Volume, volatility, and leverage:
A dynamic analysis

George Tauchen\textsuperscript{\ast, a}, Harold Zhang\textsuperscript{b}, Ming Liu\textsuperscript{a}

\textsuperscript{a} Department of Econometrics, Duke University, Durham, NC 27708, USA
\textsuperscript{b} Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

This paper uses dynamic impulse response analysis to investigate the interrelationships among stock price volatility, trading volume, and the leverage effect. Dynamic impulse response analysis is a technique for analyzing the multi-step-ahead characteristics of a nonparametric estimate of the one-step conditional density of a strictly stationary process. The technique is the generalization to a nonlinear process of Sims-style impulse response analysis for linear models. In this paper, we refine the technique and apply it to a long panel of daily observations on the price and trading volume of four stocks actively traded on the NYSE: Boeing, Coca-Cola, IBM, and MMM.

Key words: Dynamic impulse response; Financial time series; Nonlinear processes
JEL classification: G10

1. Introduction

Financial time series are widely understood to be nonlinear processes. Although there is evidence for some dependence in the conditional mean of asset returns, there is substantially more structure in the second and higher moments. A short list of documented features includes ARCH/GARCH (Bollerslev, Chou, and Kroner, 1992), leptokurtosis (Clark, 1973; Gallant, Hsieh, and Tauchen, 1991), various asymmetries (Nelson, 1991; Campbell et al., 1993; LeBaron, 1992, a, b),

\textsuperscript{\ast} Corresponding author.
The material is based upon work supported by the National Science Foundation under Grant No. SES-9022083. We thank Tim Bollerslev, A. Ronald Gallant, Peter Rossi, and seminar and conference participants for helpful comments at various stages of this research.
and volume–volatility relations (Tauchen and Pitts, 1983; Karpoff, 1987; Gallant, Rossi, and Tauchen, 1992; Lamoureux and Lastrapes, 1994). Much of the existing evidence on the nonlinear character of financial time series pertains to the one-step-ahead conditional distribution. Furthermore, this evidence is adduced from parametric models that incorporate a priori constraints on the shape and low-order moments of the conditional distribution.

In this paper, we investigate the nonlinear dynamics of daily price and volume movements, 1982–1987, for four stocks: Boeing (BA), International Business Machine (IBM), Coca-Cola (KO), and Minnesota Mining and Manufacturing (MMM). In contrast to earlier studies, we investigate multi-step dynamics and we use a nonparametric approach. The objectives are to examine in detail the persistence properties of stochastic volatility, the asymmetric character of the conditional variance function, and the characteristics of various nonlinear volume–price interactions. Unlike previous studies using aggregated price index and total volume, we use a cross-section of individual securities, which provides a means for assessing robustness of findings. Also, the use of disaggregated data avoids some logical difficulties with adding up volume across various securities, and it avoids some modeling issues as well. Models of trading volume that incorporate a market microstructure and latent information process (Tauchen and Pitts, 1983; Andersen, 1995; Foster and Viswanathan, 1995) do not aggregate well, which motivates testing volume–price relations using individual security data.

The investigation entails use and refinement of the strategy for performing nonlinear impulse response analysis described in Gallant, Rossi, and Tauchen (1993). Nonlinear impulse response analysis is the extension to the nonlinear case of methods proposed by Sims (1980) and refined by Doan, Litterman, and Sims (1984) and others. For a general nonlinear model, the dynamic properties can be elicited by perturbing the vector of conditioning arguments in the conditional density function. The dynamic response to the perturbation can be traced out by computing multi-step-ahead expectations of nonlinear functions such as conditional moments. Much of the focus of this paper is on the first two conditional moments, though in Section 6 we report on some work with the conditional third and fourth moments as well.

The approach is completely nonparametric in that it does not confine the conditional probability distribution of the financial time series to any particular parametric family. The conditional probability distribution of a strictly stationary, possibly nonlinear, process is completely summarized by its one-step-ahead conditional density function. From the nonparametric perspective, the conditional density represents the process and is the fundamental object of interest. Nonlinear impulse response analysis is used to uncover the implications for the multi-step-ahead dynamics that are subsumed within the one-step-ahead dynamics.

The organization of the remainder of the paper is as follows. Section 2 outlines a nonlinear impulse-response methodology and discusses computations. Section 3 describes the strategy for computing bootstrap sup-norm confidence bands.
Section 4 describes the raw and adjusted price and volume series for the four stocks. Section 5 undertakes seminoparametric (SNP) estimation of univariate and bivariate conditional densities for each of the four stocks. Section 6 presents the results of the impulse response analysis. Section 7 contains concluding remarks.

2. Impulse-response analysis of nonlinear models

2.1. Notation

Let \( \{y_t\}_{t=-\infty}^{\infty} \) with \( y_t \in R^M \) be a strictly stationary process with a conditional density function that depends upon at most \( L \) lags. Collect the \( L \) lags of \( y_{t+1} \) into \( x_t = (y_{t-L+1}, \ldots, y_t) \in R^d \), and write \( f(y_{t+1} | x_t) \) for the one-step-ahead conditional density. Due to the strict stationarity assumption, the functional form of the conditional density is time-invariant.

Let \( g(y_{r-J}, y_{r-J+1}, \ldots, y_t) \) be a function of a stretch of the process of length \( J + 1 \). The conditional moment profile function, \( \{\hat{g}_j(x)\}_{j=1}^{\infty} \), is defined as the conditional expectation given \( x_t = x \) of the trajectory \( \{g(y_{r-J+j}, \ldots, y_{t+j})\}_{j=1}^{\infty} : \)

\[
\hat{g}_j(x) = \delta [g(y_{r+j-J}, \ldots, y_t) | x_t = x] = \int \cdots \int g(y_{-J}, \ldots, y) \left[ \prod_{i=0}^{j-1} f(y_{t-i+1} | y_{t-L+i}, \ldots, y_t) \right] dy_1 \cdots dy_j,
\]

for \( j = 1, 2, \ldots \). When \( x \) is fixed at some particular value, say \( x = x^* \), then we will use the term conditional moment profile to refer to the conditional moment profile function evaluated at that value, i.e., \( \{\hat{g}_j^*\}_{j=1}^{\infty} = \{\hat{g}_j(x^*)\}_{j=1}^{\infty} \) is conditional moment profile.

2.2. Perturbed initial conditions

The conditional moment profile is the basic concept for nonlinear impulse response analysis. The idea is to compare the conditional moment profile starting from a perturbed initial condition to the conditional moment profile starting from a baseline initial condition. The difference across the profiles is the impulse response. Recall the interpretation of \( x = (y_{-L+1}, y_{-L+2}, \ldots, y_0)' \), where \( y_0 \in R^M \) represents a contemporaneous value, the \( y_{-k} \in R^M, 1 \leq k \leq L - 1 \), represent lags. Let \( \delta y^+, \delta y^- \in R^M \) represent perturbations to the contemporaneous \( y_0 \) where \( \delta y^+ \) is ‘positive’ shock and \( \delta y^- \) is a ‘negative’ shock, in a sense to be made precise later. Put

\[
\begin{align*}
x^+ &= (y_{-L+1}, y_{-L+2}, \ldots, y_0)' + (0, 0, \ldots, \delta y^+)'', \\
x^0 &= (y_{-L+1}, y_{-L+2}, \ldots, y_0)', \\
x^- &= (y_{-L+1}, y_{-L+2}, \ldots, y_0)' + (0, 0, \ldots, \delta y^-)''.
\end{align*}
\]
The conditional moment profiles \( \{ \hat{g}_j(x^+) \}_{j=1}^\infty \), \( \{ \hat{g}_j(x^0) \}_{j=1}^\infty \), and \( \{ \hat{g}_j(x^-) \}_{j=1}^\infty \) are the conditional expectations of the trajectory \( \{ g(y_{t+j}, y_{t-j+1}, \ldots, y_{t+1}) \}_{j=1}^\infty \) starting from initial conditions \( x^+ \), \( x^0 \), and \( x^- \), respectively. Profiles compared to baseline

\[
\{ \hat{g}_j(x^+) - \hat{g}_j(x^0) \}_{j=1}^\infty, \quad \{ \hat{g}_j(x^-) - \hat{g}_j(x^0) \}_{j=1}^\infty
\]

are the dynamic impulse responses of \( g(y_{t+j}, y_{t-j+1}, \ldots, y_{t+1}) \) to shocks \( \delta y^+ \) and \( \delta y^- \).

The impulse response of the conditional mean is the special case obtained by letting

\[
g(y_{t-L+1}, \ldots, y_t) = \mathcal{E}(y_{t+1} \mid y_{t-L+1}, \ldots, y_t),
\]

with \( J = L - 1 \). For this case put

\[
\hat{y}_j(x) = \mathcal{E}[g(y_{t-j}, \ldots, y_{t+1}) \mid x_t = x] = \mathcal{E}[\mathcal{E}(y_{t+j} \mid y_{t-L+1}, \ldots, y_{t+1}) \mid x_t = x],
\]

and put

\[
\hat{y}_j^+ = \hat{y}_j(x^+), \quad \hat{y}_j^0 = \hat{y}_j(x^0), \quad \hat{y}_j^- = \hat{y}_j(x^-),
\]

for \( j = 1, 2, \ldots \). The conditional mean profiles \( \{ \hat{y}_j^+ \}_{j=1}^\infty \), \( \{ \hat{y}_j^0 \}_{j=1}^\infty \), and \( \{ \hat{y}_j^- \}_{j=1}^\infty \), are the conditional expectations starting from \( x^+ \), \( x^0 \), and \( x^- \), of the trajectories of the one-step conditional mean. The sequence, \( \{ \hat{y}_j^+ - \hat{y}_j^0 \}_{j=1}^\infty \), represents the impulse response to the positive impulse while \( \{ \hat{y}_j^- - \hat{y}_j^0 \}_{j=1}^\infty \) represents the impulse response to the negative impulse. An analogous notion of an impulse response is developed and refined in a line of research initiated by Simon Potter (Potter, 1994; Pesaran and Potter, 1993; and references therein). The impulse responses depend upon the initial \( x_t \), which reflects the nonlinearity of the system.

The law of iterated expectations implies that \( \hat{y}_j(x) = \mathcal{E}(y_{t+j} \mid x_t = x) \). The sequences \( \{ \hat{y}_j^+ - \hat{y}_j^0 \}_{j=1}^\infty \) and \( \{ \hat{y}_j^- - \hat{y}_j^0 \}_{j=1}^\infty \), therefore represent the effects of the shocks \( \delta y^+ \) and \( \delta y^- \) on the trajectory of the process itself. We work with perturbed conditional moment profiles, instead of perturbed trajectories of \( y_t \), because the conditional moment profiles provide the generalization of linear impulse response analysis to the general nonlinear case. Invoking the law of iterated expectations conceals the underlying structure.

To see the connection to the linear case, suppose the underlying dynamics are those of a VAR: \( A(\mathcal{L}) y_t = u_t \), where \( A(\mathcal{L}) = I - \sum_{k=1}^L A_k \mathcal{L}^k \) is a matrix polynomial in the lag operator \( \mathcal{L} \), and \( \{ u_t \}_{t=-\infty}^\infty \) is a sequence of iid innovations with mean \( \mathbb{E} u_t = 0 \) and variance matrix \( \text{Var}(u_t) = \mathbb{E} u_t u'_t = \Omega \). Suppose also that \( A(\mathcal{L}) \) is invertible, and \( B_j(\mathcal{L}) = \sum_{k=0}^\infty b_{j,k} \mathcal{L}^k \) is the \( j \)th element of \( B(\mathcal{L}) = [A(\mathcal{L})]^{-1} \). It can be shown that, if \( \delta y^+ = e_j \), where \( e_j \in \mathbb{R}^M \) is the \( j \)th unit vector,
with unity in the $j$th place and zeros elsewhere, then $b_{jk} = y_k^k$, $k = 0, 1, 2, \ldots$, is the dynamic response of the $i$th variable to a positive, one-unit movement in the $j$th element of the innovation vector $\mu_t$. Thus, in the linear case, this notion of a nonlinear impulse response correctly specializes to the VAR impulse methodology of Sims (1980).

Similarly, by suitably defining the function $g(\cdot)$, we can measure the effects of shocks on volatility. Take

$$g(y_{t-L+j}, \ldots, y_{t-1+j}) = \text{Var}(y_{t+j} \mid y_{t-L+j}, \ldots, y_{t-1+j})$$

and put

$$\hat{\psi}_j(x) = \mathcal{E}[g(y_{t-L+j}, \ldots, y_{t+j}) \mid x_t = x]$$
$$= \mathcal{E}[\text{Var}(y_{t+j} \mid x_{t+j-1}) \mid x_t = x],$$

for $j = 1, 2, \ldots$, where $x = (y_{t-L+1}, \ldots, y_0)'$. $\hat{\psi}_j(x)$ is the conditional expectation of the one-step variance matrix $j$ steps ahead, conditional on $x_t = x$. As before, let $x^0$ define baseline initial conditions, $x^+$ correspond to a positive impulse $\delta y^+$, and $x^-$ to a negative impulse $\delta y^-$. In a manner similar to that for the conditional mean, put

$$\psi_j^+ = \hat{\psi}_j(x^+), \quad \psi_j^0 = \hat{\psi}_j(x^0), \quad \psi_j^- = \hat{\psi}_j(x^-).$$

The net effects of perturbations $\delta y^+$ and $\delta y^-$ on volatility are assessed by plotting the profiles compared with the baseline:

$$\{\psi_j^+ - \psi_j^0, \psi_j^- - \psi_j^0\}.$$  

Note that the conditional volatility profile, $\{\hat{\psi}_j(x)\}_{j=1}^\infty$, is different from the path described by the $j$-step-ahead mean square error of the process, which is defined by $\hat{\mu}_j(x) = \text{Var}(y_{t+j} \mid x_t = x)$. The contrast is apparent from noting that $\text{Var}(y_{t+j} \mid x_{t+j})$ is a central moment about $\mathcal{E}(y_{t+j} \mid x_{t+j})$, while $\text{Var}(y_{t+j} \mid x_t)$ is a central moment about $\mathcal{E}(y_{t+j} \mid x_t)$. Thus $\hat{\psi}_j(x)$ and $\hat{\mu}_j(x)$ differ in the centering for a mean square error computation. Which of the two is of primary interest depends upon the character of the application. As argued in Gallant, Rossi, and Tauchen (1993), $\hat{\mu}_j(x)$ is a confounding of first- and second-moment properties of the process, while $\hat{\psi}_j(x)$ corresponds precisely to the notion of volatility in an ARCH model. Thus, $\hat{\psi}_j(x)$ is appropriate for volatility analysis, where we are interested in fluctuations around a conditional mean, though $\hat{\mu}_j(x)$ could be appropriate in other applications such as macroeconomic forecasting. Impulse response analysis can readily be extended to examine the effects of shocks on $\hat{\mu}_j(x)$, but at the expense of introducing a more complicated form of the $g(\cdot)$ function considered above.
2.3. Computations

Analytical evaluation of the integrals in the definition of a conditional moment profile is intractable, though, evaluation is well suited to Monte Carlo integration. The steps involved in Monte Carlo integration are outlined for the general case. Specializations to the conditional mean and volatility profiles are obvious.

Let \( \{y_j^r\}_{j=1}^{\infty}, r = 1, 2, \ldots, R \) denote \( R \) simulated realizations of the process starting from \( x_0 = x \). In other words, \( y_j^r \) is a random drawing from \( f(y \mid x) \) with \( x = (y_{-L+1}^r, \ldots, y_{-1}^r, y_0^r) \); \( y_j^r \) is a random draw from \( f(y \mid x) \), with \( x = (y_{-L+2}^r, \ldots, y_0^r, y_1^r) \), and so forth.

As above, \( g(y_{-J}, \ldots, y_0) \) denotes a time-invariant function of a stretch of \( \{y_t\} \) of length \( J + 1 \). Then

\[
\hat{\theta}_j(x) = \int \cdots \int g(y_{-J}, \ldots, y_j) \left[ \prod_{t=0}^{j-1} f(y_{t+1} \mid y_t, \ldots, y_1) \right] \, dy_1 \cdots dy_j
\]

\[
\approx \frac{1}{R} \sum_{r=1}^{R} g(y_{-J}^r, \ldots, y_j^r),
\]

with the approximation error tending to zero almost surely as \( R \to \infty \), under mild regularity conditions on \( f \) and \( g \).

3. Sup-norm confidence bands

As described in Gallant, Rossi, and Tauchen (1993), sup-norm bands can be constructed by bootstrapping, which is a method that uses simulation to take into account the sampling variation in the estimation of \( \hat{\theta}(y \mid x) \); see Efron (1992). In this case, the method is implemented as follows: Additional data sets of the same length as the original data are generated from the fitted conditional density \( \hat{f}(y \mid x) \) using the initial conditions of the original data. A conditional density is estimated from each data set and a profile computed from it. A 95% sup-norm confidence band is an \( \varepsilon \)-band around the profile from \( \hat{f}(y \mid x) \) that is just wide enough to contain 95% of the simulated profiles.

To illustrate, consider, as is done Section 6.2, setting a sup-norm band on a difference between the conditional volatility profiles of the first element \( y_{1t} \) of \( y_t \),

\[
\mathcal{N}_j = \delta[\text{Var}(y_{1j} \mid x_{j-1}) \mid x^+] - \delta[\text{Var}(y_{1j} \mid x_{j-1}) \mid x^-], \quad j = 1, \ldots, J.
\]

The computations proceed as follows. Let \( \{y_t^r\}_{t=L+1}^{\infty} \) denote the \( r \)th simulated data set from the conditional density \( \hat{f}(y \mid x) \), \( r = 1, \ldots, R \), computed as described in Section 5.2. In these simulations, the initial conditions \( x = (y_1^r, \ldots, y_L^r) \) are held the same for each \( r \) where as the seed of the random number generator is reset.
randomly for each \( r \). Let \( \hat{f}'(y \mid x) \) denote a nonparametric estimate of \( f(y \mid x) \) fitted to \( \{y_j^r\} \). Compute

\[
M' = \max_{1 \leq j \leq J} |\hat{N}_j^r - \tilde{N}_j^r|
\]

for each \( r \), where \( \hat{N}_j^r \) is computed from \( \hat{f}'(y \mid x) \) as described in Section 5 and \( \tilde{N}_j^r = (1/R) \sum_{r=1}^{R} \hat{N}_j^r \). Lastly, letting \( M^{0.95} \) denote the 0.95 quantile of the \( \{M'_r\} \), the sup-norm confidence band on \( \{\hat{N}_j : j = 1, \ldots, J\} \) is

\[
\hat{N}_j^r \pm M^{0.95}, \quad j = 1, 2, \ldots, J.
\]

A sample from \( \hat{f}(y \mid x) \) will usually appear more homogeneous than a sample from the true conditional density (Efron and LePage, 1992). This would tend to bias the bandwidth determined from the data upward for a kernel estimator and the truncation point of a series estimator downward. Thus, we take the truncation point of \( \hat{f}'(y \mid x) \) be that of \( \hat{f}(y \mid x) \).

4. Data and adjustments

The basic raw data consists of the daily closing prices and the number of shares traded on the New York Stock Exchange for the common stocks of four companies: International Business Machine (IBM), Boeing (BA), Coca-Cola (KO), and Minnesota Mining & Manufacturing (MMM). The sample period is from March 1, 1982 to December 29, 1988, for 1729 observations on each series.

The data sets were constructed from two main sources: FITCH NYSE transaction data tapes and ISSM transaction data tapes. Data for 1982 to 1986 were extracted from FITCH tapes and the remainder taken from the ISSM tapes. We used the Daily Stock Price Record (DSPR) published by Standard and Poor’s to fill out (occasional) missing values in each data set. For missing price observations, we used the prices quoted in the DSPR. For missing volume observations, we used the DSPR figure with a correction because volume as reported in DSPR also includes trades on the Midwest, Pacific, Philadelphia, Boston, and Cincinnati stock exchanges, and trades reported by the National Association of Securities Dealers and Instinet. We also corrected the price and volume series for BA, KO, and MMM for stock splits. IBM did not split over the sample period.

The raw price series, \( P_i^t \), is differenced in the logs to create the price change series, \( 100(\log P_i^t - \log P_{i-1}^t) \), and the raw volume series is also transformed into the log \( V_i^t \), where \( i = \text{IBM, BA, KO, and MMM} \).

Many empirical studies have documented systematic calendar effects and seasonalties in both the mean and variance of stock price and trading volume series. A comprehensive exposition is given in Gallant, Rossi, and Tauchen (1992). To adjust for these documented effects, we performed an adjustment process in which systematic effects are first removed from the mean and then from the variance.
We use the following set of dummy variables in the adjustment regressions to capture these systematic effects:

(1) Day of the week dummies (one for each day, Tuesday through Friday).

(2) Dummy variables for each number of nontrading days preceding the current trading day (dummies for each of two, three, and four nontrading days since the preceding trading day); these variables capture the effects of holidays and weekends.

(3) Dummy variables for months of March, April, May, June, July, August, September, October, and November.

(4) Dummy variables for each week of December and January; These variables are used to accommodate the January effect in both the mean and variance of prices and volume.

This list comprises $4 \cdot 3 + 9 + 8 = 24$ dummies in total, estimated on 1729 observations, so there are many degrees of freedom in the adjustment regressions.

Volume data tend to show long persistent movements about a very gentle upward trend. After extensive elementary data analysis and inspection of plots, we decided not to adjust for trend in BA, KO, or MMM. In effect, this decision attributes the very low-frequency movements in volume over the seven-year sample period to low-frequency components of a stationary process. The nonparametric estimation procedure has to adapt to these movements. In the case of IBM, however, the trend was much more pronounced, making implausible the assumption of (approximate) stationarity over the seven-year sample period. Thus, for the IBM series we included $t$ and $t^2$ time trend variables in the mean regressions for the volume.

We adjust rescaled log first-differences of the stock prices and log volume. Using log first-differences is established practice for financial prices; using log volume is motivated by the fact that volume is nonnegative and tends to be volatile at higher levels and quiescent at low levels, so logging achieves some variance stabilization that lessen the burden on the nonparametric procedures. To perform the adjustment to $100(\log P_t^i - \log P_{t-1}^i)$ or $\log V_t^i$, where $i = $IBM, BA, KO, and MMM, we first run the regression:

$$ y = x'\beta + u $$  \hspace{1cm} (mean \ equation),

where $y$ is the series to be adjusted and $x$ is the set of adjustment regressors as described above. We then take the least squares residuals from the mean equation to construct a variance equation:

$$ \log(u^2) = x'y + \varepsilon $$  \hspace{1cm} (variance \ equation).

This regression is used to standardize the residuals from the mean equation and then a final linear transformation is performed to create adjusted $y$. 
\[ y_{\text{adj}} = a + b \left( \frac{\hat{\sigma}}{\exp(x^2/2)} \right), \]

where \( a \) and \( b \) are chosen such that the sample means and variances of series are the same before and after the adjustment, which eases interpretation of the empirical results.

The choice of dummies for these adjustments was motivated by prior information about the seasonal nature of stock returns and volume. The dummies were generally statistically significant, but actually contributed little to the explanation of the series, because price and volume movements are intrinsically very noisy. With eight series to adjust, there are sixteen adjustment regressions to run for the means and variances, and, of course, not every coefficient in every equation is statistically significant. We resisted making judgmental decisions on suppressing dummies, so as to keep the procedure simple, replicable, and to avoid risk of specification error at such an early stage of the research.

The top half of Fig. 1 shows the adjusted price change data for each of bottom half the four securities; the bottom half shows the adjusted volume data. The general appearance of the series is typical for financial data, indicating that the data are not overprocessed by the adjustments. Leaving the in low-frequency movements of volume for BA, KO, and MMM is perhaps the most contentious aspect of what we do. As a robustness check, we also added to the volume adjustments a dummy for a level shift after the crash of 1987 and redid the entire analysis. The results were nearly the same as reported in Section 6 below; the only major difference was that the impulse response of MMM volume shows more rapid return to baseline in response to own shocks. This is not surprising from looking at the series in the lower right panel of Fig. 1, because a major low-frequency movement is omitted when the 1987 crash dummy is included. We also experimented with breaking trends in the adjustment equations, but the results suggested little evidence for trend breaks.

5. Estimation of the conditional density

5.1. Semiparametric (SNP) estimators

The SNP method is based on the notion that a Hermite expansion can be used as a general purpose nonparametric estimator of a density function. Letting \( z \) denote an \( M \)-vector, the particular Hermite expansion employed has the form

\[ f(z) \propto [P(z)]^2 \phi(z), \]

where \( P(z) \) denotes a multivariate polynomial of degree \( K_z \) and \( \phi(z) \) denotes the density function of the (multivariate) Gaussian distribution with mean zero and the identity matrix as its variance–covariance matrix. The constant of proportionality is the divisor \( \int [P(z)]^2 \phi(z) \, dz \) which makes \( f(z) \) integrate to one. Because of this division, the density is a homogeneous function.
Fig. 1. Adjusted data.
of the coefficients of the polynomial \( P(z) \), and these coefficients can only be determined to within a scalar multiple. To achieve a unique representation, the constant term of the polynomial part is put to one.

The location/scale shift \( y = Rz + \mu \), where \( R \) is an upper triangular matrix and \( \mu \) is an \( M \)-vector, followed by a change of variables leads to a parameterization that is easy to interpret: \( f(y \mid \theta) \propto \{ P[R^{-1}(y - \mu)] \}^2 \{ \phi[R^{-1}(y - \mu)]/\vert \det(R) \vert \}. \) Because \{ \phi[R^{-1}(y - \mu)]/\vert \det(R) \vert \} is the density function of the \( M \)-dimensional multivariate Gaussian distribution with mean \( \mu \) and variance–covariance matrix \( \Sigma = RR' \), and because the leading term of the polynomial part equals unity, the leading term of the entire expansion is the multivariate Gaussian density function; denote it by \( n_M(y \mid \mu, \Sigma) \). When \( \Sigma \) is put to zero, one gets \( n_M(y \mid \mu, \Sigma) \) exactly. When \( \Sigma \) is positive, one gets a Gaussian density whose shape is modified due to multiplication by a polynomial in the normalized error \( z = R^{-1}(y - \mu) \).

The shape modifications thus achieved are rich enough to accurately approximate under a weighted Sobelov norm densities from a large class that includes densities with fat \( t \)-like tails, densities with tails that are thinner than Gaussian, and skewed densities. The weighted Sobelov norm puts most weight on the central part and the near tails of the density, where most or all of the data tend to lie in applications, and little weight on the extreme tails, where little or no data lie. Gallant and Nychka (1987) and Andrews (1991) provide the theoretical justification for using series expansion estimators. Fenton and Gallant (1995) present striking evidence on the capability of the SNP family to approximate extremely well quite strangely shaped densities.

The parameters \( \theta \) of \( f(y \mid \theta) \) are made up of the coefficients of the polynomial \( P(z) \) plus \( \mu \) and \( R \) and are estimated by maximum likelihood. Equivalent to maximum likelihood but more stable numerically is to estimate \( \theta \) in a sample of size \( n \) by minimizing \( s_n(\theta) = (-1/n) \sum_{i=1}^{n} \log[f(y_i \mid \theta)] \). If the number of parameters \( p_\theta \) grows with the sample size \( n \), then the true density and various features of it such as derivatives and moments are estimated consistently. Because the method is parametric yet has nonparametric properties, it is termed semiparametric to suggest that it lies halfway between parametric and nonparametric procedures.

This basic approach is adapted to the estimation of the conditional density of a multiple time series that has a Markovian structure as follows. By a Markovian structure, one means that the conditional density of the \( M \)-vector \( y_t \) given the entire past \( y_{t-1}, y_{t-2}, \ldots \) depends only on \( L \) lags from the past. For notational convenience, we collect these lags together in a single vector denoted as \( x_{t-1} \) which has length \( M \cdot L \).

As above, a density is obtained by a location/scale shift \( y_t = Rx_t + \mu_t \) off a sequence of normalized errors \( \{z_t\} \). Here \( \mu_t \) is a linear function of \( x_{t-1} \), specifically \( \mu_t = b_0 + Bx_{t-1} \) where \( b_0 \) is \( M \times 1 \) and \( B \) is \( M \times M \cdot L \), making the leading term of the expansion \( n_M(y \mid \mu_t, \Sigma) \), \( \Sigma = RR' \) which is a Gaussian vector autoregression (VAR).
The leading term of the expansion can be put to a Gaussian ARCH, instead of a Gaussian VAR, by letting $R$ be a linear function of the absolute values of the elements of $L_r$ lags of $y_t$, centered and scaled to have mean zero and identity covariance matrix. This differs from the classical ARCH (Engle, 1982) which has $\Sigma_t$ depending on a linear function of squared lagged residuals; the SNP version of ARCH is more akin to the suggestions of Nelson (1991) and Davidian and Carroll (1987). The SNP specification is $\Sigma_t = R_t R_t'$, with $\text{vech}(R_t) = P_0 + P_1 \text{abs}(x_{t-1}^r)$, where $P_0$ and $P_1$ are coefficient matrices of dimension $M \cdot (M + 1)/2 \times 1$ and $[(M \cdot (M + 1)/2)] \times M \cdot L$, respectively, $x_{t-1}^r$ is centered and rescaled $x_{t-1}$, and $\text{abs}(x_{t-1}^r)$ is the element-wise absolute value of $x_{t-1}^r$.

In time series analysis, the $\{z_t\}$ are usually referred to as linear innovations. In order to permit the innovations to be conditionally heterogeneous, the coefficients of the polynomial $P(z)$ are, themselves, polynomials of degree $K_z$ in $x_{t-1}$. This polynomial is denoted as $P(z, x)$. When $K_z = 0$, the $\{z_t\}$ are homogeneous, as the conditional density of $z_t$ does not depend upon $x_{t-1}$. When $K_z > 0$, the $\{z_t\}$ are conditionally heterogeneous. The tuning parameter $K_z$ controls the extent to which the model deviates from normality while $K_x$ controls the extent to which these deviations vary with the history of the process.

We distinguish between the total number of lags under consideration, which is $L$, the number of lags of $y_t$ in the $x$ part of the polynomial $P(z, x)$, which we denote by $L_a$, and the number of lags of $y_t$ in $R_z$, which is $L_r$. The vector $x$ has length $M \cdot L$ where $L = \max(L_r, L_a)$.

Large values of $M$ can generate a large number of interactions (cross product terms) for even modest settings of degree $K_z$; similarly, for $M \cdot L_a$ and $K_x$. Accordingly, there are two additional tuning parameters, $I_x$ and $I_r$, to represent filtering out of these high-order interactions. $I_x = 0$ means no interactions are suppressed, $I_x = 1$ means the highest-order interactions are suppressed, namely those of degree exceeding $K_z - 1$. In general, a positive $I_x$ means all interactions of order exceeding $K_z - I_x$ are suppressed; similarly for $K_x - I_x$.

To illustrate, consider first the model with $M = 1$, $L_r = 4$, $L_a = 0$, $K_z = 4$, $K_x = 0$; for these settings, $I_x = I_r = 0$ are irrelevant. The polynomial is of the form

$$P(z) = \sum_{t=0}^{4} a_t z^t,$$

where the $a_0, a_1, \ldots, a_4$ are the polynomial coefficients with, as noted above, the constant term $a_0 = 1$ to achieve a unique representation. Both $\mu_t$ and $R_z$ are linear in $y_{t-j}$ back to four lags. This model has fourteen free parameters: the four free polynomial parameters, the intercept and four slope parameters in $\mu_t$, and the intercept and four slope parameters of $R_z$.

Now consider $L_a = 1$ and $K_x = 1$ but everything else the same. The polynomial becomes

$$P(z, x) = \sum_{t=0}^{4} (a_{0t} + a_{1t, y_{t-1}}) z^t.$$
Recall that $x = (y_{-4}, y_{-3}, y_{-2}, y_{-1})$, and with $L_a = 1$ and $K_x = 1$ the polynomial is linear in first lag. The normalization is $a_{00} = 1$, so the polynomial has nine free parameters, yielding nineteen free parameters in total.

For a multivariate example, consider the case where $M = 2$, $L_r = 4$, $L_a = 1$, $K_z = 4$, $L_z = 2$, $K_x = 2$, and $I_x = 1$. Then the polynomial $P(z, x)$ would take the form

$$P(z, x) = \sum_{\lambda_1, \lambda_2} a_{\lambda_1, \lambda_2}(x) z_1^{\lambda_1} z_2^{\lambda_2},$$

where the $a_{\lambda_1, \lambda_2}(x)$ are the coefficients of the polynomial in $z \in \mathbb{R}^2$, and the sum is over all pairs of nonnegative integers $(\lambda_1, \lambda_2)$ such that $\lambda_1 + \lambda_2 \leq 4$, excluding the quartic interactions, $(1,3)$, $(2,2)$, and $(3,1)$, and the cubic interactions, $(2,1)$ and $(1,2)$, suppressed by $I_z = 2$. With $L_a = 1$, $K_x = 2$, and $I_x = 1$, the $a_{\lambda_1, \lambda_2}(x)$ are quadratic functions of the elements of $y_{-1}$, a $2 \times 1$ vector, with all cross-products suppressed, since these are quadratic interactions. Recall that $x = (y_{-4}, \ldots, y_{-1})$. 

In general, $L_r$ and $L_a$ determine the location/scale shift $y = R_x z_t + \mu_x$ and hence determine the nature of the leading term of the expansion. The number of lags in the location shift $\mu_x$ is the overall lag length $L$ which is the maximum of $L_r$ and $L_a$. The number of lags in the scale shift $R_x$ is $L_r$. The number of lags that go into the $x$ part of the polynomial $P(z, x)$ is $L_a$. The parameters $K_z$ and $K_x$ determine the degree of $P(z, x)$ and hence the nature of the innovation process $\{z_t\}$. $I_z$ and $I_x$ determine filters that suppress interactions when set to positive values. The SNP software package automatically counts the number of free parameters in any model given the settings of the tuning parameters.

The form of the SNP conditional density is $f(y \mid x, \theta) \propto [P(z, x)]^2 n_M(y \mid \mu_x, \Sigma_x)$ where $z = R_x^{-1}(y - \mu_x)$ and $\theta$ denotes the coefficients of the polynomial $P(z, x)$ and the Gaussian ARCH $n_M(y \mid \mu_x, \Sigma_x)$ collected together. This form provides a flexible family for modeling nonlinear time series controlled by the tuning parameters $L_r$, $L_a$, $K_z$, $I_z$, $K_x$, and $I_x$. In a recent forecasting comparison (Ellner et al., 1995), SNP was found to perform very well relative to more parametric approaches. Putting certain of the tuning parameters to zero implies sharp restrictions on the process $\{y_t\}$, the more interesting of which are given in Table 1.

<table>
<thead>
<tr>
<th>Parameter setting</th>
<th>Characterization of ${y_t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_r = 0, L_a = 0, K_z = 0, K_x = 0$</td>
<td>iid Gaussian</td>
</tr>
<tr>
<td>$L_r = 0, L_a &gt; 0, K_z = 0, K_x = 0$</td>
<td>Gaussian VAR</td>
</tr>
<tr>
<td>$L_r = 0, L_a &gt; 0, K_z &gt; 0, K_x = 0$</td>
<td>Non-Gaussian VAR, homogeneous innovations</td>
</tr>
<tr>
<td>$L_r &gt; 0, L_a \geq 0, K_z = 0, K_x = 0$</td>
<td>Gaussian ARCH</td>
</tr>
<tr>
<td>$L_r &gt; 0, L_a \geq 0, K_z &gt; 0, K_x = 0$</td>
<td>Non-Gaussian ARCH, homogeneous innovations</td>
</tr>
<tr>
<td>$L_r &gt; 0, L_a &gt; 0, K_z &gt; 0, K_x &gt; 0$</td>
<td>Full nonlinear non-Gaussian</td>
</tr>
</tbody>
</table>
5.2. SNP fitting of the conditional density

To determine an appropriate SNP specification within this hierarchy of models, we follow a strategy similar to that of Gallant and Tauchen (1992). The basic idea is to use the Schwarz model selection criterion (Schwarz, 1978) together with a battery of specification tests on the conditional mean and variance. The model selection criterion determines the initial candidate model. The candidate model is then subjected to a battery of specification tests on the conditional mean and variance. Typically, because of the conservative nature of the Schwarz criterion, the candidate model fails some or all the specification tests. The candidate model is expanded as needed until a model is found that is not too much further up in the hierarchy from the initial set and also does well on the specification tests. We also considered the Hannan–Quinn model selection criterion (Hannan, 1987) to choose the initial model, but, for our data sets, the model selected under Hannan–Quinn was found to differ little from that selected under Schwarz.

The Schwarz criterion is

\[ s_n(\hat{\theta}) + (p_0/n)\log(n)/2, \]

where

\[ s_n(\hat{\theta}) = (-1/n) \sum_{i=1}^{n} \log f(y_i \mid x_{i-1}, \theta) \]

is the SNP optimization criterion and \( p_0 \) is the number of free parameters in the model. The criterion is of the form \( s_n(\hat{\theta}) + (p_0/n)c_n \), where \( c_n \) is a penalty, and a smaller value of a criterion indicates a better fit. Asymptotic analysis (Hannan, 1987; Potscher, 1989) indicates that were the true model a stationary finite-dimensional ARMA, then a criterion such that \( c_n/n \to 0 \) and \( c_n \to \log[\log(n)] \) for large enough \( n \) will select the true ARMA model with probability one as \( n \to \infty \). On the other hand, criteria such that \( \limsup_{n \to \infty} c_n/n \to 0 \) will underfit and those such that \( \liminf_{n \to \infty} c_n/n/\log[\log(n)] \to 1 \) will overfit with positive probability. From this analysis, Schwarz is \( \varepsilon \) consistent criterion while the Hannan–Quinn defines a lower boundary on consistent criteria for ARMA models. For SNP estimation, these asymptotic results are only suggestive, because there is no claim that the true model is a finite-dimensional SNP, and there is always some bias associated with any selected finite-dimensional model. For our purposes, Schwarz might be expected to select too small of a model, a conjecture borne out in practice because it almost always selects a model that does poorly on specification tests (Gallant and Tauchen, 1992).

The specification tests are constructed from the standardized residuals. The standardized residuals are the elements of \([\text{Var}_{t-1}(y_t)]^{-1/2} [y_t - E_{t-1}(y_t)] \in \mathbb{R}^d\). The specification test on the conditional mean is a regression of each of the standardized residuals on a constant and the distinct elements of \( \{y_{t-k}, y_{t-k} \otimes y_{t-k}, y_{t-k} \otimes y_{t-k} \otimes y_{t-k}\}_{k=1}^{3}\). The specification test on the conditional variance is a
regression of the square of each standardized residual on a constant and \(\{y_{t-k},\ y_{t-k} \otimes y_{t-k}, y_{t-k} \otimes y_{t-k} \otimes y_{t-k}\}_{k=1}^2\). The presence of lagged linear, quadratic, and cubic terms provides power against general nonlinear misspecification of either the conditional mean and variance function.

This specification search procedure entails three steps: (i) estimate many SNP models to determine the initial tentative model; (ii) run the diagnostics on this model; (iii) estimate additional SNP models and run diagnostics as need be to determine an adequate fit. Space constraints preclude reporting the full objective function surface. An earlier version of this paper [available electronically upon request until two years after publication] reports the full surface for all estimations in eight full page tables. We report only the initial tentative model and the final preferred model, bearing in mind that the underlying specification search entails consideration of substantially more models than reported on here.

In some cases, the models turn out to be quite large, which reflects the complicated nonlinear structure that financial data can display. It is now recognized that simple parametric models like GARCH can miss features of the data; for instance, Nelson (1994) reports that conditional (i.e., time-varying) leptokurtosis is present in stock returns. A flexible, nonparametric procedure like SNP automatically expands the leading term to accommodate such features, when they are present in the data. On the other hand, model expansion can stop at the leading term, as happens immediately below for univariate BA series, when the evidence suggests no further expansion is warranted.

5.3. Univariate estimation

We start with estimation of the conditional density of the univariate price change process. In this case the process \(\upsilon_t = \Delta p_t\) is scalar. Table 2 displays, for each of the four series, the objective function values and the diagnostic tests for initial tentative model selected under the Schwarz criterion and the preferred final model. In the univariate case there are no \(z\) interactions present, and it turned out that there never was a need to suppress \(x\)-interactions; thus \(L_z = 0\) and \(I_x = 0\) throughout the univariate estimation.

5.3.1. Univariate: Boeing (BA)

As reported in Table 2, the model selected under the Schwarz criterion is defined by \(L_x = 1, L_a = 0, K_z = 6, K_x = 0\). This model has one lag in the ARCH component and in the linear AR component, and a highly non-Gaussian error structure reflected by the fact that a polynomial of degree 6 in \(z\) is selected. It is essentially an ARCH model with a nonparametric error density and is very similar to the model of Engle and Gonzales-Rivera (1991). This model easily passes the diagnostic test at conventional significance levels, and is the final selected model for the impulse response analysis.
Table 2
Univariate SNP estimation

<table>
<thead>
<tr>
<th>Objective function values</th>
<th>( L_r )</th>
<th>( L_a )</th>
<th>( K_z )</th>
<th>( K_x )</th>
<th>( p_0 )</th>
<th>( s_0(\theta) )</th>
<th>Schwarz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected under the Schwarz criterion</td>
<td>( 1 )</td>
<td>( 0 )</td>
<td>( 6 )</td>
<td>( 0 )</td>
<td>( 10 )</td>
<td>1.362870</td>
<td>1.384429</td>
</tr>
<tr>
<td>BA</td>
<td>( 4 )</td>
<td>( 0 )</td>
<td>( 8 )</td>
<td>( 0 )</td>
<td>( 18 )</td>
<td>1.339072</td>
<td>1.377880</td>
</tr>
<tr>
<td>IBM</td>
<td>( 5 )</td>
<td>( 0 )</td>
<td>( 9 )</td>
<td>( 0 )</td>
<td>( 17 )</td>
<td>1.240193</td>
<td>1.274689</td>
</tr>
<tr>
<td>KO</td>
<td>( 6 )</td>
<td>( 0 )</td>
<td>( 9 )</td>
<td>( 0 )</td>
<td>( 22 )</td>
<td>1.253055</td>
<td>1.300486</td>
</tr>
<tr>
<td>MMM</td>
<td>( 1 )</td>
<td>( 0 )</td>
<td>( 6 )</td>
<td>( 0 )</td>
<td>( 10 )</td>
<td>1.362870</td>
<td>1.384429</td>
</tr>
<tr>
<td>Final model selected</td>
<td>( 4 )</td>
<td>( 1 )</td>
<td>( 8 )</td>
<td>( 1 )</td>
<td>( 27 )</td>
<td>1.334260</td>
<td>1.392470</td>
</tr>
<tr>
<td>BA</td>
<td>( 5 )</td>
<td>( 1 )</td>
<td>( 6 )</td>
<td>( 1 )</td>
<td>( 26 )</td>
<td>1.226748</td>
<td>1.280647</td>
</tr>
<tr>
<td>IBM</td>
<td>( 6 )</td>
<td>( 1 )</td>
<td>( 8 )</td>
<td>( 1 )</td>
<td>( 31 )</td>
<td>1.249770</td>
<td>1.316605</td>
</tr>
</tbody>
</table>

Specification tests

<table>
<thead>
<tr>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_r )</td>
<td>( L_a )</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>Selected under the Schwarz criterion</td>
<td>( 1 )</td>
</tr>
<tr>
<td>BA</td>
<td>( 4 )</td>
</tr>
<tr>
<td>IBM</td>
<td>( 5 )</td>
</tr>
<tr>
<td>KO</td>
<td>( 6 )</td>
</tr>
<tr>
<td>Final model selected</td>
<td>( 1 )</td>
</tr>
<tr>
<td>BA</td>
<td>( 4 )</td>
</tr>
<tr>
<td>IBM</td>
<td>( 5 )</td>
</tr>
<tr>
<td>KO</td>
<td>( 6 )</td>
</tr>
</tbody>
</table>

\( p_0 \) is the number of free parameters in the model; \( l_z = l_x = 0 \) in all univariate estimation.

5.3.2. Univariate: IBM

The model selected under the Schwarz criterion has \( L_r = 4, L_a = 0, K_z = 8, K_x = 0 \). As seen from Table 2, this model passes the specification test of the conditional variance at the 1% but not at the 5% level, while model \( L_r = 4, L_a = 0, K_z = 8, K_x = 1 \) easily passes all tests. In what follows, we use the latter as the primary model, noting that it is only marginally preferred.

5.3.3. Univariate: Coca-Cola (KO)

The model selected under the Schwarz criterion is defined by \( L_r = 5, L_a = 0, K_z = 4, K_x = 0 \). This model performs very poorly on both the conditional mean and the variance test. As seen in Table 2, the model defined by \( L_r = 5, L_a = 1, K_z = 6, K_x = 1 \) passes both tests at the 1% level and the variance test at the 5% level, and is taken as the primary model for subsequent analysis.
5.3.4. Univariate: MMM

The model selected under the Schwarz criterion is $L_r = 6$, $L_a = 0$, $K_z = 8$, $K_x = 0$. This model just barely passes the diagnostics (at 1% but not 5%), while the model $L_r = 6$, $L_a = 1$, $K_z = 8$, $K_x = 1$, does better, particularly on the specification test on the variance, and is taken as the primary model.

5.4. Bivariate estimation

We now implement this model selection procedure to estimate the conditional density of the bivariate process $y_t = (\Delta p_t, v_t)'$. Table 3 displays the objective function values and specification test results for the model selected under the Schwarz criterion and the final selected model.

5.4.1. Bivariate: Boeing (BA)

As seen in Table 3, the model selected under Schwarz just barely pass the specification test on the conditional variance of $\Delta p_t$ and fails the specification test on the conditional mean of $v_t$. After proceeding through an extensive series of SNP models, it was found that further expansion to $L_r = 2$, $L_a = 2$, $K_z = 4$, $I_z = 0$, $K_x = 2$, $I_x = 1$ provides a final model that passes all the specification tests at the 1% level.

The final model is quite large—160 parameters—with a rich non-Gaussian/nonlinear dynamic structure. It is characterized by a fourth-degree polynomial in $z_t \in \mathbb{R}^2$ with a full set of interactions. The coefficients of the polynomial depend upon a quadratic function of two lags of the bivariate process with quadratic interactions suppressed by $I_z = 1$. Examining the specification tests of nearby models show there is little scope to suppress interactions in $z_t$ and reduce to unity the lag length of the polynomial part and still pass the test on the conditional mean of $v_t$.

5.4.2. Bivariate: IBM

In Table 3, except for the test on the conditional variance of $v_t$, the model selected under Schwarz does very poorly on all of the diagnostics. Moving up from that model, it was found that $L_r = 4$, $L_a = 1$, $K_z = 4$, $I_z = 0$, $K_x = 1$, $I_x = 0$ provides a model that passes all four specification tests at the 1% level.

5.4.3. Bivariate: Coca-Cola (KO)

The bivariate Coca-Cola series proved the most difficult to fit. Initial results indicated some numerical instability. The instability was of the type occasionally encountered in multivariate SNP estimation where extreme values of one series can be used to fit nearly exactly the conditional variance of another series. Re-specifying the ARCH-type model of the leading term so that it is a diagonal ARCH (the diagonal elements of the covariance matrix depend upon own lags only) helped considerably to stabilize the calculations.
Table 3
Bivariate SNP estimation

<table>
<thead>
<tr>
<th>Objective function values</th>
<th>$L_r$</th>
<th>$L_a$</th>
<th>$K_z$</th>
<th>$I_z$</th>
<th>$K_x$</th>
<th>$I_x$</th>
<th>$p_\theta$</th>
<th>$s_n(\hat{\theta})$</th>
<th>Schwarz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected under the Schwarz criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>2.589843</td>
<td>2.652366</td>
</tr>
<tr>
<td>IBM</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>2.523387</td>
<td>2.598846</td>
</tr>
<tr>
<td>KO</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>2.484385</td>
<td>2.555531</td>
</tr>
<tr>
<td>MMM</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>59</td>
<td>2.416757</td>
<td>2.543958</td>
</tr>
<tr>
<td>Final model selected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>160</td>
<td>2.500243</td>
<td>2.843040</td>
</tr>
<tr>
<td>IBM</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>89</td>
<td>2.473524</td>
<td>2.665404</td>
</tr>
<tr>
<td>KO</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>63</td>
<td>2.458653</td>
<td>2.594479</td>
</tr>
<tr>
<td>MMM</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>89</td>
<td>2.399210</td>
<td>2.591090</td>
</tr>
</tbody>
</table>

Specification tests

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th>Volume</th>
<th>Veriance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L_r$</td>
<td>$L_a$</td>
<td>$K_z$</td>
<td>$I_z$</td>
</tr>
<tr>
<td></td>
<td>$L_r$</td>
<td>$L_a$</td>
<td>$K_z$</td>
<td>$I_z$</td>
</tr>
<tr>
<td>Selected under the Schwarz criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>IBM</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>KO</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>MMM</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Final model selected</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>IBM</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>KO</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>MMM</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

194
The model selected under Schwarz fails the diagnostics for both the conditional mean and variance of \( A_{\Delta p_t} \), and is borderline on the specification test for the conditional mean of \( v_t \). Unlike just about all other fitting, further expansion of the model provided only little help on the specification tests. In what follows we use the fit \( L_r = 4, L_a = 1, K_z = 4, I_z = 1, K_x = 1 \), bearing in mind that any interpretation is tenuous in view of the outcome of the specification tests shown in Table 3.

5.4.4. Bivariate: MMM

For the model selected under the Schwarz criterion, the specification tests on the conditional variance of \( \Delta p_t \) and the conditional mean of \( v_t \) are on the 5% borderline while the other two are alright. Expansion to \( L_r = 4, L_a = 1, K_z = 4, I_z = 0, K_x = 1, I_x = 0 \) helps substantially on the specification test on the conditional variance of \( \Delta p_t \). In what follows, we use this model as the preferred model, keeping in mind that the simpler model with \( L_r = 4, K_z = 4, K_x = 0 \) is nearly adequate.

6. Results

6.1. Univariate

Fig. 2 shows, for each of the four securities, the impulse response of the conditional variance of \( \Delta p_t \) to 5% price shocks. Here \( y_t = \Delta p_t \). The four panels contain the conditional variance profiles \( \{ v_j^{\Delta p_t} \}_{j=1}^{20}, \{ v_j^0 \}_{j=1}^{20}, \{ v_j^{-} \}_{j=1}^{20}, \{ v_j^{0-} \}_{j=1}^{20} \), computed as

\[
\hat{v}_j^{\Delta p_t} = \hat{v}_j(x^+), \quad \hat{v}_j^0 = \hat{v}_j(x^0), \quad \hat{v}_j^{-} = \hat{v}_j(x^-),
\]

where \( \hat{v}_j(x) \) is the conditional variance profile function defined in Section 2.2. The baseline is computed at the mean of data, \( x^0 = \bar{x} = \frac{1}{T-L+1} \sum_{t=L+1}^{T} x_t \), where \( L \) is the length of \( x \). The perturbed initial conditions are 5% price movements: \( x^+ = \bar{x} + (0, \ldots, 0, 5.0) \) and \( x^- = \bar{x} + (0, \ldots, 0, -5.0) \) [keep in mind the interpretation \( x_t = (y_t-L+1, \ldots, y_T) \)]. The contrasts, \( \{ v_j^{\Delta p_t} - v_j^0 \}_{j=1}^{20} \) and \( \{ v_j^{-} - v_j^{0-} \}_{j=1}^{20} \), are the impulse responses, the net effects of the shocks. We also computed, but do not report, the impulse response of the conditional mean \( \delta(y_{t+1} \mid x_t) \) to the same price shocks. For the conditional mean, the effects are symmetric and totally dissipated within a day of the shock, suggesting very little evidence for nonlinearity in the conditional mean of the price process.

As seen from the four panels, the responses of volatility to price shocks are slowly damped for three of the four securities: IBM, KO, and MMM. For these three securities, it takes between 10 and 20 days for the effects of shock on volatility to dissipate, with KO and MMM on the high end of that range and
IBM on the short end. In contrast, for BA the volatility response damps out just a few days after the shock.

Interestingly, the volatility responses of all four stock prices to own-price shocks are sluggish, though much less persistent than those reported in Gallant, Rossi, and Tauchen (1993) for the S&P Index. Indeed, Nelson (1991) and others report that the index is nearly integrated in variance, while our four securities are not. We emphasize that SNP can detect extreme volatility persistence when it is an important feature of the series. In fact, Bollerslev and Mikkelsen (1993) use the extreme volatility persistence detected by SNP in Gallant, Rossi, and Tauchen (1993) as evidence that discredits GARCH and points toward fractionally integrated volatility.

Our sample period is 1982–1988, which is a subperiod of that used by others in fitting the index. The relatively quick decay of volatility detected here could be a feature of this subperiod, though we note that the period includes the crash and the surrounding period of high financial volatility. Tim Bollerslev (private communication) noted that contrast between the findings on volatility persistence for individual stocks and the S&P Index might be explained by a volatility setup patterned after Granger (1980), where an aggregation of AR(1) processes displays fractional integration. In this setup, the volatility of a cross-sectional aggregate of stock prices can be expected to display persistent, fractionally integrated behavior like that reported in Bollerslev and Mikkelsen (1993) for the S&P Index. We
leave this point as an issue to be taken up in future research with a broader cross-section of securities covering other time periods.

The leverage effect (Black, 1976; Christie, 1982; Nelson, 1991; Campbell, Grossman, and Wang, 1993) refers to an asymmetry in the conditional variance function of speculative price changes. Price movements in either direction are known to lead to increased volatility. The leverage effect is the tendency for a price decline to lead to a subsequent volatility increase that is larger than the increase in volatility associated with a price rise of the same magnitude. In other words, a crash has a larger impact on volatility than does a boom of the same magnitude.

Since nonlinear impulse response analysis traces out the dynamic effect of shocks, it is very well suited for assessing the empirical importance of the leverage effect. In terms of our notation, the leverage effect is the prediction that

\[ \hat{\sigma}_j^+ - \hat{\sigma}_j^0 > \hat{\sigma}_j^- - \hat{\sigma}_j^0 \quad \text{for} \quad j = 1, 2, \ldots . \]

From Fig. 2, it is seen that only IBM shows evidence of a persistent leverage effect, and even then the asymmetry is gone after three or four days. BA shows no asymmetry at all, because the preferred model from the specification search is, in effect, an ARCH model with a nonparametric error density, which is constrained to be symmetric. On the other hand, there are no symmetry constraints imposed on the preferred fits for KO and MMM. Still, there is evidence in the univariate results for at most a small leverage effect that lasts only a day or two at most.

6.2. Bivariate

We now assess the effects of price and volume shocks in the bivariate case where \( y_t = (\Delta p_t v_t)' \).

6.2.1. Shocking strategies

With multivariate data that exhibit contemporaneous cross-correlations, it can be very unrealistic and misleading to perturb one of the variables while leaving the others fixed. This point has been well appreciated in the literature since the original Sims (1980) paper and has lead to the extensive work on orthogonalized shocks. In the nonlinear case, the situation is even more complex, because there can be nonlinear contemporaneous relations among variables that need to be taken into account.

As is known from Clark (1973), Tauchen and Pitts (1983), and Karpoff (1987), the volume is contemporaneously related to the magnitude of the price movement. In other words, a large movement in either \( v_t \) or \( v_t - \delta_{t-1}(v_t) \) is positively related to \( (\Delta p_{t-1} - \delta_{t-1}(\Delta p_t))^2 \). There is a relationship between the level of the second series and the variability of the first. No strategy based on using orthogonalized shocks can take this relationship into account, because the orthogonalization strategy only considers first-moment properties. Gallant, Rossi, and Tauchen (1993)
propose eschewing notions of orthogonalization and simply thinking in terms of the system's response to typical shocks. Following on their ideas, we investigate the effects of three types of shocks:

\begin{align*}
A \text{ shocks:} & \quad \delta y_{A1}^+ = 5.00, \quad \delta y_{A2}^+ = 2.00s_v, \\
& \quad \delta y_{A1}^- = -5.00, \quad \delta y_{A2}^- = 2.00s_v, \\
B \text{ shocks:} & \quad \delta y_{B1}^+ = 5.00, \quad \delta y_{B2}^+ = 0.00, \\
& \quad \delta y_{B1}^- = -5.00, \quad \delta y_{B2}^- = 0.00, \\
C \text{ shocks:} & \quad \delta y_{C1}^+ = 0.00, \quad \delta y_{C2}^+ = 2.00s_v, \\
& \quad \delta y_{C1}^- = 0.00, \quad \delta y_{C2}^- = -2.00s_v,
\end{align*}

where \( s_v \) is the sample standard deviation of the adjusted volume series.

In this layout, A shocks represent price movements up or down by 5% with a volume increase equal to two standard deviations. Inspect of scatter plots and simple statistics suggest that a 5% price movement accompanied by a two-standard-deviation volume increase is a typically large movement for stock market data. B shocks reflect pure price movements while C shocks reflect pure volume movements. Because A shocks are not simple rotations of B and C shocks, the above experimental design cannot be implemented using orthogonalized shocks.

For reasons of space, we cannot report the results from all nodes on the experimental design. Instead, we selectively report the more interesting impulse responses.

6.2.2. Response to A shocks at the mean

Fig. 3 shows conditional variance profiles for the price process, \( \{ \bar{\gamma}^x_{1i,j}, i=1 \}^{20}, \{ \bar{\gamma}^0_{1i,j}, i=1 \}^{20}, \{ \bar{\gamma}^-_{1i,j}, i=1 \}^{20} \), which are the upper left elements of the \( 2 \times 2 \) conditional variance matrix function \( \bar{\gamma}(x) \) evaluated at \( x^+, x^0, x^- \). Fig. 4 shows the conditional mean profiles for the volume, \( \{ \bar{\gamma}^x_{2i,j}, i=1 \}^{20}, \{ \bar{\gamma}^0_{2i,j}, i=1 \}^{20}, \{ \bar{\gamma}^-_{2i,j}, i=1 \}^{20} \), which are the second elements of the conditional mean profile function \( \gamma(x) \), also evaluated at \( x^+, x^0, x^- \). The conditional mean and variance profile functions are defined in Section 2.2. The baseline case is the mean, \( x^0 = \bar{x} \), and \( x^+ \) and \( x^- \) are the perturbed initial conditions corresponding to \( (\delta y_{A1}^+, \delta y_{A2}^+) \) and \( (\delta y_{A1}^-, \delta y_{A2}^-) \), respectively. The computations are done using the preferred specifications for each of the four bivariate estimations.

With the exception of KO, Fig. 3 generally indicates some evidence for a leverage effect. The effect appears to be small and to die out quickly, though. As for the volume, the responses of the volume to A shocks are quite symmetric and very slowly damped. There is remarkable uniformity across the series in the volume responses.
Fig. 3. Bivariate estimation: Impulse response of $\text{Var}(A_p)$ to A shocks; --- baseline, -------- response to positive price shocks, ------ response to negative price shock.

Fig. 4. Bivariate estimation: Impulse response of $v$ to A stocks; --- baseline, -------- response to positive price shock, ------ response to negative price shock.
6.2.3. Averaged response to B shocks

So far we have conducted the impulse response analysis by examining conditional moment profiles emanating from the unconditional mean of the data. The unconditional mean is arguably a central point in the data cloud for the conditional mean profiles, but it is not for the conditional variance. At the conditioning value \( x = \bar{x} \), the process is unusually quiet and volatility exceptionally low. The fact volatility is unusually low at the mean explains why the baseline volatility in Fig. 3 shows upward drift towards the unconditional variance of the process.

A more reasonable notion of typical impulse response can be obtained by averaging over all possible histories in the data set. To this end, define the averaged conditional mean profiles

\[
y_j^+ = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \hat{y}_j(x_t + \delta x^+),
\]

\[
y_j^0 = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \hat{y}_j(x_t),
\]

\[
y_j^- = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \hat{y}_j(x_t + \delta x^-),
\]

where \( \delta x^+ = (0, \ldots, \delta y^{+T}) \) and \( \delta x^- = (0, \ldots, \delta y^{-T}) \). Likewise, define the averaged conditional variance profiles

\[
y^+ = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \nu(x_t + \delta x^+),
\]

\[
y^0 = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \nu(x_t),
\]

\[
y^- = \frac{1}{(T - L + 1)} \sum_{t=L+1}^{T} \nu(x_t + \delta x^-),
\]

Figs. 5 and 6 show averaged conditional moment profiles. To conserve space we only show those for B shocks, which correspond to pure contemporaneous price movements without an accompanying contemporaneous volume movement; this allows us to show on one set of figures the results from using the different types of shocks (B versus A) and the results from using averaged impulse responses. The main effect of using average impulse responses is to eliminate the baseline drift.

Fig. 5 shows a clear characterization of the leverage effect. The effect is present in three out of four plots, but it is quite small, however, at least in comparison to the magnitudes of the total deviations from baseline. Furthermore, across each of the plots its effects are essentially gone two days after the shock.
Fig. 5. Bivariate estimation: Averaged impulse response of Var(Δp) to B shocks; — baseline, ——— response to positive price shock, --- response to negative price shock.

Fig. 6. Bivariate estimation: Averaged impulse response of v to B shocks; — baseline, ——— response to positive price shock, --- response to negative price shock.
Fig. 6 provides a clear and consistent picture of the response of volume to pure price shocks. Regardless of the sign of the price movement, volume shows a sharp rise one day after the price movement and, with the exception of BA, it shows a very slow decline back towards baseline. In the case of BA, the effect dampens nearly immediately. The directions of the short-term impacts of the price movements on volume are consistent with those reported by Gallant, Rossi, and Tauchen (1993) for the S&P Composite Index. On the other hand, the long-term monotonic decline contrasts with their finding of long-term slowly damped negative response of volume to price shocks.

The fact that the responses of volume to these large positive and negative price shocks is nearly the same in Fig. 6 regardless of the sign of the price shock reflects a nonlinearity detected by the nonparametric procedure. A linear model such as a VAR, or a VAR with ARCH errors, imposes symmetry on the conditional mean profile. The conditional moment profiles for shocks that are symmetric in the sense that \( \delta y^+ = -\delta y^- \), are constrained to be mirror images about the baseline. Since B shocks are symmetric in this way, the mirror image constraint would enforce be seen in VAR or VAR-ARCH impulse response analysis. Here, the polynomial part of the SNP model compensates and produces responses that are not reflections about the baseline.

6.2.4. Confidence bands for impulse responses to B shocks

The preceding analysis suggests two things. First, looking across the four stocks and a variety of shocking strategies and reporting strategies, point estimates indicate that the leverage effect is small and dies out within two or three days of a shock. Second, the dynamic response of volume to price shocks is positive and slowly damped regardless of the sign of the price shock.

Figs. 7 and 8 show 95% confidence bands for the differential effects of positive and negative B shocks. The bands were computed using the bootstrap procedure described in Section 3 above with 500 refittings of the SNP model. Unfortunately, it is at present computationally infeasible to compute such bands for averaged responses, despite the apparent superiority of averaging across many histories. Instead, the bands are computed for the case where the baseline is the mean, \( \mu^0 = \bar{x} \).

Fig. 7 shows that the size of the leverage effect is small relative to the statistical noise in estimation. In short, the leverage effect is insignificant at the 95% level, because the bands always include the null profile of zero everywhere.

Fig. 8 shows a 95% sup-norm confidence band about the point estimate of \( \{ \hat{y}_{2,i}^+ - \hat{y}_{2,i}^- \}_{i=1}^{20} \). The results show that the difference between the positive and negative volume responses always include the null profile. One cannot reject the null hypothesis that the differential effects are the same, which is evidence in support of nonlinearity over and above ARCH.
Fig. 7. Bivariate estimation: 95% confidence bands for the impulse responses of $\text{Var}(\Delta p)$ to positive B shocks minus the impulse responses to negative B shocks.

Fig. 8. Bivariate estimation: 95% confidence bands for the impulse responses of $\nu$ to positive B shocks minus the impulse responses to negative B shocks.
We also computed confidence estimates of effects relative to baseline. Specifically, we computed 95\% bands about \(\{\hat{\gamma}_{11, j} - \hat{\gamma}_{11, j}^{0}\}_{j=1}^{20}\), \(\{\hat{\gamma}_{12, j} - \hat{\gamma}_{12, j}^{0}\}_{j=1}^{20}\), \(\{\hat{\gamma}_{21, j} - \hat{\gamma}_{21, j}^{0}\}_{j=1}^{20}\), and \(\{\hat{\gamma}_{22, j} - \hat{\gamma}_{22, j}^{0}\}_{j=1}^{20}\). Space constraints preclude reporting all of these plots. In nearly every case, however, the point estimates of deviations relative to baseline exclude (or just barely include) the null profile in the first few days, indicating statistical significance (or nearly so) at the 95\% level.

6.3. Higher-order nonlinearities

Hiemstra and Jones (1994) investigate nonlinear Granger-causal relationships between aggregate volume and stock prices. Their test statistics indicate that volume nonlinearly Granger-causes price movements. These findings stand in contrast to those of Gallant, Rossi, and Tauchen (1993), who find no evidence for feedback from volume shocks to either the conditional mean or volatility of stock prices. In additional work (not reported) we were able to replicate the GRT results for the four stocks considered here.

Reconciliation of these sets of findings entails examination of feedback from volume shocks to higher-order moments of stock prices. To investigate the impact of volume shocks on the skewness of stock prices, we take \(g(\cdot)\) to be the conditional skewness function of the price change:

\[
g(y_{t-L+1}, \ldots, y_{t-1+j}) = \mathcal{E}[\{y_{t+1+j} - \mathcal{E}(y_{t+1+j} | y_{t-L+1}, \ldots, y_{t-1+j})\}^3 | y_{t-L+1}, \ldots, y_{t-1+j}].
\]

Recall that \(y_t = (\Delta p_t, v_t)'\). Given this \(g(\cdot)\), the skewness conditional moment profile function is

\[
\hat{\gamma}_3(x) = \mathcal{E}[g(y_{t-L+1}, \ldots, y_{t-1+j}) | x_t = x], \quad \text{for} \quad j = 1, 2, \ldots.
\]

The impulse responses of the skewness to pure volume shocks (C shocks) are traced out in exactly the manner described in Section 2.2 above, with Monte Carlo integration used to compute conditional expectations at \(x = \bar{x}\).

Likewise, for the conditional kurtosis, we take \(g(\cdot)\) to be

\[
g(y_{t-L+1}, \ldots, y_{t-1+j}) = \mathcal{E}[\{y_{t+1+j} - \mathcal{E}(y_{t+1+j} | y_{t-L+1}, \ldots, y_{t-1+j})\}^4 | y_{t-L+1}, \ldots, y_{t-1+j}].
\]

We define the kurtosis conditional moment profile function

\[
\hat{\gamma}_4(x) = \mathcal{E}[g(y_{t-L+1}, \ldots, y_{t-1+j}) | x_t = x], \quad \text{for} \quad j = 1, 2, \ldots
\]

and proceed in the same manner.

The two panels of Fig. 9 show the impulse responses of the conditional skewness and kurtosis to volume shocks for Boeing (BA). Volume shocks are seen
to cause the conditional skewness and kurtosis to deviate from baseline, but the effects dissipate in four to six days. These plots are representative. For the cases of IBM and KO (not shown), the impulse responses are more attenuated and damped. For MMM (also not shown), the impulse response of the conditional skewness is more slowly damped than in Fig. 9 while the kurtosis is more heavily damped. Taken together, our findings, along with those of Gallant, Rossi, and Tauchen (1993) and Hiemstra and Jones (1994) reveal that volume nonlinearly Granger-causes stock prices, but the effects are concentrated in the higher-order moments (above two) and are transient.

Finally, we investigate the extent to which the local time path of volume affects the serial correlation properties of stock price changes. This effort is motivated by the evidence reported in LeBaron (1992b) that the stock index appears more
Fig. 10. Bivariate estimation: Impulse response to \( \Delta p \) to rising volume (top left) and falling volume (top right); volume paths shown in bottom two panels; \( \cdots \cdots \) baseline, \( \cdots \cdots \cdots \) response to positive price shock, \( \cdots \cdots \) response to negative price shock.

persistent on rising than falling volume. For BA, the top left panel of Fig. 10 shows the impulse response of the stock price change to its own shock on rising volume. The time path of rising volume, shown in the lower panel, starts at two standard deviations below the mean four days before the price shock, and it rises linearly to two standard deviations above the mean on the day of the price shock. Similarly, the right-hand side of the figure shows the analogous impulse response on falling volume. Comparing the two top panels of these two figures reveals little evidence for the local path of volume affecting the serial correlation properties of the BA stock price series. The plots are similar for the other three stocks in our data set. The evidence indicates that volume-persistence relations are not characteristics of actively traded individual securities like our four stocks.

7. Conclusion

Our nonparametric, nonlinear analysis leads to sharp empirical findings regarding the multi-step-ahead dynamics of the four daily stock and volume series. First, each stock price displays persistent response of volatility to large price shocks. However, consistent with recent evidence on fractionally integrated volatility, the degree of persistence in volatility is less than that detected in broad stock price
indexes. Second, the asymmetric leverage effect is a characteristic of the point estimates, but it is quite small compared with the direct effects of the price shocks on volatility and is barely statistically significant. Also, the leverage effect is extremely rapidly damped, and in particular dies out within two or three days after a price shock, while the effect of the price shock itself takes weeks to damp out. Third, volume responds nonlinearly to price shocks, and in a way that could not be detected using standard symmetric parametric models—volume increases whatever the sign of the price shock and then damps very slowly back to baseline. Fourth, volume shocks affect stock prices, but the effects are very transient and confined to the higher-order (above the second) conditional moments, with basically no effects on the first two conditional moments. Fifth, in contrast to some recent evidence based on broad market indexes, the history of volume appears not to influence the serial correlation properties of our four stock price series.

These findings are quite robust across the four securities used in the panel. The first, fourth, and fifth findings are new to this study; the second and third are extensions of results obtained in Gallant, Rossi, and Tauchen (1993).

References

Andersen, Torben, 1995, Return volatility and trading volume: An information flow interpretation of stochastic volatility, Unpublished manuscript (Northwestern University, Evanston, IL).
Doan, Thomas, Robert Litterman, and Christopher Sims, 1984, Forecasting and conditional projection using realistic prior distributions, Econometric Reviews 3, 1–100.

Fenton, Victor M. and A. Ronald Gallant, 1995, Qualitative and asymptotic performance of SNP density estimators, Unpublished manuscript (University of North Carolina, Chapel Hill, NC).


Gallant, A. Ronald and George Tauchen, 1994, Which moments to match?, Unpublished manuscript (Duke University, Durham, NC).


LeBaron, Blake, 1992a, Forecast improvements using a volatility index, Journal of Applied Econometrics 7 (Suppl.), s137-s149.

LeBaron, Blake, 1992b, Persistence of the Dow Jones index on rising volume, Working paper (Department of Economics, University of Wisconsin, Madison, WI).


Nelson, Daniel, 1994, Asymptotic filtering and smoothing theory for multivariate ARCH models, Unpublished manuscript (University of Chicago, Chicago, IL).


Potter, Simon M., 1994, Nonlinear impulse-response functions, Unpublished manuscript (University of California, Los Angeles, CA).

