized to aid interpretation. Interestingly, the daily forward variance is quite active, as is the monthly, while the yearly is nearly constant.

Figure 7 shows this variance-extraction calculation using the estimated three-factor model. As is perhaps not surprising, both the two- and the three-factor models give nearly the same conditional means of the daily and monthly integrated forward variance. However, the two models differ substantially for the yearly integrated forward variance, with the three-factor model still showing the effects of shocks one year out. This contrast is directly attributable to the capability of the three-factor model to mimic the appearance of long memory in volatility.

The reprojection technique presented here is a model-based procedure that makes use of the daily price movement and range data. It is different from the direct measurement of realized volatility approach pioneered by French et al. (1987) and followed to the ultrahigh frequency by Andersen and Bollerslev (1998) and Andersen, Bollerslev, et al. (1998). Direct measurement relies on the fact that σ_{pt}^2 could, in principle, be extracted from continuous record $\{p_s\}_{s\in[0,t]}$. Doing so is not actually feasible in practice as the continuous record is not available. However, for a very large K—i.e., for small $\delta = 1/K$—one has

$$\sum_{k=1}^K (p_{t-1+k\delta} - p_{t-1+(k-1)\delta})^2 \approx \int_{t-1}^t \sigma_{ps}^2 \, ds.$$

Clearly, there is more information in ultrahigh-frequency data than there is in close-to-close and range data, so direct measurement produces a more accurate integrated variance measure, as documented by simulation in Andersen and

FIGURE 6—TWO-FACTOR MODEL: ROOT EXPECTED INTEGRATED FORWARD VARIANCE

Daily Expected Integrated Volatility

Monthly Expected Integrated Volatility

Yearly Expected Integrated Volatility
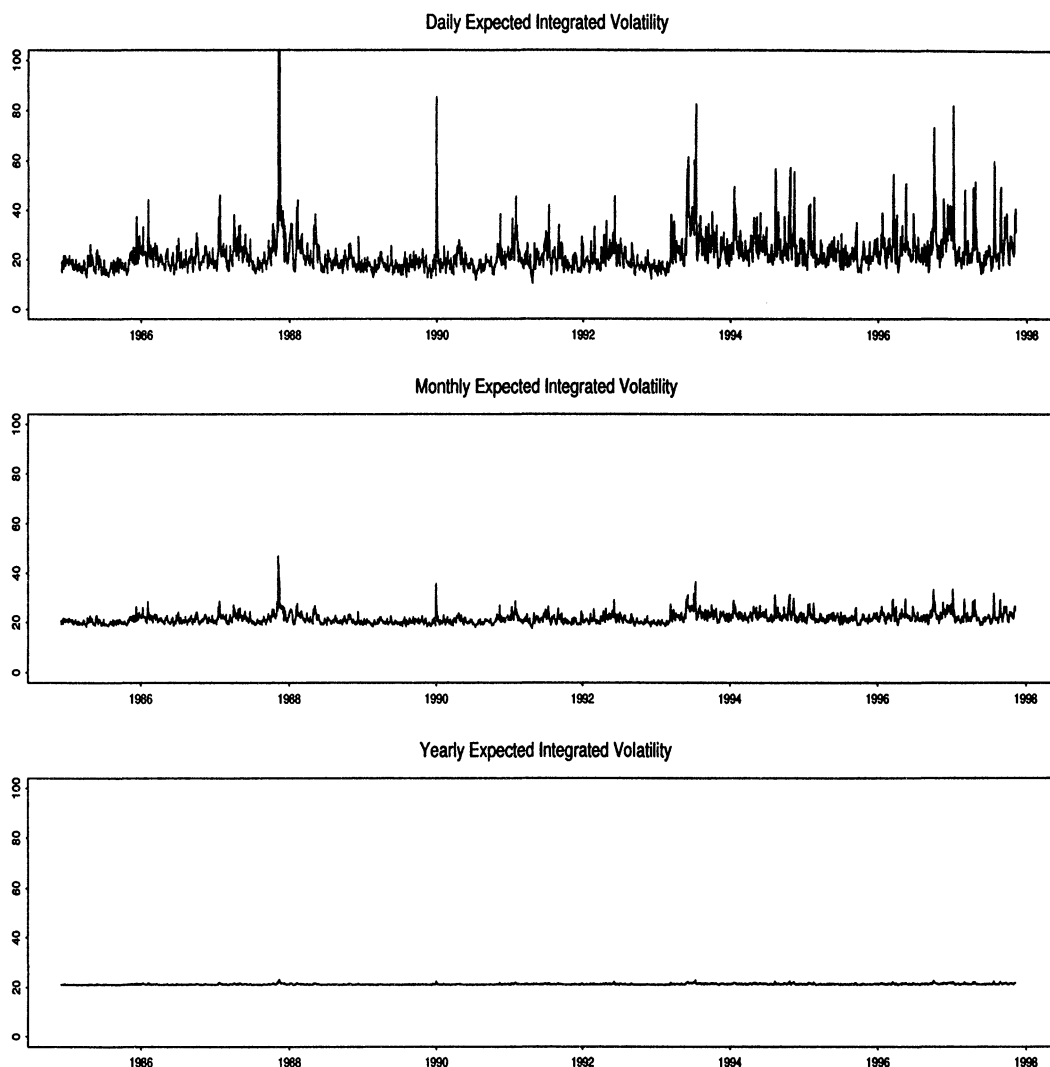
$(\int v_j \hat{p}(v_j | \bar{x}_t, \hat{\theta}) \, dv_j)^{1/2}$, one day, one month, and one year forward, annualized.

Bollerslev (1998). A potential drawback, however, is that direct measurement entails storage and manipulation of huge data sets, and it still leaves open the matter of obtaining the conditional distribution of the integrated variance $v_{t,t+J}$ which is needed to compute the expectation required to implement equation (14). On the other hand, our approach is potentially useful for understanding volatility dynamics using very long historical data sets in which the daily range has been recorded but the within-day tick data were discarded long ago. Finally, we emphasize the reprojection approach delivers a diffusion model's predicted integrated variance. The prediction can indeed disagree with actual integrated variance, if the model is misspecified. This situation is one the model's user certainly needs to know about, and figures like figure 7 would be useful for making such assessments.

VIII. Conclusion

Implementation of the Hull and White options-pricing formula requires both an estimate of the stochastic volatility model governing price dynamics and an estimate of the conditional distribution of the forward integrated variance given the observed variables. This paper addressed both of these requirements. Using daily stock price data on close-to-close log price movement and the high/low range, we found that the standard two-factor stochastic volatility models did not fit all aspects of the data very well and that a more general three-factor model did better. In particular, the three-factor model captured the extreme persistence in financial volatility attributed to fractional integration (long-memory) by Ding et al. (1993), Baillie et al. (1996), Bollerslev and Mikkelsen (1996), and others. We also

FIGURE 7—THREE-FACTOR MODEL: ROOT EXPECTED INTEGRATED VARIANCE

Daily Expected Integrated Volatility

Monthly Expected Integrated Volatility

Yearly Expected Integrated Volatility
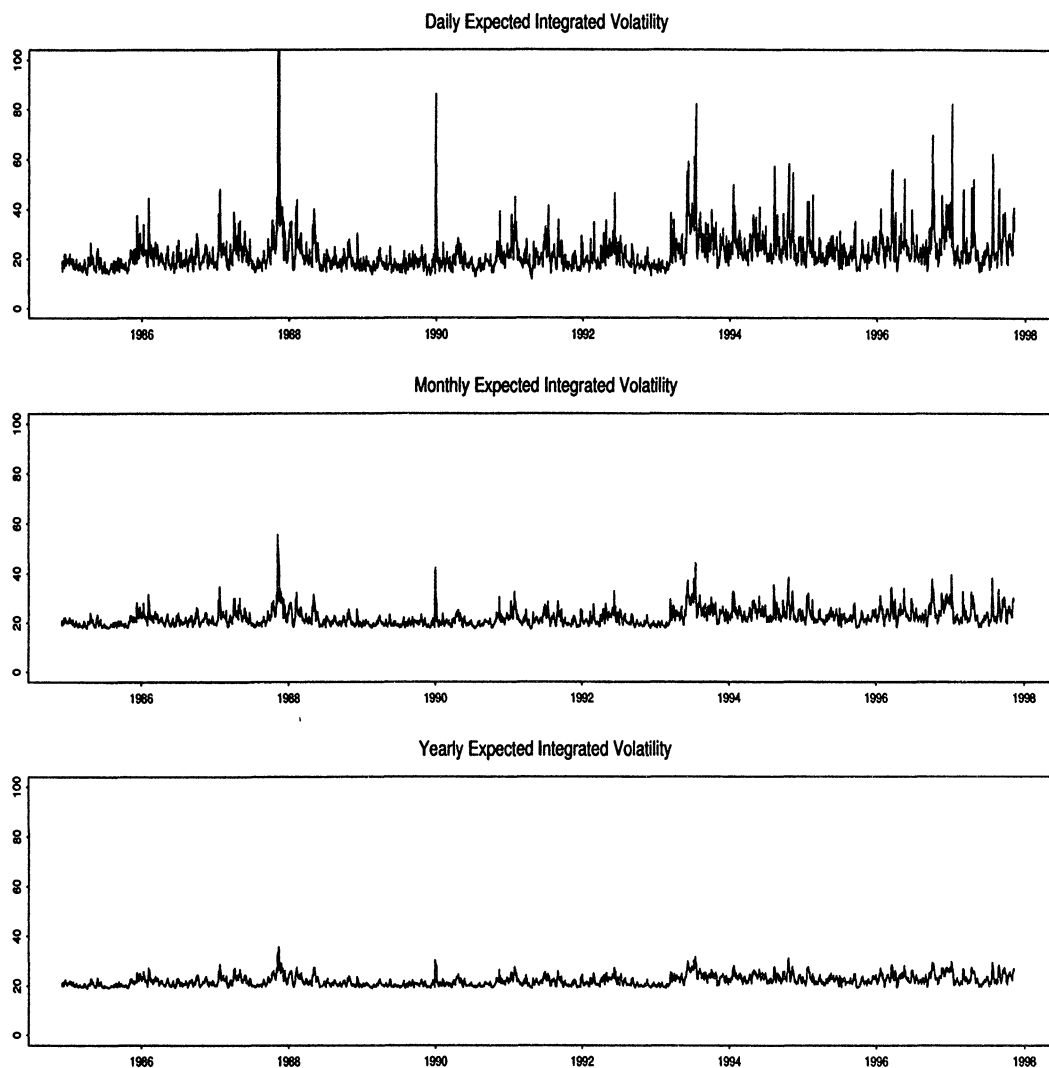
$(\int v_j \hat{p}(v_j | \tilde{x}_t, \hat{\theta}) \, dv_j)^{1/2}$, one day, one month, and one year forward, annualized.

adapted and extended the reprojection technique of Gallant and Tauchen (1998) to estimate the conditional distribution of the forward integrated variance given the observed variables.

REFERENCES

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

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

Andersen, T. G., H-J. Chung, and B. E. Sorensen, "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study," working paper (1998).

Andersen, T. G., T. Bollerslev, F. Diebold, and P. Labys, "The Distribution of Exchange Rate Volatility," NBER Working Paper 6861 (1998).

Andersen, T. G., and J. Lund, "Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate," *Journal of Econometrics* 77 (1997), 343–378.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen, "Fractionally Integrated Generalized Autoregressive Heteroskedasticity," *Journal of Econometrics* 73 (1996), 3–33.

Beckers, S., "Variances of Security Price Returns Based on High, Low, and Closing Prices," *Journal of Business* 56 (1983), 97–112.

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

Chumacero, R., "Finite Sample Properties of the Efficient Method of Moments," *Studies in Nonlinear Dynamics and Econometrics* 2 (1997), 35–51.

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

Ding, Z., and C. W. J. Granger, "Modeling Volatility Persistence of Speculative Returns: A New Approach," *Journal of Econometrics* 74 (1996) 185–215.

Drost, F. C., and T. E. Nijman, "Temporal Aggregation of GARCH Processes," *Econometrica* 61 (1993), 909–927.

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

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

Engle, R. F., and G. G. J. Lee, "A Permanent and Transitory Component Model of Stock Return Volatility," Discussion paper 92-44R, (1997) University of California at San Diego.

——— "Estimating Diffusion Models of Stochastic Volatility," in P. E. Rossi, (ed.), *Modeling Stock Market Volatility: Bridging the Gap to Continuous Time* (New York: Academic Press, 1996), 333–384.

French, K. R. G., W. Schwert, and R. F. Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics* 19 (1987), 3–29.

Gallant, A. R., D. A. Hsieh, and G. Tauchen, "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics* 81 (1997), 159–192.

Gallant, A. R., and J. R. Long, "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square," *Biometrika* 84 (1997), 125–141.

Gallant, A. R., and D. W. Nychka, "Seminonparametric Maximum Likelihood Estimation," *Econometrica* 55 (1987), 363–390.

Gallant, A. R., and G. Tauchen, "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica* 57 (1989), 1091–1120.

——— "Which Moments to Match?," *Econometric Theory* 12 (1996a), 657–681.

——— "Specification Analysis of Continuous Time Models in Finance," in P. E. Rossi, (ed.), *Modeling Stock Market Volatility: Bridging the Gap to Continuous Time* (New York: Academic Press, 1996b), 333–384.

——— "Estimation of Continuous Time Models for Stock Returns and Interest Rates," *Macroeconomic Dynamics* 1 (1997a), 135–168.

——— "SNP: A Program for Nonparametric Time Series Analysis: A User's Guide" (1997b), available via ftp at ftp.econ.duke.edu, subdirectory pub/arg/snp.

——— "EMM: A Program for Efficient Method of Moments Estimation," 1997c, available via ftp at ftp.econ.duke.edu, subdirectory pub/get/emm.

——— "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions," *Journal of the American Statistical Association* 93 (1998), 10–24.

Garman, M. B., and M. J. Klass, "On the Estimation of Security Price Volatilities from Historical Data," *Journal of Business* 53 (1980), 67–78.

Gourieroux, C., A. Monfort, and E. Renault, "Indirect Inference," *Journal of Applied Econometrics* 8 (1993), S85–S118.

Hansen, L. P., and J. A. Scheinkman, "Back to the Future: Generating Moment Implications for Continuous Time Markov Process," *Econometrica* 50 (1995), 1029–1054.

Hsieh, D. A., "Chaos and Nonlinear Dynamics: Application to Financial Markets," *Journal of Finance* 46 (1991), 307–317.

Hull, J., and A. White, "The Pricing of Options on Assets with Stochastic Volatility," *Journal of Finance* 42, (1987), 381–400.

Ingram, B. F., and B. S. Lee, "Simulation Estimation of Time Series Models," *Journal of Econometrics* 47 (1991). 197–250.

Karatzas, I., and S. E. Shreve, *Brownian Motion and Stochastic Calculus* (Berlin: Springer-Verlag, 1991).

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

Kloeden, P. E., and E. Platen, 1992, *Numerical Solution of Stochastic Differential Equations* (Berlin: Springer-Verlag, 1992).

Liesenfeld, R., "Trading Volume and the Short and Long-Run Components of Volatility," working paper, Eberhard-Karls University (1997).

Melino, A., and S. M. Turnbull, "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics* 45, (1990) 239–266.

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

——— "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," *Journal of Econometrics* 52 (1992), 61–90.

Parkinson, M., "The Extreme Value Method for Estimating the Variance of the Rate of Return," *Journal of Business* 53 (1980), 61–65.

Schwarz, G., "Estimating the Dimension of a Model," *Annals of Statistics* 6 (1978), 461–464.

Tauchen, G., "New Minimum Chi-Square Methods in Empirical Finance," in Kreps, David, and Kenneth Wallis (eds.), *Advances in Econometrics, Seventh World Congress,* (Cambridge, UK: Cambridge University Press, 1997), 279–317.

——— "The Objective Function of Simulation Estimators Near the Boundary of the Unstable Region of the Parameter Space," this REVIEW 80 (1998), 389–398.

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

Zhou, H., "Finite Sample Properties of Efficient Method of Moments and Maximum Likelihood Estimations for a Square-Root Diffusion Process," working paper, Duke University (1998).