

## NONLINEAR DYNAMIC STRUCTURES

BY A. RONALD GALLANT, PETER E. ROSSI, AND GEORGE TAUCHEN<sup>1</sup>

The paper develops an approach for analyzing the dynamics of a nonlinear time series that is represented by a nonparametric estimate of its one-step ahead conditional density. The approach entails examination of conditional moment profiles corresponding to certain shocks; a conditional moment profile is the conditional expectation evaluated at time  $t$  of a time invariant function evaluated at time  $t + j$  regarded as a function of  $j$ . Comparing the conditional moment profiles to baseline profiles is the nonlinear analog of conventional impulse-response analysis. The approach includes strategies for laying out realistic perturbation experiments in multivariate situations and for undertaking statistical inference using bootstrap methods. It also includes examination of profile bundles for evidence of damping or persistence. The empirical work investigates a bivariate series comprised of daily changes in the Standard and Poor's composite price index and daily NYSE transactions volume from 1928 to 1987. The effort uncovers evidence showing the heavily damped character of the "leverage effect" and the differential response (short-term increase, long-term decline) of trading volume to "common-knowledge" price shocks.

KEYWORDS: Impulse-response, nonlinear time series, nonparametric, financial market volatility, trading volume.

### 1. INTRODUCTION

THE 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. However, the conditional density is a complicated nonlinear function of a large number of arguments and is difficult to interpret. Thus, it is important to develop methods to elicit the dynamics embodied in the conditional density.

For a linear process with homogeneous errors, there is a comprehensive tool kit of methods for exploring the dynamics of a process and comparing them to the predictions of an economic model. A central component of this tool kit is the impulse-response or "error shock" methodology put forth in Sims (1980) and refined by Doan, Litterman, and Sims (1984) and others. The key idea of impulse-response analysis is to trace through the system the effects of small movements in the innovations, or linear combinations of the innovations. There are various graphical and numerical techniques for tracing through the effects of innovations. These techniques provide a means for exploring the characteristics of the conditional density, which can be quite complicated even for linear processes.

There are some extensions of these ideas to volatility. Using a recursive system of GARCH models, Engle, Ito, and Lin (1990), perform an error shock

<sup>1</sup> Research supported by National Science Foundation Grants SES-8808015, SES-8810357, SES-9023083, North Carolina Agricultural Experiment Station Project NCO-6134, and the PAMS Foundation; we thank two referees and many seminar participants for their helpful comments.

analysis of the effects of price shocks on subsequent volatility. GARCH models have additive innovations defined by a linear difference equation.

In this paper, we outline a strategy for performing an impulse-response analysis of nonlinear time series models. For the general nonlinear model, it seems less useful to think in terms of an innovation. Rather, the dynamic properties of the nonlinear model appear best 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 forecasts of the conditional mean and conditional variance functions. The method can also be applied to general functions, such as those used to study turning points.

If the time series is multivariate, then it is important to design a series of perturbation experiments that take into account the contemporaneous relationships between series. It may be unrepresentative to perturb one of the variables in the conditioning set without simultaneously adjusting the values of others. Because the interpretation of a direct perturbation to conditioning variables is more transparent than perturbing the errors of a linear system, it can be easier to lay out a representative experimental design in the space of conditioning variables than in the space of additive errors to a linear system.

Nonlinear impulse-response analysis, as described above, involves a comparison of a conditional moment profile to a baseline profile. A conditional moment profile is the forecast made at time  $t$  of the time  $t + j$  value of a time-invariant function regarded as a function of  $j$ . Equivalently, a conditional moment profile is the conditional expectation evaluated at time  $t$  of a time-invariant function evaluated at time  $t + j$  regarded as a function of  $j$ . Other conditional moment profiles are of interest. In particular, bundles of long-term profiles run out from a subsequence of data points from the sample can be examined for evidence of persistence. The statistical significance of a conditional moment profile can be assessed by comparing a sup-norm confidence band on the profile with a null profile.

The methods we propose require a means to sample the estimated conditional density efficiently. Having an efficient algorithm to simulate a sample path, a conditional moment profile can be obtained by running a time-invariant function out over many simulated sample paths and then averaging. Bootstrap estimates, which are used to compute sup-norm confidence bounds on profiles, are obtained by simulating the sample path of the entire data and re-estimating the density.

We utilize these ideas to study the dynamics of daily price and volume movements on the NYSE. We employ the conditional density that was estimated via the SNP nonparametric technique in Gallant, Rossi, and Tauchen (1992) for the period 1928 to 1987. This paper examines multi-step ahead dynamics, while the former concentrates exclusively on one-step ahead characteristics. We uncover evidence that the so-called "leverage effect" (Nelson (1991)) is a weak transient phenomenon that dies away within six to ten days after a price shock. We also uncover evidence indicating that volume responds

to price shocks much differently in the short run than in the long run. Interestingly, this evidence would not have been available from fitting a linear VAR model with ARCH errors. A VAR-ARCH model imposes symmetry on the response of volume to price shocks, and this symmetry directly conflicts with the intrinsic nature of the price-volume relationship.

The organization of the paper is as follows. Section 2 outlines a nonlinear impulse-response methodology and discusses computations. Section 3 specializes the concepts to an AR(1) model with ARCH and GARCH errors to provide a comparison with existing methods for linear difference equations. Section 4 discusses strategies for defining representative impulse-response sequences. Section 5 presents estimates of the impulse-response analysis for stock price and volume data. Section 6 contains confirmatory analysis. Section 7 investigates long term persistence in volatility and volume. Section 8 contains concluding remarks.

2. IMPULSE-RESPONSE ANALYSIS OF NONLINEAR MODELS

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. Denote the  $L$  lags of  $y_{t+1}$  by  $x_t = (y'_{t-L+1}, \dots, y'_t) \in R^{ML}$  and write  $f(y_{t+1}|x_t)$  for the (one-step ahead) conditional density. Due to the strict stationarity assumption, the functional form  $f(y|x)$  of the conditional density does not depend on the index  $t$ ; that is, the density is time invariant.

In this section we shall describe strategies for eliciting the dynamics of the process  $\{y_t\}$  as represented by  $f(y|x)$ . To provide a familiar framework, we first summarize VAR error-shock analysis. We then discuss conditional mean profiles, which are closely related to VAR error shock analysis. Next, the ideas are extended to conditional volatility profiles, which are of particular interest to the finance literature. In the remaining subsections we discuss general conditional moment profiles, computational issues, sup-norm confidence bands, and profile bundles.

Because it is critically important to differentiate between one-step ahead and multi-step ahead conditional moments, the terminology throughout the remainder of the paper adheres to the following conventions. One-step mean is short for the "one-step ahead forecast of the mean conditioned on the history of the process" which is  $\mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty})$  in general or  $\mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{L-1})$  for a Markovian process as above. Similarly, the one-step variance, also called the volatility, is the one-step ahead forecast of the variance conditioned on history; that is

$$\begin{aligned} \text{Var}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty}) &= \mathcal{E}\left\{ \left[ y_{t+1} - \mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty}) \right] \right. \\ &\quad \left. \times \left[ y_{t+1} - \mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty}) \right]' \middle| \{y_{t-k}\}_{k=0}^{\infty} \right\} \end{aligned}$$

or  $\text{Var}(y_{t+1}|\{y_{t-k}\}_{k=0}^{L-1})$  for a Markovian process. With respect to a conditional

moment profile,

$$\mathcal{E} \left[ g(y_{t+j-J}, \dots, y_{t+j}) \mid \{y_{t-k}\}_{k=0}^{L-1} \right] \quad (j = 0, 1, 2, \dots),$$

the word moment refers to the time-invariant function  $g(y_{-j}, \dots, y_0)$ . Thus, the term conditional mean profile is short for “the conditional moment profile of the one-step mean” and conditional volatility profile is short for “the conditional moment profile of the one-step variance.”

### 2.1. VAR Error-Shock Analysis

Since Sims’ (1980) paper, impulse-response functions have been widely used to study the dynamics of a linear process. Briefly, the ideas are as follows.

Suppose

$$A(\mathcal{L})y_t = u_t$$

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  $\mathcal{E}u_t = 0$  and variance matrix  $\text{Var}(u_t) = \mathcal{E}u_t u_t' = \Omega$ . Suppose also that  $A(\mathcal{L})$  is invertible so that

$$y_t = B(\mathcal{L})u_t$$

where  $B(\mathcal{L}) = [A(\mathcal{L})]^{-1} = \sum_{k=0}^{\infty} B_k \mathcal{L}^k$ . Denote the  $ij$ th element of  $B(\mathcal{L})$  by  $B_{ij}(\mathcal{L}) = \sum_{k=0}^{\infty} b_{ijk} \mathcal{L}^k$ . The sequence  $\{b_{ijk}\}_{k=0}^{\infty}$  can be viewed as the dynamic response of the  $i$ th variable to a positive, one-unit movement in the  $j$ th element of the innovation vector  $u_t$ .

The sequence  $\{b_{ijk}\}_{k=0}^{\infty}$  can be computed as follows: Put  $\hat{y}_t = 0$  for negative  $t$ . Put  $\hat{y}_0 = \xi_j$  for  $t = 0$ , where  $\xi_j$  is a vector with 1 in the  $j$ th position and 0 in the others. Iterate the difference equation  $\hat{y}_t = \sum_{s=1}^L A_s \hat{y}_{t-s}$  over positive  $t$ . The  $i$ th element of the sequence  $\{\hat{y}_t\}_{t=0}^{\infty}$  is the sequence  $\{b_{ijk}\}_{k=0}^{\infty}$ . This method is a well known computational technique for obtaining the  $ij$ th element of  $[A(\mathcal{L})]^{-1}$ .

For later reference, note that due to linearity one could also compute  $\{\hat{y}_t\}_{t=1}^{\infty}$  as follows: Let  $\{\hat{y}_t^0\}_{t=-\infty}^{\infty}$  be an arbitrary sequence. Set  $\hat{y}_0^+ = \hat{y}_0^0 + \xi_j$  for  $t = 0$  and  $\hat{y}_t^+ = \hat{y}_t^0$  for negative  $t$ . Iterate the difference equations  $\hat{y}_t^0 = \sum_{s=1}^L A_s \hat{y}_{t-s}^0$  and  $\hat{y}_t^+ = \sum_{s=1}^L A_s \hat{y}_{t-s}^+$  for positive  $t$ . Put  $\hat{y}_t = \hat{y}_t^+ - \hat{y}_t^0$ .

Also for later reference, note that  $\hat{y}_t^+$  is the conditional mean of the process  $\{y_t\}_{t=1}^{\infty}$  with initial conditions  $\{\hat{y}_t^+\}_{t=-\infty}^{\infty}$ . By the law of iterated expectations,  $\hat{y}_t^+$  is also the forecast of the one-step mean  $\mathcal{E}(y_{t+1} \mid \{y_{t-j}\}_{j=0}^{\infty})$  given initial conditions  $\{\hat{y}_t^+\}_{t=-\infty}^{\infty}$ .

In applied work, the contemporaneous covariance matrix  $\Omega$  of the linear system is usually not diagonal. In this case, a perturbation of one unit in  $u_{jt}$  holding the other elements of  $u_t$  constant is not considered representative of the typical shocks that impinge on the system. Common practice in the literature is to restrict attention to orthogonalized shocks. An orthogonalization is obtained by choosing a lower triangular matrix  $H$  such that  $H\Omega H' = I$ , and writing  $y_t = B^*(\mathcal{L})u_t^*$ , where  $B^*(\mathcal{L}) = B(\mathcal{L})H^{-1}$ , and  $u_t^* = Hu_t$ . The  $ij$ th element of  $B^*(\mathcal{L})$  represents the response of variable  $i$  to orthogonalized

innovation  $u_{jt}^*$ , which is a linear combination of the elements  $u_t$ . As is well known, the orthogonalization of errors by factorization of  $\Omega$  is not unique, making it difficult to provide an economic interpretation of shocks to the orthogonalized errors. This identification problem has led to a large literature on “structural VAR” analysis using both short-run restrictions (Blanchard and Watson (1986)), and long-run restrictions (Blanchard and Quah (1989); King, Plosser, Stock, and Watson (1991)) to identify the structural rather than-reduced-form errors.

In the general nonlinear case, there are various notions of an innovation. One could center  $y_t$  on a conditional mean or median and leave it unscaled or scale by the conditional variance or range. It may be difficult to compute an impulse response for any of these notions of an innovation. However, if instead of viewing perturbation of  $u_t$  as the primitive concept, the computational technique of perturbing  $y_t$  is viewed as the primitive, then the ideas from linear VARs extend directly as we shall see in the subsections that follow. Regardless of whether impulse responses are computed by direct perturbation or through innovations, a structural model is still needed to link shocks to endogenous variables back to shocks in underlying exogenous or policy variables. Also, there are complications related to the task of obtaining realistic perturbations, especially for multivariate data. These issues are addressed more fully in Section 4 and Subsection 5.2. Our strategy uses graphical methods to develop an experimental design for the shocks.

### 2.2. Conditional Moment Profiles—Means

Consider the general case in which  $\{y_t\}$  is a stationary process represented by the one-step ahead conditional density  $f(y|x)$ . Define the conditional mean profile  $\{\hat{y}_j(x)\}_{j=0}^\infty$  corresponding to initial condition  $x$  by

$$\begin{aligned} \hat{y}_j(x) &= \mathcal{E}(y_{t+j}|x_t = x) \\ &= \int y f^j(y|x) dy \end{aligned}$$

where  $f^j(y|x)$  denotes the  $j$ -step ahead conditional density

$$f^j(y|x) = \int \cdots \int \left[ \prod_{i=0}^{j-1} f(y_{i+1}|y_{i-L+1}, \dots, y_i) \right] dy_1 \cdots dy_{j-1}$$

with  $x = (y'_{-L+1}, \dots, y'_0)'$ . (If a dummy variable of integration coincides with an element of  $x$ , that integration is omitted.)

In empirical work,  $f^j(y|x)$  is approximated by using a nonparametric estimate  $\hat{f}(y|x)$  in place of  $f(y|x)$ . Given an efficient algorithm for sampling  $\hat{f}(y|x)$ ,  $\hat{y}_j(x)$  is easily computed using Monte Carlo integration as discussed in Subsection 2.5 below.

Recall the interpretation of  $x = (y'_{-L+1}, y'_{-L+2}, \dots, y'_0)'$ :  $y_0 \in R^M$  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 small perturbations to the contemporaneous  $y_0$  where  $\delta y^+$  is “positive” and  $\delta y^-$  is “negative” in a sense to be made precise later. Put

$$x^+ = (y'_{-L+1}, y'_{-L+2}, \dots, y'_0)' + (0, 0, \dots, \delta y^+)',$$

$$x^0 = (y'_{-L+1}, y'_{-L+2}, \dots, y'_0)',$$

$$x^- = (y'_{-L+1}, y'_{-L+2}, \dots, y'_0)' + (0, 0, \dots, \delta y^-)'$$

Thus,  $x^+$  is an initial condition corresponding to a positive impulse or shock  $\delta y^+$  added to contemporaneous  $y_0$ ,  $x^-$  corresponds to a negative impulse, while  $x^0$  represents the base case with no impulse.

Now put

$$\hat{y}_j^+ = \hat{y}_j(x^+),$$

$$\hat{y}_j^0 = \hat{y}_j(x^0),$$

$$\hat{y}_j^- = \hat{y}_j(x^-),$$

for  $j = 0, 1, 2, \dots$ . The conditional mean profiles  $\{\hat{y}_j^+\}_{j=0}^\infty$ ,  $\{\hat{y}_j^0\}_{j=0}^\infty$ , and  $\{\hat{y}_j^-\}_{j=0}^\infty$ , are forecasts of subsequent  $y_{t+j}$  for each of these three initial  $x$  values. The profile  $\{\hat{y}_j^0\}_{j=0}^\infty$  is the baseline forecast.

A natural definition of the nonlinear impulse response is the net effect of the impulse  $\delta y^+$  (or  $\delta y^-$ ). The net effect is obtained by comparing the profile for  $\delta y^+$  (or  $\delta y^-$ ) to the baseline. Specifically the sequence,  $\{\hat{y}_j^+ - \hat{y}_j^0\}_{j=0}^\infty$  represents the net response to the positive impulse while  $\{\hat{y}_j^- - \hat{y}_j^0\}_{j=0}^\infty$  represents the net response to the negative impulse. The impulse responses depend upon the initial  $x$ , which reflects the nonlinearities of the system. A similar development of a nonlinear shocking strategy is undertaken independently in Potter (1991), who develops various theoretical concepts of nonlinearity and persistence of time series. Our approach below places more emphasis on the complications generated by defining “typical” shocks for multivariate data, on impulse responses for higher-order functions of the data (for example, volatility), and on methods of statistical inference.

Tracing out the impulse responses of a nonlinear system in the way just described is the exact analogue of VAR error-shock analysis, as was noted in the remarks regarding computations in Subsection 2.1.

For multivariate nonlinear impulse-response analysis, the same issues arise as in the linear case regarding the contemporaneous correlation structure among the variables. To be concrete, suppose  $\delta y^+$  contains a perturbation of unity for the first variable, so one is tracing out the effects of a shock in the first variable. For the impulse responses to be realistic, the remaining elements of  $\delta y^+$  should be adjusted to take account of contemporaneous covariance. One possible way to do this would be to enter for the remaining elements their predicted values given the movement in the first variable. These can be calculated either from

linear projections, as in Sims (1980), or from the conditional expectations implied by the fitted density of the data.

An alternative, which is used in Subsection 5.2 below, is to inspect a scatter plot of the data cloud  $\{y_t\}$  and visually determine shocks  $\delta y^+$  and  $\delta y^-$  that appear typical relative to the historical dispersion of the data. This strategy is available for two or three dimensional data, but, of course, will not work directly for four or more dimensions. An interesting topic for future research is exploration of the feasibility of using cluster analysis to determine reasonable shocks from higher dimensional point clouds. Of particular interest might be the robust clustering strategies of Liu (1990).

### 2.3. Conditional Moment Profiles—Volatility

So far we have only considered tracing out the effects of shocks on the means of subsequent  $y$ 's. But there are macro and finance applications where one is interested not only in the effects of shocks on the means of subsequent  $y$ 's but also the effects on subsequent volatility. In financial applications, for instance,  $y$  is a price change, which is nearly unpredictable, but large price swings have strong implications for subsequent volatility (see Nelson (1991), Bollerslev and Engle (1993), and the references therein). In macro applications, one might be interested in the effects of monetary disturbances on subsequent output volatility.

In the familiar case of a model that has a notion of an innovation, such as a VAR,

$$A(L)y_t = u_t,$$

volatility is defined as the one-step variance of the innovation  $u_t$ . This is the concept usually applied to ARCH and GARCH specifications of  $\{u_t\}$ . The one-step variance of  $u_t$  is, of course, the same as the one-step variance of  $y_t$ , which is computed as

$$\text{Var}(y|x) = \int [y - \mathcal{E}(y|x)][y - \mathcal{E}(y|x)]' f(y|x) dy,$$

$$\mathcal{E}(y|x) = \int yf(y|x) dy.$$

It is of interest to trace out the effects of a shock on subsequent volatility by extending the analysis above. To do so, note that the law of iterated expectations disguised the fact that it is actually the effect of a shock on the one-step mean  $\mathcal{E}(y|x)$  of  $y$  that is traced out in Subsection 2.2. That is,  $\hat{y}_j(x) = \mathcal{E}[\mathcal{E}(y_{t+j}|x_{t+j-1})|x_t = x]$  and a conditional mean profile is actually a forecast of the one-step mean of  $y$  given contemporaneous  $x$ . The extension to volatility is immediate.

Define

$$\begin{aligned} \hat{v}_j(x) &= \mathcal{E}[\text{Var}(y_{t+j}|x_{t+j-1})|x_t = x] \\ &= \int \cdots \int \text{Var}(y_j|y_{j-1-L}, \dots, y_{j-1}) \\ &\quad \times \left[ \prod_{i=0}^{j-1} f(y_{i+1}|y_{i-L+1}, \dots, y_i) \right] dy_1 \cdots dy_j \end{aligned}$$

for  $j = 1, 2, \dots$  where  $x = (y'_{-L+1}, \dots, y'_0)'$ . (If a dummy variable of integration coincides with an element of  $x$ , that integration is omitted.)  $\hat{v}_j(x)$  is the forecast of the one-step variance (or variance matrix)  $j$  steps ahead, conditional on  $x_t = x$ .

The analysis proceeds as before.  $x^0$  defines baseline initial conditions,  $x^+$  corresponding to a positive impulse  $\delta y^+$ , and  $x^-$  a negative impulse  $\delta y^-$ . The net effect of an impulse is assessed by plotting its profile relative to the baseline. Examples are presented in Subsections 5.1 and 5.2.

The conditional volatility profile,  $\{\hat{v}_j(x)\}_{j=1}^\infty$ , is different from the path described by the  $j$ -step ahead mean square error of the process which is defined by

$$\hat{\mathcal{M}}_j(x) = \text{Var}(y_{t+j}|x_t = x).$$

The contrast is best seen by writing

$$\begin{aligned} \hat{v}_j(x) &= \mathcal{E}\left\{ \left[ y_{t+j} - \mathcal{E}(y_{t+j}|x_{t+j-1}) \right] \left[ y_{t+j} - \mathcal{E}(y_{t+j}|x_{t+j-1}) \right]' \middle| x_t = x \right\}, \\ \hat{\mathcal{M}}_j(x) &= \mathcal{E}\left\{ \left[ y_{t+j} - \mathcal{E}(y_{t+j}|x_t) \right] \left[ y_{t+j} - \mathcal{E}(y_{t+j}|x_t) \right]' \middle| x_t = x \right\}. \end{aligned}$$

Thus  $\hat{v}_j(x)$  and  $\hat{\mathcal{M}}_j(x)$  are seen to differ only in the centering for a mean square error computation. The conditional volatility profile is centered at  $\mathcal{E}(y_{t+j}|x_{t+j-1})$ , while the  $j$ -step ahead mean square error is centered at  $\mathcal{E}(y_{t+j}|x_t)$ .

Which concept is of primary interest depends upon the character of the application. The conditional volatility profile appears to be the one better suited for studying the structure of the second moment properties of the process separately from first moment properties. This intuition is based on examples from parametric models (see Section 3) and the identity

$$\begin{aligned} \hat{\mathcal{M}}_j(x) &= \hat{v}_j(x) + \sum_{k=1}^{j-1} \mathcal{E}\left\{ \left[ \mathcal{E}(y_{t+j}|x_{t+k}) - \mathcal{E}(y_{t+j}|x_{t+k-1}) \right] \right. \\ &\quad \left. \times \left[ \mathcal{E}(y_{t+j}|x_{t+k}) - \mathcal{E}(y_{t+j}|x_{t+k-1}) \right]' \middle| x_t = x \right\}. \end{aligned}$$

Thus,  $\hat{\mathcal{M}}_j(x)$  is a confounding of  $\hat{v}_j(x)$  with the variability of the conditional expectation of  $y_{t+j}$  as information accumulates between times  $t$  and  $t+j-1$ . This confounding comingles first and second moment characteristics of  $\{y_t\}$ , which complicates the task of understanding the second moment properties of the process.



On the other hand, for analyzing the properties of forecasts of  $y_{t+j}$  given the history of the process up through time  $t$ ,  $\hat{\mathcal{H}}_j(x)$  would be the appropriate concept. This type of analysis, while important in some applications, is of secondary interest in this paper.

2.4. *Conditional Moment Profiles—General Functions*

The extension of the preceding analysis, which only considers the first two conditional moments of  $y_t$ , to general, possibly nonlinear, functions is straightforward.

Let  $g(y_{-j}, y_{-j+1}, \dots, y_0)$  denote a time-invariant function of a stretch of  $y$ 's of length  $J + 1$ . Put

$$\begin{aligned} \hat{g}_j(x) &= \mathcal{E} \left[ g(y_{t+j-j}, \dots, y_{t+j}) \mid x_t = x \right] \\ &= \int \cdots \int g(y_{j-j}, \dots, y_j) \left[ \prod_{i=0}^{j-1} f(y_{i+1} \mid y_{i-L+1}, \dots, y_i) \right] dy_1 \cdots dy_j \end{aligned}$$

for  $j = 0, 1, \dots$  where  $x = (y'_{-L+1}, \dots, y'_0)'$ . (If a dummy variable of integration coincides with an element of  $x$ , that integration is omitted.)

The profiles  $\{\hat{g}_j(x^+)\}_{j=0}^\infty$ ,  $\{\hat{g}_j(x^0)\}_{j=0}^\infty$ , and  $\{\hat{g}_j(x^-)\}_{j=0}^\infty$  are the forecasts of  $g$ 's starting from the initial conditions  $x^+$ ,  $x^0$ , and  $x^-$ , respectively. Profiles compared to baseline

$$\begin{aligned} &\{\hat{g}_j(x^+) - \hat{g}_j(x^0)\}_{j=0}^\infty, \\ &\{\hat{g}_j(x^-) - \hat{g}_j(x^0)\}_{j=0}^\infty, \end{aligned}$$

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

This general setup subsumes a wide variety of cases. By suitably defining the function  $g$ , it covers the earlier cases where the impulse responses are defined for the one-step means and variances. Another potential application is turning point analysis. Under one notion of a downturn, the function

$$g(y_{t-3}, y_{t-2}, y_{t-1}, y_t) = I(y_{t-2} \geq y_{t-3}, y_{t-1} < y_{t-2}, y_t < y_{t-1}),$$

where  $I(\cdot)$  is the zero-one indicator function, takes the value unity if a downturn occurs between  $t - 3$  and  $t$ . Hence, a forecast of  $g(y_{t-3+j}, y_{t-2+j}, y_{t-1+j}, y_{t+j})$ ,  $j \geq 1$ , is the conditional probability of a downturn between  $t - 3 + j$  and  $t + j$ . Examination of the impact that shocks to  $y_t$  have on these conditional probabilities can possibly provide insight in the character of business cycle fluctuations and, in particular, insight into the asymmetric character of business cycles.

### 2.5. Computations

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

Let  $\{y_j^r\}_{j=1}^\infty$ ,  $r = 1, 2, \dots, R$ , denote  $R$  simulated realizations of the process starting from  $x_0 = x$ . In other words,  $y_1^r$  is a random drawing from  $f(y|x)$  with  $x = (y'_{-L+1}, \dots, y'_{-1}, y'_0)'$ ;  $y_2^r$  is a random draw from  $f(y|x)$ , with  $x = (y'_{-L+2}, \dots, y'_0, y_1^r)'$ , and so forth.

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

$$\hat{g}_j(x) = \int \cdots \int g(y_{j-J}, \dots, y_j) \left[ \prod_{i=0}^{j-1} f(y_{i+1} | y_{i-L+1}, \dots, y_i) \right] dy_1 \cdots dy_j$$

$$\doteq \frac{1}{R} \sum_{r=1}^R g(y_{j-J}^r, \dots, y_j^r)$$

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

To illustrate, a nonparametric estimator of the one-step ahead conditional density of a nonlinear, stationary time series is the kernel estimate (Robinson (1983))

$$\hat{f}(y|x) = \frac{\sum_{t=1}^n K_2(y - y_t, x - x_t)}{\sum_{t=1}^n K_1(x - x_t)}$$

$$= \frac{\sum_{t=1}^n K_1(x - x_t) K_2(y - y_t | x - x_t)}{\sum_{t=1}^n K_1(x - x_t)}$$

where  $K_2(y, x)$  is a multivariate density function such as the normal,  $K_1(x) = \int K_2(y, x) dy$ , and  $K_2(y|x) = K_2(y, x) / \int K_2(y, x) dy$ . To sample  $y$  from this density for given  $x$ , randomly select  $t$ ,  $1 \leq t \leq n$ , with probability  $K_1(x - x_t) / \sum_{t=1}^n K_1(x - x_t)$  and then sample  $y$  from the density  $p(y) = K_2(y - y_t | x - x_t)$ .

### 2.6. Sup-Norm Confidence Bands

The significance of a profile (or of a linear combination of profiles) may be assessed by comparing its sup-norm confidence band with a null profile. A null profile describes a null response to an impulse and would usually be a horizontal line through zero or some other unconditional moment. If the band includes the null profile, the effect of the impulse is judged insignificant.

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

$\hat{f}(y|x)$ ; 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  $\hat{f}(y|x)$  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  $\hat{f}(y|x)$  that is just wide enough to contain 95% of the simulated profiles.

To illustrate, consider setting a sup-norm band on a difference between conditional volatility profiles of the first element  $y_{1t}$  of  $y_t$ ,

$$\mathcal{N}_j = \mathcal{E}[\text{Var}(y_{1j}|x_{j-1})|x^+] - \mathcal{E}[\text{Var}(y_{1j}|x_{j-1})|x^-] \quad (j = 1, \dots, J),$$

as in Subsection 6.3. The computations proceed as follows. Let  $\{y_t^r\}_{t=L+1}^n$  denote the  $r$ th simulated data set from the conditional density  $\hat{f}(y|x)$ ,  $r = 1, \dots, R$ , computed as described in Subsection 2.5. In these simulations, the initial conditions  $x = (y'_1, \dots, y'_L)'$  are held the same for each  $r$  whereas the seed of the random number generator is reset randomly for each  $r$ . Let  $\hat{f}^r(y|x)$  denote a nonparametric estimate of  $f(y|x)$  fitted to  $\{y_t^r\}$ . Compute

$$M^r = \max_{1 \leq j \leq J} |\hat{\mathcal{N}}_j^r - \hat{\mathcal{N}}_j|$$

for each  $r$  where  $\hat{\mathcal{N}}_j^r$  is computed from  $\hat{f}^r(y|x)$  as described in Subsection 2.3. Lastly, letting  $M^{0.95}$  denote the 0.95 quantile of the  $\{M^r\}_{r=1}^R$ , the sup-norm confidence band on  $\{\hat{\mathcal{N}}_j: j = 1, \dots, J\}$  is

$$\hat{\mathcal{N}}_j \pm M^{0.95} \quad (j = 1, 2, \dots, J).$$

To see this, note that for probabilities computed using  $\hat{f}(y|x)$

$$\begin{aligned} & \Pr \left[ (\mathcal{N}_1^r, \dots, \mathcal{N}_J^r) \in (\hat{\mathcal{N}}_1 \pm M^{0.95}, \dots, \hat{\mathcal{N}}_J \pm M^{0.95}) \right] \\ &= \Pr \left[ \max_{1 \leq j \leq J} |\hat{\mathcal{N}}_j^r - \hat{\mathcal{N}}_j| \leq M^{0.95} \right] \\ &= \Pr [M^r \leq M^{0.95}] \\ &= 0.95. \end{aligned}$$

A sample from  $\hat{f}(y|x)$  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 recommend that the bandwidth or truncation point of  $\hat{f}^r(y|x)$  be that of  $\hat{f}(y|x)$ .

At this level of generality, the bootstrap as described above, or one of its variants (Efron and LePage (1992)), is perhaps all that is available. However, when the estimator of the conditional density is based on a parametric model or a series expansion, there are some alternative approaches based on the usual finite dimensional asymptotics.

Consider the case of a series expansion. An estimator of the one-step ahead conditional density based on a series expansion has the representation  $p_K(y|x, \hat{\theta})$

where  $\theta$  is a finite dimensional vector whose elements are the coefficients of the series expansion. The values  $\hat{\theta}$  of these coefficients are determined by some statistical procedure such as maximum likelihood or method of moments. The length  $p_\theta$  of  $\theta$  depends on the point  $K$  at which the underlying series expansion is truncated. If  $K$  is regarded as fixed and  $p_K(y|x, \theta^0)$  is regarded as the true data generating model for some  $\theta^0$ , then standard finite dimensional asymptotic results apply (Gallant (1987, Chapter 7)) and

$$\sqrt{n}(\hat{\theta} - \theta^0) \xrightarrow{\mathcal{L}} N[0, V(\theta^0)]$$

where  $V(\theta)$  is of the form

$$V(\theta) = [\mathcal{J}(\theta)]^{-1} [\mathcal{I}(\theta)] [\mathcal{J}(\theta)]^{-1}$$

with  $\mathcal{I}(\theta)$  and  $\mathcal{J}(\theta)$  determined by the statistical estimation procedure used to compute  $\hat{\theta}$ .

A conditional moment profile for fixed  $x$  is a sequence of functionals of the one-step ahead conditional density  $p_K(y|x, \theta)$ . As a consequence of the parametric representation of the one-step ahead conditional density, this sequence of functionals can be represented as a sequence of parametric functions:  $\{g_j(x, \theta): j = 1, \dots, J\}$ . Let

$$g(\theta) = [g_1(x, \theta), \dots, g_J(x, \theta)]'$$

An asymptotic 95% sup-norm confidence band for  $g(\theta^0)$  is  $g_j(x, \hat{\theta}) \pm \varepsilon$  where  $\varepsilon$  is determined such that

$$\int_{\{u \in \mathfrak{R}^{p_\theta}: g_i(x, \hat{\theta}) - \varepsilon < g_i(u) < g_i(x, \hat{\theta}) + \varepsilon, i = 1, \dots, J\}} N[du | \hat{\theta}, V(\hat{\theta})/n] = 0.95.$$

This integral is easiest to compute by drawing a random sample  $\{\theta^r\}_{r=1}^R$  from  $N[u | \hat{\theta}, V(\hat{\theta})/n]$ , using  $p_K(y|x, \theta^r)$  to compute  $g(\theta^r)$  for each  $r$  by Monte-Carlo integration as described in Subsection 2.5, and putting an  $\varepsilon$ -band around  $g(\hat{\theta})$  just wide enough to include 95% of the  $g(\theta^r)$ ,  $r = 1, \dots, R$ .

Alternatively, if the form of  $g(\theta)$  were not too complicated, one could find the distribution of  $g(u)$  from  $N[u | \hat{\theta}, V(\hat{\theta})/n]$  with a change of variables or use a Taylor expansion to get an approximation and compute  $\varepsilon$  analytically. With an analytical expression, one could also use the Bonferroni inequality to set simultaneously valid, equal width confidence intervals on the  $g_j(x, \theta)$  to get a sup-norm band although the use of such a conservative method would probably give unnecessarily wide bands. It is important to notice, however, that either analytical or Monte Carlo computations based on  $N[u | \hat{\theta}, V(\hat{\theta})/n]$  require inversion of the large and imprecisely determined matrix  $\mathcal{J}(\hat{\theta})$ .

Were it possible to get the exact, small sample distribution  $F_{\hat{\theta}}(u | \theta^0)$  of  $\hat{\theta}$ , one could sample from  $F_{\hat{\theta}}(u | \theta^0)$  instead of  $N[u | \hat{\theta}, V(\hat{\theta})/n]$  and avoid the inversion. Since  $\theta^0$  is not known, the obvious approach is to sample from  $F_{\hat{\theta}}(u | \hat{\theta})$  instead of  $F_{\hat{\theta}}(u | \theta^0)$ . This is the bootstrap as described at the beginning of this subsection. Its use would be justified if  $F_{\hat{\theta}}(u | \hat{\theta})$  could be shown to be asymptotically

$N[u|\theta^0, V(\theta^0)/n]$  since this would imply  $F_{\hat{\theta}}(u|\hat{\theta})$  is asymptotically  $F_{\hat{\theta}}(u|\theta^0)$  (because  $F_{\hat{\theta}}(u|\theta^0)$  is asymptotically  $N[u|\theta^0, V(\theta^0)/n]$ ). Unfortunately, the asymptotic theory that implies  $F_{\hat{\theta}}(u|\theta^0)$  is asymptotically  $N[u|\theta^0, V(\theta^0)/n]$  is not adequate to imply that  $F_{\hat{\theta}}(u|\hat{\theta})$  is asymptotically  $N[u|\theta^0, V(\theta^0)/n]$ .

One needs a central limit theorem of the sort that states  $F_{\hat{\theta}}(u|\theta_n)$  is asymptotically  $N[u|\theta^0, V(\theta^0)/n]$  for any sequence  $\theta_n$  that converges to  $\theta^0$ ; that is, one needs a continuously convergent central limit theorem. These are available for nonlinear models with regression structure (Gallant (1987, Chapter 3)) but are not available for nonlinear dynamic models (Gallant (1987, Chapter 7)). This point is discussed in Efron (1992), Bickel (1992), and their references.

Nonetheless, in view of the statistical and numerical instability in the computation of  $V(\hat{\theta})$  when  $p_{\theta}$  is large, we elect not to use methods based on  $N[u|\hat{\theta}, V(\hat{\theta})/n]$  despite the lack of results supporting use of  $F_{\hat{\theta}}(u|\hat{\theta})$ . See Efron and LePage (1992) for a more complete discussion of this point and further argument supporting this strategy.

When  $K$  grows with the sample size, either adaptively or deterministically, and  $\theta^0$  is regarded as the truncation of an infinite dimensional parameter, sup-norm bands constructed from  $N[u|\hat{\theta}, V(\hat{\theta})/n]$  are asymptotically valid and have a nonparametric interpretation in the situations considered by Andrews (1991), Eastwood (1991), and Gallant and Souza (1991). However, the processes and estimators they studied are simpler than those discussed here and their results have not been extended to our more complicated setting.

We use sup-norm bands, that is, equal width, simultaneously valid confidence bands, because they are easy to interpret and to display graphically. One could set an elliptical band on  $\{g_j(x)\}_{j=1}^J$  or bands whose width is proportional to  $\text{Var}[g_j(x)]$ . Proportional bands could be made simultaneously valid or they could be constructed to be valid only for each  $j$  considered separately as advocated by Runkle (1987) for use with linear VAR impulse-response analysis. The bootstrap strategy outlined above is easily adapted to any shape and can be made either simultaneously valid or separately valid as desired.

### 2.7. Profile Bundles

Bollerslev and Engle (1993) develop interesting theoretical notions of co-integration and persistence in variance. Following Bollerslev and Engle (1993), we will say that  $\{y_t\}$  is not integrated in mean if

$$\lim_{j \rightarrow \infty} \mathcal{E} \left[ \mathcal{E}(y_{t+j} | \mathcal{F}_{t+j-1}) | \mathcal{F}_t \right] = \text{const} \quad \text{for all } t$$

with probability one, where  $\mathcal{F}_t$  is the sigma-field generated by  $\{y_{t-j}\}_{j=0}^{\infty}$ . The process is integrated in mean if the condition fails. Similarly, the process is not integrated in variance if

$$\lim_{j \rightarrow \infty} \mathcal{E} \left[ \text{Var}(y_{t+j} | \mathcal{F}_{t+j-1}) | \mathcal{F}_t \right] = \text{const} \quad \text{for all } t$$

with probability one. The process is integrated in variance if the condition fails.

The idea is that the process is integrated in either mean or variance if the corresponding long-term forecasts of the one-step mean or variance remain sensitive to the initial condition in the limit as  $j \rightarrow \infty$ .

For example, consider  $y_{t+1} = \rho y_t + u_{t+1}$  where  $y_t$  is real valued and  $u_t$  is iid  $(0, \sigma^2)$ .  $\mathcal{E}[\mathcal{E}(y_{t+j} | \mathcal{F}_{t+j-1}) | \mathcal{F}_t] = \rho^j y_t$ . For  $0 < \rho < 1$  a plot against  $t + j$  will damp toward zero; for  $\rho = 1$  a plot will remain at  $y_t$  indefinitely.

These notions suggest a strategy for checking for integration in mean or variance. Namely, look for excessive sensitivity to initial conditions in both the conditional mean profile  $\{\hat{y}_j(x)\}_{j=0}^\infty$  and the conditional volatility profile  $\{\hat{v}_j(x)\}_{j=1}^\infty$ . Given that the notion of a unit root is intrinsically a linear concept, this seems to be the only practical strategy for investigating issues of integration in a model fitted to a general nonlinear process. In general, extreme sensitivity to initial conditions does not imply that the time series is nonstationary. For example, Nelson (1991) shows that I-GARCH models, which are integrated in variance, are strictly stationary but lack unconditional second moments.

A reasonable empirical strategy for checking the sequences  $\{\hat{y}_j(x)\}_{j=0}^\infty$  and  $\{\hat{v}_j(x)\}_{j=1}^\infty$  for excessive sensitivity to initial conditions is to over-plot profiles for sequences over a wide range of  $x$  values and see whether the thickness of this bundle of overplotted profiles tends to collapse to zero or retain its width indefinitely. In a sample large enough to permit nonparametric estimation of  $f(y|x)$ , the  $\{x_t\}$  sequence

$$x_t = (y'_{t-L+1}, \dots, y'_t)' \quad (t = L, \dots, n)$$

obtained directly from the data should provide an adequate range of values.

While these methods might be adapted to the determination of the average slope or Lyapunov exponent of a process for the purpose of checking for near epoch dependence or chaotic dynamics, there are other, more specialized, methods that are better suited to this task; see Nychka, Ellner, Gallant, and McCaffrey (1992) and their references.

### 3. EXAMPLES BASED ON PARAMETRIC MODELS

This section uses familiar ARCH and GARCH models to examine the properties of the concepts introduced in Section 2 and discusses implications for nonlinear impulse-response analysis.

#### 3.1. *AR(1) with ARCH(1) Errors*

Following Engle (1982), suppose

$$y_{t+1} = \lambda y_t + u_{t+1}$$

where  $\{u_{t+1}\}$  are  $N(0, \sigma_{t+1}^2)$  variables such that

$$\sigma_{t+1}^2 = a + \alpha u_t^2$$

and where  $-1 < \lambda \leq 1$ ,  $a > 0$ , and  $0 \leq \alpha \leq 1$ . The mapping to the notation of Section 2 is:  $x_t = (y_{t-1}, y_t)'$ ;  $x = (y_{-1}, y_0)'$ ; and the perturbed and baseline

initial conditions are  $x^+ = (y_{-1}, y_0 + \delta y^+)$ ,  $x^- = (y_{-1}, y_0 + \delta y^-)$ , and  $x^0 = (y_{-1}, y_0)$ , respectively, with  $\delta y^- = -\delta y^+$ .

For this model, a reasonable notion of the typical dynamic response of the one-step mean to a unit shock  $\delta y^+ = 1$  is the sequence  $\{\lambda^j\}_{j=0}^\infty$ . Similarly, a reasonable notion of the typical dynamic response of volatility to a unit shock is the sequence  $\{\alpha^j\}_{j=1}^\infty$ . As we shall see, for the AR(1) model with ARCH(1) errors, the impulse-response sequence for a unit shock is  $\{\lambda^j\}_{j=0}^\infty$  regardless of initial conditions. If the baseline initial condition is appropriately chosen, then the volatility impulse-response sequence to a unit shock is  $\{\alpha^j\}_{j=1}^\infty$  as well.

The AR(1) part of the model determines the conditional mean profile which is  $\hat{y}_j(x) = \lambda^j y_0$ . The corresponding impulse-response sequences are  $\hat{y}_j^+ - \hat{y}_j^0 = \lambda^j \delta y^+$ , and  $\hat{y}_j^- - \hat{y}_j^0 = \lambda^j \delta y^-$ . The response is symmetric; that is,  $\hat{y}_j^+ - \hat{y}_j^0 = -(\hat{y}_j^- - \hat{y}_j^0)$ .

Both the AR(1) part and the ARCH(1) part determine the conditional volatility profile which is

$$\hat{v}_j(x) = \begin{cases} a \frac{1 - \alpha^j}{1 - \alpha} + \alpha^j (y_0 - \lambda y_{-1})^2, & 0 < \alpha < 1, \\ aj + (y_0 - \lambda y_{-1})^2, & \alpha = 1. \end{cases}$$

The corresponding impulse-response sequences are

$$\begin{aligned} \hat{v}_j^+ - \hat{v}_j^0 &= \alpha^j (y_0 + \delta y^+ - \lambda y_{-1})^2 - \alpha^j (y_0 - \lambda y_{-1})^2, \\ \hat{v}_j^- - \hat{v}_j^0 &= \alpha^j (y_0 + \delta y^- - \lambda y_{-1})^2 - \alpha^j (y_0 - \lambda y_{-1})^2. \end{aligned}$$

The volatility impulse-response sequences depend upon the initial  $x^0 = (y_{-1}, y_0)$ , and are not symmetric, which reflects the nonlinear character of the process. However, if the initial condition is the unconditional mean  $x^0 = [\mathcal{E}(y), \mathcal{E}(y)] = (0, 0)$  and  $\delta y^+ = -\delta y^- = 1$ , then  $\hat{v}_j^+ - \hat{v}_j^0 = \hat{v}_j^- - \hat{v}_j^0 = \alpha^j$ , the anticipated pattern for the typical response of volatility to shocks in the ARCH(1) model.

Basic calculations show that if baseline  $x^0$  were drawn randomly from the unconditional distribution of  $x_t$ , then  $\mathcal{E}_x(\hat{v}_j^+ - \hat{v}_j^0) = \alpha^j$ , where the expectation is with respect to the unconditional distribution of  $x_t$ ,  $f(x_t)$ . To put this another way, if the sequence  $\{\hat{v}_j^+ - \hat{v}_j^0\}_{j=1}^\infty$  were computed for a large number of  $x^0$  drawn randomly from  $f(x_t)$ , then their average would also plot, approximately, as  $\{\alpha^j\}_{j=1}^\infty$ . In this sense, the average of the volatility impulse responses equals what one anticipates for the typical response of volatility to shocks in the ARCH(1) model.

Interestingly, then, for the ARCH(1) model, it makes no difference whether one first sets  $x^0 = \mathcal{E}(x)$  and computes the impulse-response sequence  $\{\hat{v}_j^+ - \hat{v}_j^0\}_{j=1}^\infty$ , or whether one draws  $x^0$  randomly from  $f(x_t)$  and then averages the impulse-response sequences. Either way delivers  $\{\alpha^j\}_{j=1}^\infty$  as the response to a unit shock. This equivalence is due to the linear-quadratic structure of the ARCH(1) model. It extends directly to general ARCH and GARCH models,

but not to other models of conditional heteroskedasticity with different functional forms, for example the E-GARCH model of Nelson (1991).

The characteristics of the profile bundles used to check for integration in mean or variance are determined as follows:

$$\lim_{j \rightarrow \infty} \{ \hat{y}_j(x_{t_1}) - \hat{y}_j(x_{t_2}) \} = (y_{t_1} - y_{t_2}) \lim_{j \rightarrow \infty} \lambda^j,$$

$$\lim_{j \rightarrow \infty} \{ \hat{v}_j(x_{t_1}) - \hat{v}_j(x_{t_2}) \} = \left[ (y_{t_1} - \lambda y_{t_1-1})^2 - (y_{t_2} - \lambda y_{t_2-1})^2 \right] \lim_{j \rightarrow \infty} \alpha^j.$$

The profile bundles for the mean will thus reveal long-term dependence on the initial condition when  $\lambda = 1$  and will not otherwise. Likewise, the profile bundles for the volatility will show long-term dependence if  $\alpha = 1$ , which is integration in variance, and will not otherwise.

By putting  $\alpha = 0$  in the variance equation  $\sigma_{t+1}^2 = a + \alpha u_t^2$  so that  $\sigma_{t+1}^2 = \sigma^2 = \text{const}$ , the contrast between a conditional volatility profile and a  $j$ -step ahead conditional mean square error path becomes readily apparent. With this restriction, the conditional volatility profile becomes  $\hat{v}_j(x) = \sigma^2$ , for all  $j$ , and the corresponding impulse-response sequences are such that  $\hat{v}_j^+ - \hat{v}_j^0 = \hat{v}_j^- - \hat{v}_j^0 = 0$ , for all  $j$ . The conditional volatility profile and impulse-response sequences do not depend upon  $j$  and will plot as horizontal lines. By way of contrast, the  $j$ -step ahead conditional mean square error path is  $\hat{\mathcal{M}}_j = \mathcal{E}\{[y_{t+j} - \mathcal{E}(y_{t+j}|x_t)]^2|x_t\} = \sigma^2(1 - \lambda^{2j})/(1 - \lambda^2)$  if  $\lambda^2 < 1$ , and  $\hat{\mathcal{M}}_j = \sigma^2 j$  if  $\lambda^2 = 1$ . The dependence of  $\hat{\mathcal{M}}_j$  on  $\lambda$  reflects the confounding of mean and variance parameters characteristic of the  $j$ -step ahead mean square error path.

### 3.2. AR(1) with GARCH(1, 1) Errors

Consider the generalization to a GARCH specification of Bollerslev (1986):

$$y_{t+1} = \lambda y_t + u_{t+1}$$

where  $\{u_{t+1}\}$  are  $N(0, \sigma_{t+1}^2)$  variables such that

$$\sigma_{t+1}^2 = a + \beta \sigma_t^2 + \alpha u_t^2$$

with  $-1 < \lambda \leq 1$ ,  $a > 0$ ,  $\alpha \geq 0$ ,  $\beta \geq 0$ , and  $\alpha + \beta \leq 1$ . When  $\alpha + \beta = 1$ , then this is the I-GARCH specification with the associated complications discussed by Nelson (1990).

For this setup, the conditional volatility profile takes the form

$$\begin{aligned} \hat{v}_j(x) = & a \sum_{k=0}^{j-1} (\alpha + \beta)^k + (\beta + \alpha)^{j-1} \alpha (y_0 - \lambda y_{-1})^2 \\ & + (\beta + \alpha)^{j-1} \beta \sum_{k=0}^{\infty} \beta^k (y_{-1-k} - \lambda y_{-2-k})^2. \end{aligned}$$

Here  $x = (\dots y_{-2}, y_{-1}, y_0)$ , with the dependence extending into the indefinite past due to the non-Markovian character of the GARCH model.



Impulse-response sequences for the volatility in this case are

$$\begin{aligned}\hat{v}_j^+ - \hat{v}_j^0 &= (\alpha + \beta)^{j-1} \alpha (y_0 + \delta y^+ - \lambda y_{-1})^2 \\ &\quad - (\alpha + \beta)^{j-1} \alpha (y_0 - \lambda y_{-1})^2, \\ \hat{v}_j^- - \hat{v}_j^0 &= (\alpha + \beta)^{j-1} \alpha (y_0 + \delta y^- - \lambda y_{-1})^2 \\ &\quad - (\alpha + \beta)^{j-1} \alpha (y_0 - \lambda y_{-1})^2.\end{aligned}$$

If  $\alpha + \beta < 1$ , then the volatility response sequences decay like  $(\alpha + \beta)^{j-1}$  as  $j \rightarrow \infty$ . If  $\alpha + \beta = 1$ , the I-GARCH case, then the impulse-response sequences are straight lines, reflecting the strong dependence on initial conditions.

#### 4. REPRESENTATIVE IMPULSE-RESPONSE SEQUENCES

As seen from the preceding sections, the impulse-response sequences of a nonlinear model depend upon the conditioning argument  $x^0$  used as the initial condition. It is clearly impractical to report the impulse-response sequences for many different  $x^0$ , and it is desirable to report the “representative” or average impulse-response sequences. There are two basic strategies for accomplishing this averaging. One is to use  $\mathcal{E}(x_t)$  for  $x^0$  and start the impulse responses from that point. The other is to draw  $x^0$  from the marginal density of  $x_t$ , compute the impulse-response sequence for each  $x^0$ , and then average the sequences over the drawings. Both strategies are implemented and compared in Section 5. For an ARCH model these two strategies are equivalent, as noted in Subsection 3.1. In general, though, the strategies could possibly give different pictures of the typical impulse-response sequence. The strategies differ in the orders in which the operations of function evaluation and integration take place. Neither strategy dominates the other, and there are advantages and disadvantages to each.

An advantage of the first strategy, which simply sets  $x^0 = \mathcal{E}(x)$ , is that it ensures that the conditioning vector lies near the center of the data where the conditional density is most precisely estimated. However, the fact that all elements of the conditioning are set to the same value means that the initial condition is rather unrealistic for studying volatility; a vector with all elements constant represents an abnormally calm period in terms of market volatility. We discuss this issue more in Section 5 below. The second strategy averages over all possible initial conditions and thereby circumvents this difficulty. But it is more computationally intensive, and, in particular, sup-norm confidence bands are beyond the reach of current equipment because the computational demands of averaging over many possible  $x$  vectors are great. In addition, the second computation could be less robust because it will be influenced by extreme  $x$ 's.

#### 5. IMPULSE-RESPONSE ANALYSIS OF STOCK PRICES AND VOLUME

There are two important strands to the empirical literature on short-term price and volume movements. The first strand, the ARCH literature, focuses

almost exclusively on the dynamics of the price process alone and investigates time-varying volatility. (See Bollerslev, Chou, and Kroner (1992) for a review of ARCH in finance.) In particular, French, Schwert, and Stambaugh (1987), Schwert (1989, 1990), and Nelson (1991) examine asymmetries in the conditional variance function. This work, which uses only the price series and focuses on one-step ahead prediction, reports evidence that large negative price movements are followed by larger increases in volatility than large positive movements. This asymmetric response of volatility to crashes and booms is termed the "leverage effect" after early work by Black (1976) and Christie (1982), who ascribe the effect to changes in the riskiness of firm equity induced by changes in the debt/equity ratio. In our work, we use the term "leverage" simply to describe asymmetry in the conditional variance function as is now common in the volatility literature (see, for example, Nelson (1991)).

The second strand documents a strong contemporaneous correlation between the trading volume and price volatility movements. The finding that large movements in price typically take place on days with high volume is extremely robust across a wide variety of markets and time periods. (See Karpoff (1987), Lamoureux and Lastrapes (1992), and Tauchen and Pitts (1983).) Following Kyle (1985), a theoretical literature with nontrivial implications for the price-volume relationship has emerged. These models generate trade through the mediation of new information in a framework characterized by imperfect competition coupled with asymmetric information. In Blume, Easley, and O'Hara (1991), the trading activities of informed traders convey information to uninformed traders through observed trading volume. This induces a contemporaneous relationship between price change and volume very similar in appearance to the results in the empirical literature. Trade has also been generated in noisy rational expectations models (De Long, Shleifer, Summers, and Waldmann (1990)). Lang, Litzenberger, and Madrigal (1992) examine cross-sectional relationships between volume and price movement in an effort to discriminate among variants of these models. While the recent theoretical literature is capable of generating trade and gives implications for contemporaneous price-volume relationships, virtually nothing is understood theoretically about the dynamics of price and volume movements.

In this section, we apply the methods suggested in Section 2 to generate empirical evidence on the multi-step ahead price and volume dynamics. Our focus is on three issues: (1) the persistence of the response of volatility to price shocks, (2) the extent to which the asymmetric conditional variance function is a transient or persistent phenomenon, and (3) the extent to which the contemporaneous price-volume relationship extends to the multi-step ahead case. The analysis utilizes the nonparametric estimate of the conditional density  $f(y|x)$  for stock price changes and volume obtained by Gallant, Rossi, and Tauchen (1992) with the SNP method.

The data set consists of  $n = 16,127$  observations, 1928–1987, on the daily logarithmic price change,  $\Delta p_t = 100[\log(p_t) - \log(p_{t-1})]$ , where  $p_t$  is the Standard and Poor's Composite Price Index, and on the logarithm of the daily

trading volume on the NYSE,  $v_t$ . Both the  $\Delta p_t$  and  $v_t$  series are adjusted to remove systematic calendar and trend effects, including a quadratic trend for volume, and are taken as jointly stationary. For reporting results, the  $\Delta p_t$  series is in units of percentage changes, that is,  $\Delta p = 1.0$  is a one-percent movement in the index. The  $v_t$  series is in units of standard deviations relative to the mean, so  $v_t = 1.0$  means one standard deviation above the quadratic trend. The volume series is exceptionally volatile. One standard deviation represents 42.7 percent above the trend, which would be on the order of 65 to 85 million shares above recent normal levels of 150 to 200 million shares per day. (The volume exceeded 600 million shares on the day of the '87 crash.)

To verify stationarity of the volume series, we fitted augmented Dickey-Fuller models by regressing adjusted log volume on one lag of itself and its lagged first differences. With lags of one, five, and twenty first differences, the coefficient on the lagged level is 0.86 (0.0044), 0.92 (0.0046), 0.95 (0.0050), where conventional OLS standard errors are in parentheses. These standard errors do not reflect the adjustments to remove trend and calendar effects from both the level and variance of the volume series. However, in view of the large sample size and the distance of the point estimates from unity, it seems reasonable to treat adjusted volume as stationary.

The SNP method, summarized in Gallant and Tauchen (1992a), uses a Hermite polynomial expansion to directly approximate the conditional density. The leading term of the expansion is an ARCH. The higher-order terms in the expansion have coefficients which are functions of the conditioning data. In this manner, the polynomial expansion allows for shape deviations from normality and conditional heterogeneity of unknown form. Gallant and Tauchen (1992a) derive an efficient rejection method for sampling the fitted conditional density.

The SNP technique is applied to the univariate series price change series alone,  $y_t = \Delta p_t$ , and to the bivariate series,  $y_t = (\Delta p_t, v_t)'$ . The estimation produces two fitted conditional densities,  $\hat{f}(y_{t+1}|x_t)$ , one for the univariate series and the other for the bivariate series. In either case, specification tests and other analysis indicate that a lag length of  $L = 16$  is required to fit the data. Hence,  $x_t$  is of length 16 in the univariate fit and of length 32 in the bivariate fit.

### 5.1. Univariate Stock Price Dynamics

Figure 1 shows the dynamic impulse responses of future  $\Delta p$  to shocks in contemporaneous stock price; Figure 2 shows the corresponding impulse responses for volatility. Specifically, Figure 1 shows the three profiles,  $\{\hat{y}_j^+, \hat{y}_j^0, \hat{y}_j^-\}_{j=0}^{20}$ , obtained by evaluating the sequence

$$\hat{y}_j(x) = \mathcal{E} \left[ \mathcal{E}(y_{t+j}|x_{t+j-1}) | x_t = x \right] \quad (j = 0, 1, \dots, 20)$$

at initial conditions  $x = x^+, x^0, x^-$ , where the expectations are computed from the SNP univariate estimate  $\hat{f}(y|x)$ .  $\mathcal{E}(y_{t+j}|x_{t+j-1})$  is evaluated analytically,

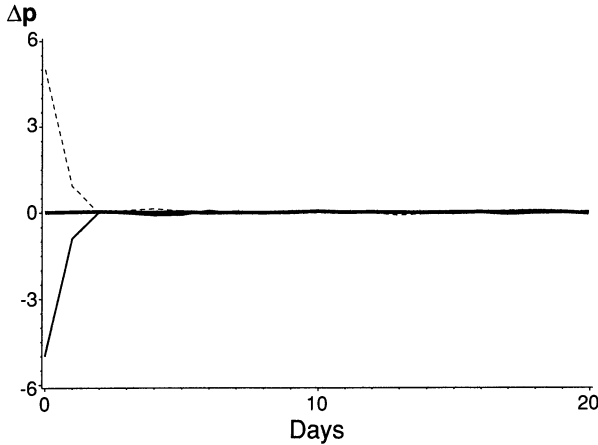


FIGURE 1.—Impulse response of  $\Delta p$  to  $\Delta p$  shock, univariate fit. Plotted is  $\hat{y}_j(x) = \mathcal{E}[\mathcal{E}(y_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 0, 1, \dots, 20$ . The heavy solid line is the baseline where  $x$  is put equal to  $x^0 = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})'$ . The solid line corresponds to a negative  $\Delta p$  shock where  $x$  is put equal to  $x^- = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, -5.0)'$ . The dashed line corresponds to a positive  $\Delta p$  shock where  $x$  is put equal to  $x^+ = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, 5.0)'$ .  $\Delta p$  and  $\hat{y}_j$  are measured in percent;  $j$  is measured in days.

$\mathcal{E}[\cdot | x_t = x]$  is evaluated by Monte Carlo integration. The initial conditions are

$$x^+ = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, 5.0)',$$

$$x^0 = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})',$$

$$x^- = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, -5.0)',$$

where  $\mu_{\Delta p} = 0.0163$  is the sample mean of the price changes, which is essentially zero on the scale of the figures. The initial condition  $x^0$  is thus the baseline where  $\Delta p_{t-j}$  is pegged to the mean for  $j \leq 0$ . The initial condition  $x^+$  corresponds to a five percent rise in the index from  $t - 1$  to  $t$ , starting from  $\Delta p_{t-j}$  set to the mean for  $j < 0$ . Similarly,  $x^-$  corresponds to a price decrease of five percent starting from the mean.

The two interesting features of Figure 1 are the extent to which the impulse responses are symmetric about the baseline and heavily damped. These features suggest that the conditional mean of the  $\{\Delta p_t\}$  series exhibits essentially no interesting higher order structure or serial dependence beyond lag one. At lag one, there is very mild linear dependence (autocorrelation less than 0.10) consistent with the asynchronous trading effect commonly displayed by market price indexes (Lo and Mackinlay (1988)).

Figure 2 shows impulse responses of price volatility to these same three shocks. The figure shows the three profiles,  $\{\hat{\nu}_j^+, \hat{\nu}_j^0, \hat{\nu}_j^-\}_{j=1}^{20}$ , for volatility. These profiles are obtained by evaluating the sequence

$$\nu_j(x) = \mathcal{E}[\text{Var}(y_{t+j}|x_{t+j-1})|x_t = x] \quad (j = 1, 2, \dots, 20)$$

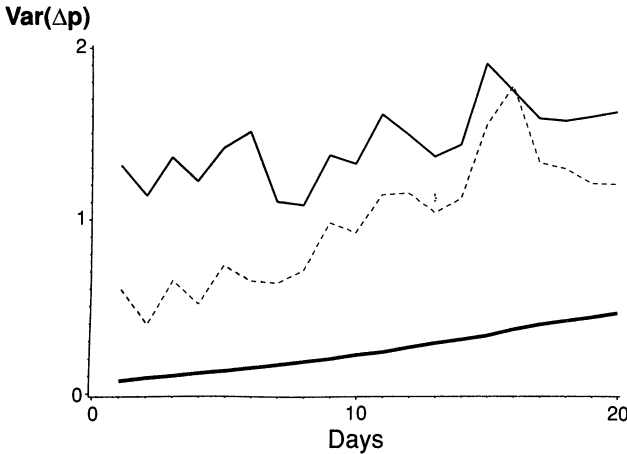


FIGURE 2.—Impulse response of volatility to  $\Delta p$  shock, univariate fit. Plotted is  $\nu_j(x) = \mathcal{E}[\text{Var}(y_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 1, \dots, 20$ . The heavy solid line is the baseline where  $x$  is put equal to  $x^0 = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})'$ . The solid line corresponds to a negative  $\Delta p$  shock where  $x$  is put equal to  $x^- = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, -5.0)'$ . The dashed line corresponds to a positive  $\Delta p$  shock where  $x$  is put equal to  $x^+ = (\mu_{\Delta p}, \mu_{\Delta p}, \dots, \mu_{\Delta p})' + (0, 0, \dots, 5.0)'$ .  $\Delta p$  is measured in percent,  $\nu_j$  is measured in percent squared, and  $j$  is measured in days.

at each of the three initial  $x$  values,  $x^+$ ,  $x^0$ , and  $x^-$ , defined above.  $\text{Var}(y_{t+j}|x_{t+j-1})$  is evaluated analytically,  $\mathcal{E}[\cdot|x_t = x]$  is evaluated by Monte Carlo integration. The impulse responses shown in Figure 2 indicate a leverage effect in which the price decrease has a larger effect on subsequent volatility than does the price increase. The wedge between the effects of positive and negative price shocks remains until about 15 days after the initial shocks. The responses also indicate that the effects of the price shocks on volatility are exceedingly slowly damped relative to the baseline, which is very close to I-GARCH behavior described by Bollerslev and Engle (1993). Note that the baseline shows a mild upward drift. The reason is that for data displaying ARCH-like behavior, volatility will be atypically low at an initial condition  $\Delta p_{t-j} = \mu_{\Delta p}$ ,  $j \geq 1$ . This situation is more quiescent than usual, so volatility will drift upwards from that point. Alternative methods, which average over the marginal distribution of  $x_t$ , can potentially avoid this drift as illustrated in Subsection 6.1 below.

## 5.2. Bivariate Price-Volume Dynamics

The next set of impulse-response simulations pertain to the effects that price and volume shocks have on subsequent volatility and volume. These results are obtained from the SNP fit to the bivariate series  $y_t = (\Delta p_t, v_t)'$ , where as before  $\Delta p_t$  is the adjusted log daily price change and  $v_t$  is the adjusted log daily volume. Recall that all results are reported with  $v_t$  expressed in units of unconditional standard deviations.

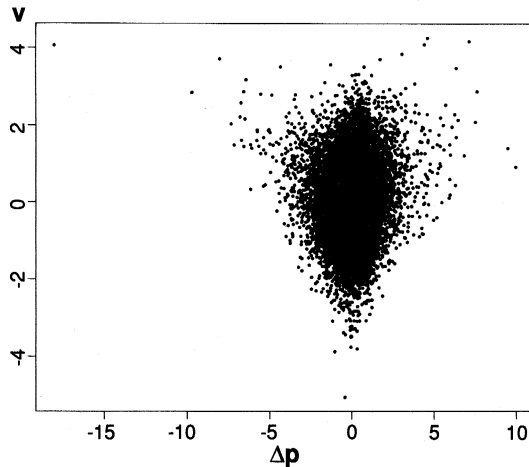


FIGURE 3.—Scatter plot of adjusted  $(\Delta p, v)$  data.  $v$  is measured in units of unconditional standard deviation, and  $\Delta p$  is measured in percent.

A complicating factor for the bivariate error shocks is the contemporaneous volume-volatility relationship. As is well known (Karpoff (1987); Tauchen and Pitts (1983)), days with large price movements in either direction are accompanied by higher trading volume. This association is analogous in some respects to the contemporaneous correlation that complicates linear VAR error shock analysis, as it should be accounted for in defining realistic shocks to the bivariate system. It differs from the usual correlation, though, in that it relates the variance of one of the variables,  $\Delta p_t$ , to the level of the other variable,  $v_t$ . There is essentially no relationship between  $\Delta p_t$  and  $v_t$ .

Figure 3 is a scatter plot of the data,  $(\Delta p_t, v_t)$ , which reveals clearly the contemporaneous volume-volatility relationship. The triangular shape of the point cloud shows that days with small price volatility tend to be days with lower than average volume, while days with large price volatility are high volume days.

The scatter plot is useful for defining shocks to prices and volume that are consistent with the historical range of the data. In particular, the scatter plot suggests the following design, with three types of error shocks labeled  $A$ ,  $B$ , and  $C$ , is typical of the variation of the data:

$$\delta y_A^+ = ( 5.0, 2.0)',$$

$$\delta y_A^- = (-5.0, 2.0)',$$

$$\delta y_B^+ = ( 5.0, 0.0)',$$

$$\delta y_B^- = (-5.0, 0.0)',$$

$$\delta y_C^+ = ( 0.0, 2.0)',$$

$$\delta y_C^- = ( 0.0, -2.0)'$$

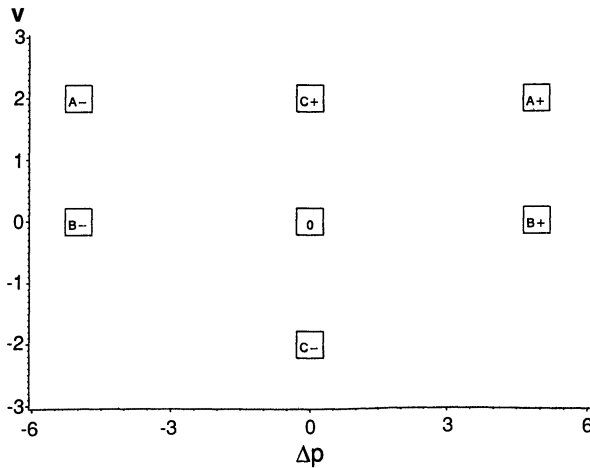


FIGURE 4.—Experimental design. Shown are *A* shocks,  $\delta y_A^+ = (5.0, 2.0)'$ ,  $\delta y_A^- = (-5.0, 2.0)'$ ; *B* shocks,  $\delta y_B^+ = (5.0, 0.0)'$ ,  $\delta y_B^- = (-5.0, 0.0)'$ ; and *C* shocks,  $\delta y_C^+ = (0.0, 2.0)'$ ,  $\delta y_C^- = (0.0, -2.0)'$ .  $v$  is measured in units of unconditional standard deviation, and  $\Delta p$  is measured in percent.

The *A* shocks are combined price-volume shocks where the price movements are  $\pm 5.0$  percent and volume is 2 standard deviations above its unconditional mean. The *B* shocks are pure price shocks of  $\pm 5.0$  percent with volume pinned at its mean. Finally, the *C* shocks are pure volume shocks of  $\pm 2.0$  standard deviations with no price movements. Figure 4 is a diagram of the three types of shocks. Comparing Figure 4 to Figure 3 indicates that each of the three classes of shocks do occur in the data set. The comparison also indicates that the layout of the design comes reasonably close to tracing out the extreme edges of the point cloud in Figure 3.

The recent theoretical literature on price-volume relationships can be used to interpret the design of the shock experiments. In the Blume-Easley-O'Hara framework, information is diffused and incorporated into prices via the trading of informed investors. The uninformed traders infer that a new packet of information has arrived in the system partially through the volume of trade. Thus, as information arrives the trading process diffuses the information, resulting in a price movement on higher than normal volume. Our *A* shock is designed to capture this situation. In some situations, a relevant piece of information may arrive as common knowledge. In this case (represented by a *B* shock), we expect to see an early consensus and a price movement on average contemporaneous volume. The positive *C* shock represents the situation where no consensus has been reached and trade occurs due to disparate beliefs. The negative *C* shock is included for symmetry.

The analysis reported here concentrates exclusively on the effects of shocks on the one-step variance of  $\Delta p_t$  (volatility) and the one-step mean of  $v_t$ . The effects of price and volume shocks on forecasts of  $\Delta p_{t+j}$ ,  $j \geq 1$ , are either very heavily damped, as in Figure 1 above, or negligible, and thus are not reported.

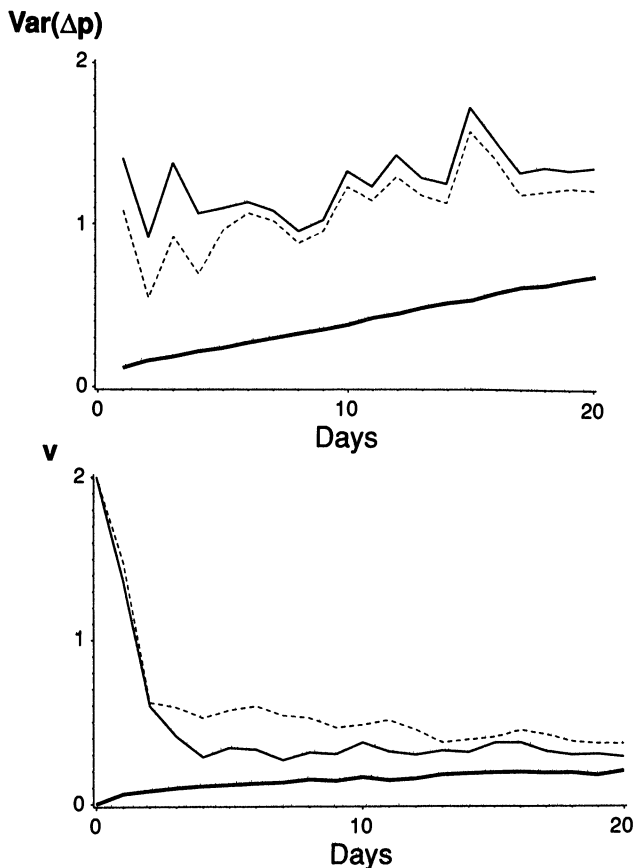


FIGURE 5.—Impulse responses of volatility and volume to  $\Delta p$  and  $v$  shock (type  $A$ ), bivariate fit. Plotted in the top panel is  $\hat{v}_{\Delta p,j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 1, \dots, 20$ . Plotted in the bottom panel is  $\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 0, 1, \dots, 20$ . In each panel, the heavy solid line is the baseline where  $x$  is put equal to  $x_A^0 = (\mu'_y, \mu'_y, \dots, \mu'_y)Y$ , the solid line corresponds to a negative  $A$  shock where  $x$  is put equal to  $x_A^- = (\mu'_y, \mu'_y, \dots, \mu'_y)Y + (0, 0, \dots, \delta y_A^-)Y$ , and the dashed line corresponds to a positive  $A$  shock where  $x$  is put equal to  $x_A^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)Y + (0, 0, \dots, \delta y_A^+)Y$ .  $v$  and  $\hat{v}_j$  are measured in units of unconditional standard deviation,  $\Delta p$  is measured in percent,  $\hat{v}_{\Delta p,j}$  is measured in percent squared, and  $j$  is measured in days.

Price and volume shocks do affect the one-step covariance of  $\Delta p_{t+j}$  with  $v_{t+j}$  and the one-step variance of  $v_{t+j}$ . But since these effects are less interesting from an economic perspective than the direct effects of shocks on price volatility and volume, they are not reported. Figures 5, 6, and 7 show the impulse responses of price volatility and volume to  $A$ ,  $B$ , and  $C$  shocks, respectively. The volatility responses, shown in the top panels, are computed as

$$\hat{v}_{\Delta p,j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j}|x_{t+j-1})|x_t = x] \quad (j = 1, 2, \dots, 20)$$

evaluated at  $x = x^+$ ,  $x^0$ , and  $x^-$ . The volume responses, shown in the bottom



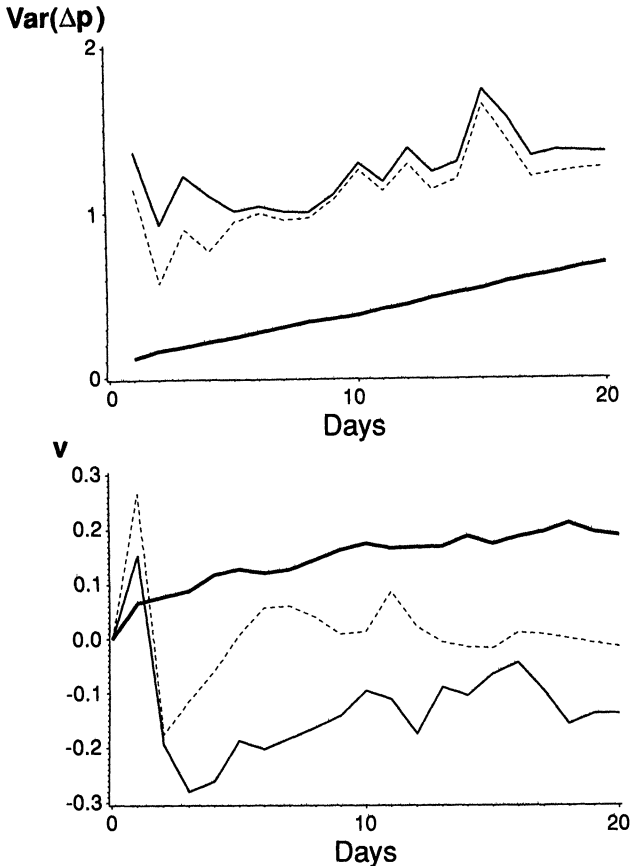


FIGURE 6.—Impulse responses of volatility and volume to pure  $\Delta p$  shock (type  $B$ ), bivariate fit. Plotted in the top panel is  $\hat{v}_{\Delta p,j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 1, \dots, 20$ . Plotted in the bottom panel is  $\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 0, 1, \dots, 20$ . In each panel, the heavy solid line is the baseline where  $x$  is put equal to  $x_B^0 = (\mu'_y, \mu'_y, \dots, \mu'_y)'$ , the solid line corresponds to a negative  $B$  shock where  $x$  is put equal to  $x_B^- = (\mu'_y, \mu'_y, \dots, \mu'_y)'$  +  $(0, 0, \dots, \delta y_B^-)'$ , and the dashed line corresponds to a positive  $B$  shock where  $x$  is put equal to  $x_B^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)'$  +  $(0, 0, \dots, \delta y_B^+)'$ .  $v$  and  $\hat{v}_j$  are measured in units of unconditional standard deviation,  $\Delta p$  is measured in percent,  $\hat{v}_{\Delta p,j}$  is measured in percent squared, and  $j$  is measured in days.

panels, are computed as

$$\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x] \quad (j = 0, 2, \dots, 20)$$

evaluated at the same three  $x$ 's. For  $A$  shocks, the initial conditions are

$$\begin{aligned} x_A^+ &= (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_A^+)', \\ x_A^0 &= (\mu'_y, \mu'_y, \dots, \mu'_y)', \\ x_A^- &= (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_A^-)'. \end{aligned}$$

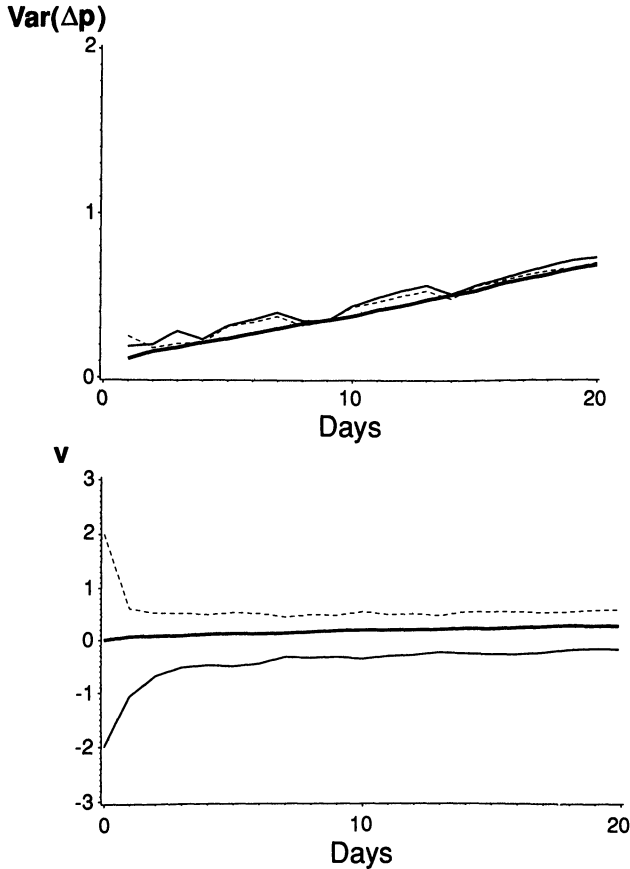


FIGURE 7.—Impulse responses of volatility and volume to pure  $v$  shock (type C), bivariate fit. Plotted in the top panel is  $\hat{v}_{\Delta p,j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 1, \dots, 20$ . Plotted in the bottom panel is  $\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x]$  against  $j$  for  $j = 0, 1, \dots, 20$ . In each panel, the heavy solid line is the baseline where  $x$  is put equal to  $x_C^0 = (\mu'_y, \mu'_y, \dots, \mu'_y)'$ , the solid line corresponds to a negative  $C$  shock where  $x$  is put equal to  $x_C^- = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_C^-)'$ , and the dashed line corresponds to a positive  $C$  shock where  $x$  is put equal to  $x_C^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_C^+)'$ .  $v$  and  $\hat{v}_j$  are measured in units of unconditional standard deviation,  $\Delta p$  is measured in percent,  $\hat{v}_{\Delta p,j}$  is measured in percent squared, and  $j$  is measured in days.

where  $\mu'_y$  is the sample average of the  $\{y_t\}$  process. For  $B$  and  $C$  shocks, the initial conditions are defined similarly with the appropriate  $\delta y$  in the rightmost place.

The three profiles shown in each of the panels of Figures 5, 6, and 7 are computed using Monte Carlo methods as described in Subsection 2.5. The computation of  $\mathcal{E}(v_{t+j}|x_{t+j-1})$  and  $\text{Var}(\Delta p_{t+j}|x_{t+j-1})$  are exact as a function of  $x_{t+j-1}$ , but their forecasts given  $x_t = x$  must be done by Monte Carlo.

The impulse-response analysis reveals four main characteristics of volatility and volume dynamics.

First, asymmetry of the volatility response (the leverage effect) is attenuated and is essentially a transient effect in the bivariate system. This can be seen by comparing the top panels of Figures 5 and 6 to Figure 2 and noting that in Figures 5 and 6 the leverage effect damps in about five or six days. The damping of leverage is much more rapid than the slow damping of volatility back towards baseline.

Second, the impulse responses of volume to volume shocks (*C* shocks) are symmetric and extremely slowly damped, as can be seen in the bottom panel of Figure 7.

Third, volume shocks (*C* shocks) have a very small effect on subsequent price volatility, as is evident in the top panel of Figure 7. This very mild feedback from volume to volatility is consistent with the findings of Schwert (1989) who applies linear methods to volume and a constructed volatility series. Weak feedback suggests that the price series is nearly Granger causally prior in the nonlinear sense discussed by Chamberlain (1982).

Fourth, large price movements to common-knowledge information events (*B* shocks) increase volume in the very short run but decrease volume over the longer term. This can be seen in the bottom panel of Figure 6. The short-run positive effect is evident in the work of Gallant, Rossi, and Tauchen (1992) who examine one-step ahead volatility. The very slowly damped long-term negative effect is a new finding of this paper.

The two most interesting and novel findings are the first, which pertains to long-term attenuation of leverage, and the fourth, which pertains to the contrast between the short- and long-term effects of price shocks on volume.

## 6. CONFIRMATION OF IMPULSE-RESPONSE RESULTS

### 6.1. Robustness to Conditioning Set

As previously noted in Section 3, there are two notions of a representative impulse-response sequence. The first, as just seen, is to use the unconditional mean of  $\{x_t\}$  as the starting point:  $x^0 = (\mu'_y, \dots, \mu'_y)'$ . The second is to average the impulse responses over the unconditional distribution of  $x_t$ . Figure 8 shows impulse responses computed for the *B*-shock computed the latter way. The top panel pertains to price volatility, where the negative, baseline, and positive response sequences are computed as

$$\begin{aligned} \bar{v}_{\Delta p, j}^- &= \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_{\Delta p, j}(x_{\tau-1} + \delta x^+), \\ \bar{v}_{\Delta p, j}^0 &= \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_{\Delta p, j}(x_{\tau-1}), \\ \bar{v}_{\Delta p, j}^+ &= \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_{\Delta p, j}(x_{\tau-1} + \delta x^-), \end{aligned}$$

for  $j = 1, 2, \dots, 20$ , where  $\delta x^+ = (0, 0, \dots, \delta y_B^+)'$  and  $\delta x^- = (0, 0, \dots, \delta y_B^-)'$ . For

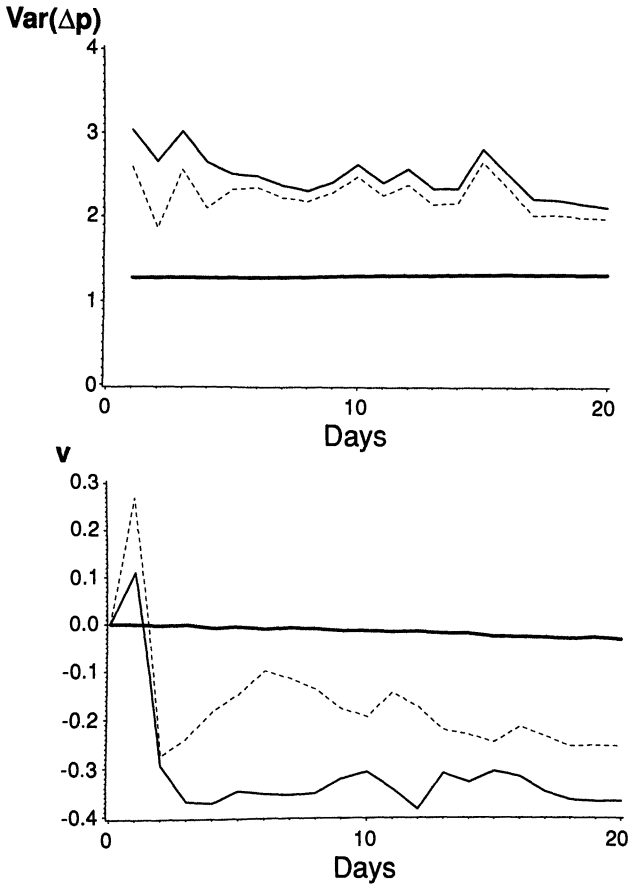


FIGURE 8.—Impulse responses of volatility and volume to pure  $\Delta p$  shock (type  $B$ ), averaged over the unconditional distribution of  $\{x_t\}$ , bivariate fit. Plotted in the top panel is  $\hat{v}_{\Delta p, j} = (1/16100)\sum_{\tau=28}^{16127} \mathcal{E}[\text{Var}(\Delta p_{t+j}|x_{t+j-1})|x_t = \hat{x}_{\tau-1}]$  against  $j$  for  $j = 1, \dots, 20$ . Plotted in the bottom panel is  $\hat{v}_j = (1/16100)\sum_{\tau=28}^{16127} \mathcal{E}[v_{t+j}|x_{t+j-1}|x_t = \hat{x}_{\tau-1}]$  against  $j$  for  $j = 0, 1, \dots, 20$ . In each panel, the heavy solid line is the baseline where  $\hat{x}_\tau = x_\tau$ , the solid line corresponds to a negative  $B$  shock where  $\hat{x}_\tau = x_\tau + (0, 0, \dots, \delta y_B^-)$ , and the dashed line corresponds to a positive  $B$  shock where  $\hat{x}_\tau = x_\tau + (0, 0, \dots, \delta y_B^+)$ .  $x_{\tau-1} = (\Delta p_{\tau-16}, v_{\tau-16}, \dots, \Delta p_{\tau-1}, v_{\tau-1})$ .  $v$  and  $\hat{v}_j$  are measured in units of unconditional standard deviation,  $\Delta p$  is measured in percent,  $\hat{v}_{\Delta p, j}$  is measured in percent squared, and  $j$  is measured in days.

the bottom panel, the averaged impulse responses for the volume are

$$\bar{v}_j^+ = \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_j(x_{\tau-1} + \delta x^-),$$

$$\bar{v}_j^+ = \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_j(x_{\tau-1}),$$

$$\bar{v}_j^+ = \frac{1}{16100} \sum_{\tau=28}^{16127} \hat{v}_j(x_{\tau-1} + \delta x^+),$$

for  $j = 0, 1, 2, \dots, 20$ . As expected, the baselines in Figure 8 are flat. The lack of baseline drift is due to averaging over many baselines with different patterns of drift. Comparing Figure 8 to Figure 6 indicates that, for this data set, the patterns for  $B$  shocks relative to the baselines are robust to the method used to compute the representative response. The same is true for  $A$  and  $C$  shocks (figures not shown).

## 6.2. Contrasts with Linear Models

A reasonable question is whether our findings could have been uncovered by fitting a more traditional linear model. To address this issue, we computed the impulse-response sequences for the three types of shocks using a highly restricted SNP model. The restricted model has a linear specification for the conditional mean function coupled with an ARCH-type specification for the conditional variance function, each with a lag length of 16. The conditional second moment specification of the restricted SNP model differs only mildly from the ARCH model of Engle (1982) in that it parameterizes the conditional standard deviation instead of the conditional variance. The error density is a modified Hermite-expansion with quartic polynomial to capture the thick-tailed character of the distribution. To be precise, the fitted model is the 1640000SNP specification reported in Table 5, p. 218, of Gallant, Rossi, and Tauchen (1992). The restricted SNP model is very close to the basic ARCH model with non-normal errors that has been utilized in many empirical finance studies (see Engle and Bollerslev (1986), and Engle and Gonzales-Rivera (1991)). Conventional model selection criteria and specification tests sharply reject this model in favor of more complicated SNP models, so it is interesting to investigate what features of the data it misses.

Figure 9 shows the impulse-response sequences for the  $B$ -shocks (common knowledge events) based on this specification of the conditional density. The sequences were computed using the second method which averages over all  $x$ , in the data set, so the figure is directly comparable to Figure 8.

In the top panel of Figure 9, the conditional variance responses to positive and negative shocks are constrained to be identical due to the symmetry of the basic ARCH specification. In contrast, the fully parameterized SNP fit used for Figure 8 displays damped asymmetry. Also, there is a noticeable baseline drift in the top panel of Figure 9 which is likely due to the misspecification intrinsic in this model.

The bottom panel of Figure 9 shows the volume responses to these positive and negative price shocks, which are constrained to be mirror images of one another due to the linear specification of the conditional mean function. This panel differs dramatically from its counterpart in Figure 8 obtained from the full SNP fit. In particular, the short-run increase in volume followed by a long-run decline for both positive and negative shocks is totally absent. The similar responses in volume irrespective of the sign of the price shock are to be expected, because of the well-known contemporaneous relationship between



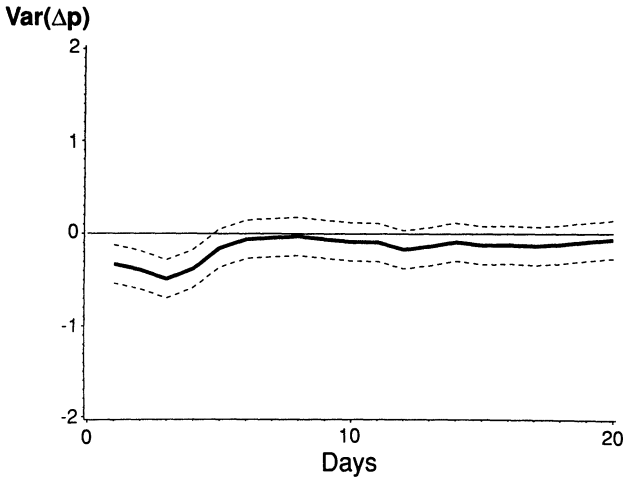


FIGURE 10.—95% confidence band for differential response of volatility to  $\Delta p$  and  $v$  shock (type A), bivariate fit. The solid line is  $\Delta \hat{v}_{\Delta p, j} = \hat{v}_{\Delta p, j}(x_A^+) - \hat{v}_{\Delta p, j}(x_A^-)$  plotted against  $j$  for  $j = 1, \dots, 20$  where  $\hat{v}_{\Delta p, j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j} | x_{t+j-1}) | x_t = x]$ ,  $x_A^- = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_A^-)'$ , and  $x_A^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_A^+)'$ . The dashed lines show a constant width, simultaneously valid 95% confidence interval on the 20 values of  $\Delta \hat{v}_{\Delta p, j}$ .  $\Delta p$  is measured in percent,  $\Delta \hat{v}_{\Delta p, j}$  is measured in percent squared, and  $j$  is measured in days.

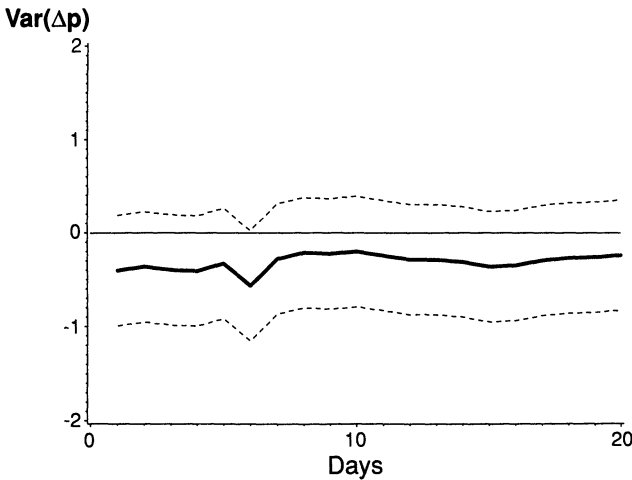


FIGURE 11.—95% confidence band for differential response of volatility to  $\Delta p$  shock, univariate fit. The solid line is  $\Delta \hat{v}_{\Delta p, j} = \hat{v}_{\Delta p, j}(x^+) - \hat{v}_{\Delta p, j}(x^-)$  plotted against  $j$  for  $j = 1, \dots, 20$  where  $\hat{v}_{\Delta p, j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j} | x_{t+j-1}) | x_t = x]$ ,  $x^- = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, -5.0)'$ , and  $x^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, 5.0)'$ . The dashed lines show a constant width, simultaneously valid 95% confidence interval on the 20 values of  $\Delta \hat{v}_{\Delta p, j}$ .  $\Delta p$  is measured in percent,  $\Delta \hat{v}_{\Delta p, j}$  is measured in percent squared, and  $j$  is measured in days.

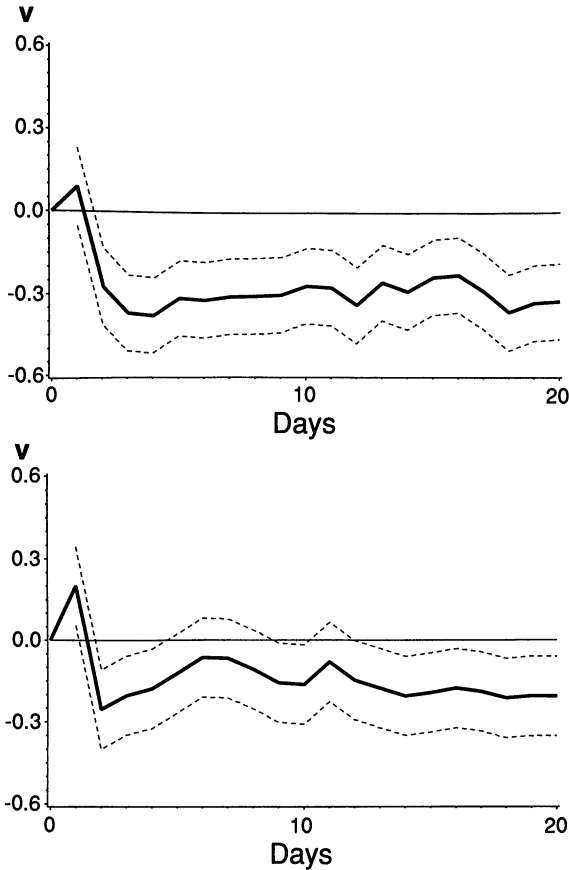


FIGURE 12.—95% confidence band for differential response of volume to pure  $\Delta p$  shock (type B), bivariate fit. The solid line in the top panel is  $\hat{v}_j(x_B^-)$  plotted against  $j$  for  $j = 0, 1, \dots, 20$  where  $x_B^- = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_B^-)'$ . The dashed lines show a constant width, simultaneously valid 95% confidence interval on the values of  $\hat{v}_j(x_B^-)$  for  $j = 1, \dots, 20$ . The solid line in the bottom panel is  $\hat{v}_j(x_B^+)$  plotted against  $j$  for  $j = 0, 1, \dots, 20$  where  $\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x]$  and  $x_B^+ = (\mu'_y, \mu'_y, \dots, \mu'_y)' + (0, 0, \dots, \delta y_B^+)'$ . The dashed lines show a constant width, simultaneously valid 95% confidence interval on the values of  $\hat{v}_j(x_B^+)$  for  $j = 1, \dots, 20$ .  $v$  and  $\hat{v}_j(x)$  are measured in units of unconditional standard deviation and  $j$  is measured in days.

bands around quantities relevant for these findings, and thereby provide an indication of the statistical significance. We compute the bands for the first type of impulse-response functions where the unconditional mean of  $x_t$  is the starting point. The computations for the bands are considerably less demanding for the first type and the robustness analysis suggested similar point estimates. The confidence bands are 95 percent sup-norm confidence bands obtained by bootstrapping as described in Subsection 2.6.

Figure 10 shows a 95 percent confidence band around the estimates

$$\hat{v}_{\Delta p, j}(x_A^+) - \hat{v}_{\Delta p, j}(x_A^-) \quad (j = 1, 2, \dots, 20).$$



If the population volatility function is symmetric, that is, the leverage effect is absent, then the above differences should be jointly insignificant. The figure suggests some evidence for asymmetry in the first five days followed by no detectable asymmetry; the sup-norm bands after day five are tightly centered about zero differential response.

Figure 11 shows a 95 percent confidence band for the differential response of volatility from the univariate estimation discussed in Subsection 5.1 above. The point estimate of the differential response shows persistent asymmetry, though the confidence band is wide and includes zero for all but one day. Overall, our interpretation from the two figures is that there is some evidence for short-run asymmetry but very little evidence for long-term persistent asymmetry.

The top and bottom panels of Figure 12 show similarly constructed 95 percent confidence bands around the effects of pure price shocks on volume relative to baseline. The top panel shows

$$\hat{v}_j(x_B^+) - \hat{v}_j(x_B^0) \pm M_+^{0.95} \quad (j = 1, 2, \dots, 20)$$

while the bottom panel shows

$$\hat{v}_j(x_B^-) - \hat{v}_j(x_B^0) \pm M_-^{0.95} \quad (j = 1, 2, \dots, 20)$$

where the  $M$ 's are computed as described in Subsection 2.6. The confidence bands indicate that the short-run increase in volume generated by the price shock is marginally statistically significant at the 95 percent level. On the other hand, the long-term decrease in volume is more strongly significant. The long-term decrease in volume presents a challenge to theoretical models of short-term speculative trading. Models based on either the notion of a random mixing process forcing a sequence of equilibria (Tauchen and Pitts (1983), Andersen (1992), Gallant and Tauchen (1992b)) or the notion of a market microstructure (Foster and Viswanathan (1992)) predict that volume remains above baseline.

### 7. PERSISTENCE IN VARIANCE

If a GARCH model is fitted to short-term financial price movements, then the implied volatility process looks integrated, or very nearly so (Engle and Bollerslev (1986)). Nelson (1990) presents theoretical evidence that one should expect to see a "unit-root" in variance as data are sampled more frequently, even if the true process is stationary. His findings are the second-moment analogues of Sims's (1984) findings regarding the martingale-like behavior of financial prices.

We utilize the methods of Subsection 2.7 to examine persistence properties of the price and volume process. The top panel of Figure 13 shows overplots of conditional mean profiles  $\{\Delta \hat{p}_j(x)\}_{j=1}^{100}$ , for  $x \in \{x_t: t = 28, 156, 284, \dots, 16028\}$ . Some reduction of the 16,100 available  $x$  values is needed to keep the plots from becoming overly dense. The choice of every 128th  $x_t$ , which generates 125 profiles, was determined by experimentation. The middle panel of Figure 13

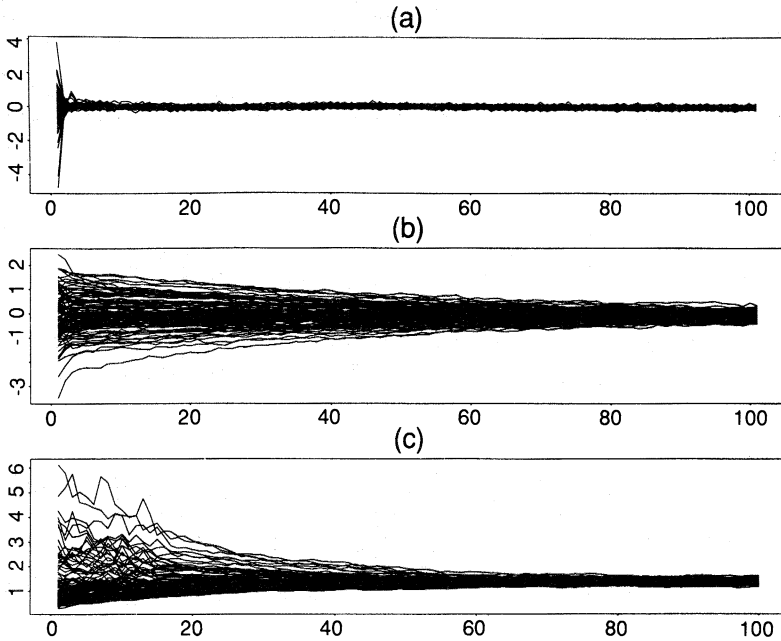


FIGURE 13.—Profiles, bivariate fit. Panel (a) shows overplots for  $\tau = 28, 156, \dots, 16028$  of the conditional mean profiles  $\{\Delta \hat{p}_j(x_t)\}_{j=1}^{100}$  where  $\Delta \hat{p}_j(x) = \mathcal{E}[\mathcal{E}(\Delta p_{t+j} | x_{t+j-1}) | x_t = x_{\tau-1}]$ . Panel (b) shows overplots for  $\tau = 28, 156, \dots, 16028$  of the conditional mean profiles  $\{\hat{v}_j(x)\}_{j=1}^{100}$  where  $\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j} | x_{t+j-1}) | x_t = x_{\tau-1}]$ . Panel (c) shows overplots for  $\tau = 28, 156, \dots, 16028$  of the conditional volatility profiles  $\{\hat{v}_{\Delta p, j}(x)\}_{j=1}^{100}$  where  $\hat{v}_{\Delta p, j}(x) = \mathcal{E}[\text{Var}(\Delta p_{t+j} | x_{t+j-1}) | x_t = x_{\tau-1}]$  against  $j$  for  $j = 1, \dots, 20$ .  $v$  and  $\hat{v}_j$  are measured in units of unconditional standard deviation,  $\Delta p$  is measured in percent,  $\hat{v}_{\Delta p, j}$  is measured in percent squared, and  $j$  is measured in days.

shows overplots of similarly constructed profiles for the volume. For each  $t$ , the profile is computed out  $J = 100$  steps.

The profile for  $\Delta p$  shown in the top panel shows that this series is clearly not integrated; there is a very strong damping of the profiles to the unconditional mean. The volume series, however, shows a damping of the profiles consistent with stationarity but at a much slower rate than for the  $\Delta p$  series. Examination of the plot shows a half rate of decay of the initial perturbation of some 40 periods. This is roughly consistent with the sort of damping expected in an AR(1) model with an autoregressive parameter of 0.90.

The bottom panel of Figure 13 shows the conditional volatility profiles  $\{\hat{v}_{\Delta p, j}(x)\}_{j=1}^{100}$  for  $x \in \{x_t : t = 28, 156, 284, \dots, 16028\}$  which is the same set of  $x$ 's used in the conditional mean profiles. The volatility profiles suggest that shocks to variance are very slowly damped, but do die out. The volatility shows little evidence of integration 100 steps out. The profile for the bivariate fits contrast sharply with very unstable and persistent profiles for the univariate fitted SNP. For example, the univariate volatility profile shown in Figure 1 shows little evidence of convergence to the baseline. Recall from Subsection 3.2, that the conditional volatility profiles for an AR(1) model with GARCH(1, 1) errors will

damp at a geometric rate governed by  $\alpha + \beta$ :

$$\begin{aligned} & \mathcal{E} \left[ \text{Var} (\Delta p_{t+j} | x_{t+j-1}) | x_t \right] \\ &= a \sum_{k=0}^{j-1} (\alpha + \beta)^k + (\beta + \alpha)^{j-1} \alpha (y_t - \lambda y_{t-1})^2 \\ & \quad + (\beta + \alpha)^{j-1} \beta \text{Var} (\Delta p_t | x_{t-1}). \end{aligned}$$

For an I-GARCH model,  $\alpha + \beta = 1$ , the profile for each  $x_t$  is

$$\begin{aligned} \mathcal{E} \left[ \text{Var} (\Delta p_{t+j} | x_{t+j-1}) | x_t \right] &= (j - 1)a + \alpha (y_t - \lambda y_{t-1})^2 \\ & \quad + \beta \text{Var} (\Delta p_t | x_{t-1}) \end{aligned}$$

which shows linear drift away from a value that depends strongly on  $x_t$ . The volatility profiles do not display this type of long-term dependence on initial condition characteristic of the I-GARCH model.

Some caution should be exercised in interpreting our findings on persistence. The Markovian structure of the SNP model allows for only a finite number of lags of the variables in the conditioning set. To capture extremely long-run persistence may require a very large number of lags. Diagnostics performed by Gallant, Rossi, and Tauchen (1992) suggest that the fitted bivariate SNP model fails to capture some of the movements of volatility at a yearly frequency; however, additional evidence shows that volatility computed from the fitted bivariate SNP closely conforms to movements in volatility on a monthly and shorter term basis.

## 8. CONCLUSION

We have developed an approach for analyzing the dynamics of a nonlinear time series that is represented by a nonparametric estimate of its one-step ahead conditional density. The approach entails examination of conditional moment profiles corresponding to certain shocks; a conditional moment profile is the conditional expectation evaluated at time  $t$  of a time invariant function evaluated at time  $t + j$  regarded as a function of  $j$ . The difference between a conditional moment profile and a baseline profile is the generalization of linear impulse-response analysis to the nonlinear situation. The approach includes strategies for laying out realistic perturbation experiments in multivariate situations and for undertaking statistical inference using bootstrap methods.

The empirical work uncovers two important new characteristics of the multi-period dynamics of the market price change and the transactions volume on the NYSE. First, the so-called leverage effect is a heavily damped transient phenomenon; it is both statistically and practically insignificant five to six days after a large movement in the market. Second, large price movements on common-knowledge information generate increased volume in the short run but reduce volume over the longer term.

The work underscores the importance of using nonlinear models to capture dynamics. In particular, the two major findings would have been missed com-

pletely by models linear in the conditional mean and/or variance functions. The work also demonstrates potential for using nonparametric density estimates to implement practicable inference procedures based on bootstrap methods.

*Dept. of Statistics, North Carolina State University, Raleigh, NC 27695, U.S.A.,  
Grad. School of Business, University of Chicago, 1101 E. 58th St., Chicago, IL  
60637, U.S.A.*

*and*

*Dept. of Economics, Duke University, Durham, NC 27708, U.S.A.*

*Manuscript received November 1990; final revision received November, 1992.*

#### REFERENCES

- ANDERSEN, TORBEN (1992): "Return Volatility and Trading Volume in Financial Markets: An Information Interpretation of Stochastic Volatility," unpublished Ph.D. dissertation, Yale University.
- ANDREWS, DONALD W. K. (1991): "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Models," *Econometrica*, 59, 307–346.
- BICKEL, PETER J. (1992): "Theoretical Comparison of Different Bootstrap  $t$  Confidence Bounds," in *Exploring the Limits of Bootstrap*, ed. by Raoul LePage and Lynne Billard. New York: Wiley.
- BLACK, F. (1976): "Studies of Stock Market Volatility Changes," *Proceedings of the American Statistical Association, Business and Economics Section*, 177–181.
- BLANCHARD, OLIVIER J., AND DANNY QUAH (1989): "The Dynamic Effects of Aggregate Supply and Demand Disturbances," *American Economic Review*, 79, 655–673.
- BLANCHARD, OLIVIER J., AND MARK W. WATSON (1986): "Are Business Cycles All Alike?" in *The American Business Cycle: Continuity and Change*, NBER Research Studies in Business Cycles, Vol. 25, ed. by R. J. Gordon. Chicago: University of Chicago Press.
- BLUME, LAWRENCE, DAVID EASLEY, AND MAUREEN O'HARA (1991): "Market Statistics and Technical Analysis: The Role of Volume," unpublished manuscript, Cornell University.
- BOLLERSLEV, TIM (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.
- BOLLERSLEV, TIM, RAY CHOU, AND KENNETH KRONER (1992): "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence, with Suggestions for Future Research," *Journal of Econometrics*, 52, 5–61.
- BOLLERSLEV, TIM, AND ROBERT ENGLE (1993): "Common Persistence in Conditional Variances," *Econometrica*, 61, 167–186.
- CHAMBERLAIN, GARY (1982): "The General Equivalence of Granger and Sims Causality," *Econometrica*, 50, 569–582.
- CHRISTIE, A. (1982): "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects," *Journal of Financial Economics*, 10, 407–432.
- DE LONG, J. BRADFORD, ANDREI SHLEIFER, LAWRENCE H. SUMMERS, AND ROBERT J. WALDMANN (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, 703–738.
- DOAN, THOMAS, ROBERT LITTELMAN, AND CHRISTOPHER SIMS (1984): "Forecasting and Conditional Projection using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.
- EASTWOOD, BRIAN J. (1991): "Asymptotic Normality and Consistency of Semi-nonparametric Regression Estimators Using an Upward  $F$  Test Truncation Rule," *Journal of Econometrics*, 48, 151–182.
- EFRON, BRADLEY (1992): "Six Questions Raised by the Bootstrap," in *Exploring the Limits of Bootstrap*, ed. by Raoul LePage and Lynne Billard. New York: Wiley.
- EFRON, BRADLEY, AND RAOUL LEPAGE (1992): "Introduction to Bootstrap," in *Exploring the Limits of Bootstrap*, ed. by Raoul LePage and Lynne Billard. New York: Wiley.
- ENGLE, ROBERT F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.

- ENGLE, ROBERT F., AND TIM BOLLERSLEV (1986): "Modeling the Persistence of Conditional Variance," *Econometric Reviews*, 5, 1–50.
- ENGLE, ROBERT F., AND GLORIA GONZALES-RIVERA (1991): "Semiparametric ARCH Models," *Journal of Business and Economic Statistics*, 9, 345–360.
- ENGLE, ROBERT F., TAKATOSHI ITO, AND WEN-LING LIN (1990): "Meteor Showers or Heat Waves?" Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market," *Econometrica*, 58, 525–542.
- FRENCH, K., G. W. SCHWERT, AND R. STAMBAUGH (1987): "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–30.
- FOSTER, DOUGLAS, AND S. VISWANATHAN (1992): "Can Speculative Trading Explain the Volume-Volatility Relation?" unpublished manuscript, Fuqua School of Business, Duke University.
- GALLANT, A. RONALD (1987): *Nonlinear Statistical Models*. New York: Wiley.
- GALLANT, A. RONALD, PETER E. ROSSI, AND GEORGE TAUCHEN (1992): "Stock Prices and Volume," *Review of Financial Studies*, 5, 199–242.
- GALLANT, A. RONALD, AND GERALDO SOUZA (1991): "On the Asymptotic Normality of Fourier Flexible Form Estimates," *Journal of Econometrics*, 50, 329–353.
- GALLANT, A. RONALD, AND GEORGE TAUCHEN (1992a): "A Nonparametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation," in *New Directions in Time Series Analysis*, ed. by D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt, and M. Taquq. New York: Springer-Verlag, pp. 71–92.
- (1992b): "Which Moments to Match?" unpublished manuscript, Duke University.
- KARPOFF, J. (1987): "The Relation between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis*, 22, 109–126.
- KING, ROBERT G., CHARLES I. PLOSSER, JAMES H. STOCK, AND MARK W. WATSON (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81, 819–840.
- KYLE, ALBERT S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1336.
- LAMOUREUX, C., AND W. D. LASTRAPES (1992): "Endogenous Trading Volume and Momentum in Stock Return Volatility," unpublished manuscript, Washington University in St. Louis.
- LANG, LARRY H. P., ROBERT H. LITZENBERGER, AND VINCENTE MADRIGAL (1992): "Testing Financial Market Equilibrium under Asymmetric Information," *Journal of Political Economy*, 100, 317–348.
- LIU, REGINA (1990): "On a Notion of Data Depth Based on Random Simplexes," *Annals of Statistics*, 18, 405–414.
- LO, ANDREW, AND A. MACKINLAY (1988): "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies*, 1, 41–66.
- NELSON, DANIEL (1990): "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," unpublished manuscript, University of Chicago.
- (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.
- NYCHKA, DOUGLAS, STEPHEN ELLNER, A. R. GALLANT, AND DANIEL MCCAFFREY (1992): "Finding Chaos in Noisy Systems," *Journal of the Royal Statistical Society B*, 54, 399–426.
- POTTER, SIMON M. (1991): "Nonlinear Impulse-Response Functions," unpublished manuscript, University of California, Los Angeles.
- ROBINSON, PETER M. (1983): "Nonparametric Estimators for Time Series," *Journal of Time Series Analysis*, 4, 185–207.
- RUNKLE, DAVID E. (1987): "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics*, 5, 437–442.
- SCHWERT, G. WILLIAM (1989): "Why Does Stock Volatility Change Over Time?" *Journal of Finance*, 44, 1115–1154.
- (1990): "Stock Volatility and the Crash of '87," *Review of Financial Studies*, 3, 77–102.
- SIMS, CHRISTOPHER (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.
- (1984): "Martingale-like Behavior of Interest Rates and Prices," Discussion Paper No. 205, University of Minnesota.
- TAUCHEN, GEORGE, AND MARK PITTS (1983): "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485–505.