Nonlinear Dynamic Structures [1]

A. Ronald Gallant
North Carolina State University

Peter E. Rossi
University of Chicago

George Tauchen
Duke University

September 1990
Revised October 1990

[1]Research supported by National Science Foundation Grants SES-8808015 and SES-8810357, North Carolina Agricultural Experiment Station Project NCO-6134, and the PAMS Foundation.

Abstract

We describe three methods for analyzing the dynamics of a nonlinear time series that is represented by a nonparametric estimate of its one-step ahead conditional density. These strategies are based on 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 first method, which compares conditional moment profiles to baseline profiles, is the nonlinear analog of conventional impulse-response analysis. The second assesses the significance of a profile by comparing its sup-norm confidence band to a null profile. The third examines profile bundles for evidence of damping or persistence. Experimental designs for choosing an appropriate set of shocks are discussed. These methods are applied to 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 findings from these data are: (i) The multi-step ahead conditional volatility profile exhibits a symmetric response to both positive and negative price shocks. In contrast, the conditional volatility profile of the univariate price change process exhibits an asymmetric response. (ii) The one-step ahead response of volume to price shocks is different than the multi-step ahead response. Price shocks produce an increase in volume one-step ahead but decrease it in subsequent steps. (iii) There is little evidence for long-term persistence in either the conditional mean or volatility of the bivariate process.

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. Economic issues related to predictability, volatility, and other properties of the time series are most naturally thought of as relating to characteristics of the conditional density. Statistical tests of an economic model may be regarded as checks for how well the model's description of selected characteristics of a process match those of 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: Using a recursive system of GARCH models, Engle, Ito, and Lin (1990), perform an error shock 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 is not possible to define an innovation and compute the impulse-response function in the manner advocated by Sims. However, the dynamic properties of the nonlinear model can be elicited by perturbing the vector of conditioning arguments in the conditional density function. The dynamic response to the perturbation can be traced out by computing multi-step ahead 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 which 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 somewhat 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 apply these ideas to the conditional density that was estimated using the SNP method by Gallant, Rossi, and Tauchen (1990) from 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. They examined the one-step ahead conditional density with the conditioning arguments set at their unconditional means. Here, we perturb the vector of conditioning variables away from the mean and trace out the effect of perturbations upon multi-step ahead conditional mean and volatility profiles. Also, we examine bundles of long-term conditional mean and volatility profiles for evidence of persistence. Of particular interest to researchers in finance is the effect of shocks on subsequent price volatility and volume.

The organization of the paper is as follows. Section 2 outlines a nonlinear impulse-response methodology and discusses computations. Section 3 applies these methods to an AR(1) model with ARCH(1) errors to provide a comparison with existing methods for linear difference equations. Section 4 presents the application to the price and volume data. Section 5 summarizes the conclusions of the study.

2 Impulse-Response Analysis of Nonlinear Models

Let $\{y_t\}_{t=-\infty}^{\infty}$ with $y_t \epsilon R^M$ 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

$$Var(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty}) = \mathcal{E}\Big\{[y_{t+1} - \mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty})][y_{t+1} - \mathcal{E}(y_{t+1}|\{y_{t-k}\}_{k=0}^{\infty})]'|\{y_{t-k}\}_{k=0}^{\infty}\Big\}$$

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

$$\mathcal{E}[g(y_{t+j-J}, \ldots, y_{t+j}) | \{y_{t-k}\}_{k=0}^{L-1}] \qquad j = 0, 1, 2, ..$$

the word moment refers to the time-invariant function $g(y_{-J}, \ldots, y_0)$. 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's (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 $Var(u_t) = \mathcal{E}u_t u_t' = \Omega$. 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 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 ξ_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 $\{\hat{y}_t^0\}_{t=-\infty}^{0}$ 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 $\{\hat{y}_t^+\}_{t=-\infty}^{0}$. By the law of iterated expectations, $\hat{y}_t^+$ is also the forecast of the one-step mean $\mathcal{E}(y_{t+1}|\{y_{t-j}\}_{j=0}^{\infty})$ given initial conditions $\{\hat{y}_t^+\}_{t=-\infty}^{0}$.

In applied work, the contemporaneous covariance matrix Ω 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, in the absence of *a priori* information about causality patterns, there is no unique way to perform this orthogonalization.

In the general nonlinear case, there is no direct way of perturbing an innovation and tracing through the effects of the perturbation. However, if instead of viewing perturbation of u_t as the primitive concept, the computational technique of perturbing y_0 is viewed as the primitive then the ideas extend directly, as we shall see in the subsections that follow.

As with VAR error shocks, issues arise regarding the task of obtaining realistic shocks in the case of multivariate data. These issues are addressed more fully in Subsections 4.2 and 3.2. The strategy, which is rather different than orthogonalizing innovations, 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

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

with $x = (y'_{-L+1}, \ldots, y'_0)'$. (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 $x = (y'_{-L+1}, y'_{-L+2}, \ldots, y'_0)' : y_0 \in R^M$ represents a contemporaneous value, the $y_{-k} \in R^M$, $1 \le k \le L-1$, represent lags. Let $\delta y^+, \delta y^- \epsilon R^M$ represent small perturbations to the contemporaneous y_0 where δy^+ is "positive" and δy^- is "negative" in a sense to be made precise later. Put

$$\begin{aligned} x^+ &= (y'_{-L+1}, y'_{-L+2}, \ldots, y'_0)' + (0, 0, \ldots, \delta y^{+\prime})' \\ x^0 &= (y'_{-L+1}, y'_{-L+2}, \ldots, y'_0)' \\ x^- &= (y'_{-L+1}, y'_{-L+2}, \ldots, y'_0)' + (0, 0, \ldots, \delta y^{-\prime})' \end{aligned}$$

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

Now put

$$\begin{aligned} \hat{y}_j^+ &= \hat{y}_j(x^+) \\ \hat{y}_j^0 &= \hat{y}_j(x^0) \\ \hat{y}_j^- &= \hat{y}_j(x^-) \end{aligned}$$

for $j = 0, 1, 2, \ldots$. 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 δy^+ (or δy^-). The net effect is obtained by comparing the profile for δy^+ (or δy^-) 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.

Tracing out the impulse responses of a nonlinear system in this way 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 δy^+ 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 δy^+ 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 4.2 below, is to inspect a scatter plot of the data cloud $\{y_t\}$ and visually determine shocks δy^+ and δy^- 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 explore 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, 1990a, Bollerslev and Engle, 1989, 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+1} which is computed as

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

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

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

for j=1,2,... where $x = (y'_{-L+1}, \ldots, y'_0)'$. (If a dummy variable of integration coincides with an element of x, that integration is omitted.) $\hat{\mathcal{V}}_j(x)$ 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 or shock δy^+, and x^- a negative impulse δy^-. The net effect of an impulse is assessed by plotting its profile relative to the baseline. Examples are presented in Subsections 4.1 and 4.2.

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

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

The contrast is best seen by writing

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

Thus $\hat{\mathcal{V}}_j(x)$ 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

$$\hat{\mathcal{M}}_j(x) = \hat{\mathcal{V}}_j(x) + \sum_{k=1}^{j-1} \mathcal{E}\{[\mathcal{E}(y_{t+j}|x_{t+k}) - \mathcal{E}(y_{t+j}|x_{t+k-1})][\mathcal{E}(y_{t+j}|x_{t+k}) - \mathcal{E}(y_{t+j}|x_{t+k-1})]'|x_t = x\}$$

Thus, $\hat{\mathcal{M}}_j(x)$ is a confounding of $\hat{\mathcal{V}}_j(x)$ with the variability of the conditional expectation of y_{t+j} as information accumulates between times t and $t+j-1$. This confounding co-mingles 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 forecasts of y_{t+j} given the history of the process up through time t, $\hat{\mathcal{M}}_j(x)$ 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 $g(y_{-J}, y_{-J+1}, \ldots, y_0)$ denote a time-invariant function of a stretch of $y's$ of length $J+1$. Put

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

for j=0,1,... where $x = (y'_{-L+1}, \ldots, y'_0)'$. (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

$$\{\hat{g}_j(x^+) - \hat{g}_j(x^0)\}_{j=0}^{\infty}$$

$$\{\hat{g}_j(x^-) - \hat{g}_j(x^0)\}_{j=0}^{\infty}$$

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

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 $\{y_j^r\}_{j=1}^{\infty}$, $r = 1, 2, \ldots, R$, 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 $x = (y'_{-L+1}, \ldots, y'_{-1}, y'_0)'$; y_2^r is a random draw from $f(y|x)$, with $x = (y'_{-L+2}, \ldots, y'_0, y_1^{r\prime})'$, and so forth.

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

$$\begin{aligned} \hat{g}_j(x) &= \int \cdots \int g(y_{j-J}, \ldots, y_j)[\prod_{i=0}^{j-1} f(y_{i+1}|y_{i-L+1}, \ldots, y_i)]\, dy_1 \cdots dy_j \\ &\doteq \frac{1}{R} \sum_{r=1}^{R} g(y_{j-J}^r, \ldots, y_j^r) \end{aligned}$$

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

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 are constructed by bootstrapping: 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 estimated from each data set and a profile computed from it. A 95% sup-norm confidence band is an ϵ-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

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

as in Subsection 4.2. The computations proceed as follows. Let $\{y_t^r\}_{t=L+1}^n$ denote the r^{th} simulated data set from the conditional density $\hat{f}(y|x)$, $r = 1, \ldots, R$, computed as described in Subsection 2.5. In these simulations, the initial conditions $x = (y_1', \ldots, y_L')'$ 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 $\{y_t^r\}$. For each r, the quantities $\hat{\mathcal{N}}_j^r$, $j = 1, \ldots, J$, are computed from $\hat{f}^r(y|x)$ as described in Subsection 2.3. Compute

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

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

$$\hat{\mathcal{N}}_j \pm M^{0.95} \qquad j = 1, 2, \ldots, J$$

2.7 Profile Bundles

Bollerslev and Engle (1989) develop interesting theoretical notions of co-integration and persistence in variance. Following Bollerslev and Engle (1989), we will say that $\{y_t\}$ is not integrated in mean if

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

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

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

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 $j \to \infty$.

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 prof'‿ $\{\hat{\mathcal{V}}_j(x)\}_{j=1}^{\infty}$. 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.

A reasonable empirical strategy for checking the sequences $\{\hat{y}_j(x)\}_{j=0}^{\infty}$ and $\{\hat{\mathcal{V}}_j(x)\}_{j=1}^{\infty}$ 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 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

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

obtained directly from the data should provide an adequate range of values. An application is in Subsection 4.3.

3 Examples Based on Parametric Models

This section uses a familiar parametric model 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

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

The corresponding impulse-response sequences are

$$\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0 = \alpha^j (y_0 + \delta y^+ - \lambda y_{-1})^2 - \alpha^j (y_0 - \lambda y_{-1})^2$$

$$\hat{\mathcal{V}}_j^- - \hat{\mathcal{V}}_j^0 = \alpha^j (y_0 + \delta y^- - \lambda y_{-1})^2 - \alpha^j (y_0 - \lambda y_{-1})^2$$

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 $\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0 = \hat{\mathcal{V}}_j^- - \hat{\mathcal{V}}_j^0 = \alpha^j$, 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 $\mathcal{E}_x(\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0) = \alpha^j$, where the expectation is with respect to the unconditional distribution of x_t, $f(x_t)$. To put this another way, if the sequence $\{\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0\}_{j=1}^{\infty}$ 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 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 $\{\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0\}_{j=1}^{\infty}$, 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.

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

$$\begin{aligned}
\lim_{j\to\infty}\{\hat{y}_j(x_{t_1}) - \hat{y}_j(x_{t_2})\} &= (y_{t_1} - y_{t_2})\lim_{j\to\infty}\lambda^j \\
\lim_{j\to\infty}\{\hat{\mathcal{V}}_j(x_{t_1}) - \hat{\mathcal{V}}_j(x_{t_2})\} &= \left[(y_{t_1} - \lambda y_{t_1-1})^2 - (y_{t_2} - \lambda y_{t_2-1})^2\right]\lim_{j\to\infty}\alpha^j
\end{aligned}$$

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 $\sigma_{t+1}^2 = \sigma^2 = const$, 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 $\hat{\mathcal{V}}_j(x) = \sigma^2$, for all j, and the corresponding impulse-response sequences are such that $\hat{\mathcal{V}}_j^+ - \hat{\mathcal{V}}_j^0 = \hat{\mathcal{V}}_j^- - \hat{\mathcal{V}}_j^0 = 0$, 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 λ reflects the confounding of mean and variance parameters characteristic of the j-step ahead mean square error path.

3.2 Computing Typical Impulse-Response Sequences

Unlike a linear model, 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 "typical" 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; this is the strategy employed in Section 4. The other is to draw x^0 randomly from the marginal density of x_t, compute the impulse-response sequence for each simulated x^0, and then average the sequences over the drawings. For an ARCH model these two strategies are equivalent, as noted in Subsection 3.1. In general, though, the strategies can be expected to give somewhat 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. Also, it is far less computationally demanding than the second. On the other hand, as noted in the discussion of Figures 5 through Figure 7 below, the baselines for the volatility and volume panels show upward drift. The baseline drift is due the fact that a vector with all elements constant represents an abnormally calm period in terms of market volatility. The second strategy will not show such baseline drift, and in that sense might come closer to providing a typical impulse response. But, this strategy is considerably more demanding computationally and, in particular, sup-norm confidence bands are beyond the reach of current equipment. In addition, the second computation is arguably less robust because it will be influenced be extreme $x's$.

There are some interestingly possibilities for modifying and extending the second approach. Rather than drawing the x's from the marginal stationary distribution and then applying the shock to each draw, we could search through the data for x configurations "close" to the perturbed unconditional mean vector. We could then average the conditional moment profiles over these x configurations.

No matter what, the computational demands of averaging over many possible x vectors are potentially very great. For each x vector, we must compute the conditional moment profile by Monte Carlo integration. In addition, the parametric bootstrapping would require the same n-fold increase in computations where n is the number of x configurations sampled. With our existing computing equipment it would be difficult to compute conditional moment profiles for much more than 50 configurations.

4 Applications: Stock Price and Volume Dynamics

A thorough review of the literature on the contemporaneous price-volume relationship is in Karpoff (1987). This literature documents a positive correlation between volume and the magnitude of price changes and a peaked and thick tailed price change distribution. Bollerslev, Ray, Jayaraman, and Kroner (1990) review the literature on the dynamics of volatility. That literature finds persistent volatility when past prices are the only conditioning variables. Gallant, Rossi, and Tauchen (1990) study the one-step ahead conditional density of price change and volume. They find that large price movements are followed by large volume one-step ahead. They also find that the "leverage effect" (asymmetry of the one-step ahead forecast of the variance when plotted against contemporaneous price) is sharply attenuated when volume is modeled jointly with prices. From an economic perspective, the empirical findings established in this literature represent the stylized facts that a fully articulated, general equilibrium model will have to confront.

In this section, we extend these findings to multi-step ahead dynamics. Specifically, we apply the methods suggested in Section 2 to the nonparametric estimate of the conditional density $f(y|x)$ for stock price changes and volume obtained by Gallant, Rossi, and Tauchen (1990) using 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 Δp_t and v_t series are adjusted to remove systematic calendar and trend effects, including a quadratic trend for volume, and are reasonably taken as jointly stationary. For reporting results, the Δp_t series is in units of percentage changes, that is, $\Delta p = 1.0$ is a 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 SNP method, which is summarized in Gallant and Tauchen (1990), 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 (1990) derive an efficient rejection method for sampling the fitted conditional density.

Gallant, Rossi, and Tauchen (1990) applied the SNP technique 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 thus 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.

4.1 Univariate Stock Price Dynamics

Figure 1 shows the dynamic impulse responses of future Δp 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

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

at initial conditions $x_t = x = x^+$, 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, $\mathcal{E}[\,\cdot\,|x_t = x]$ is evaluated

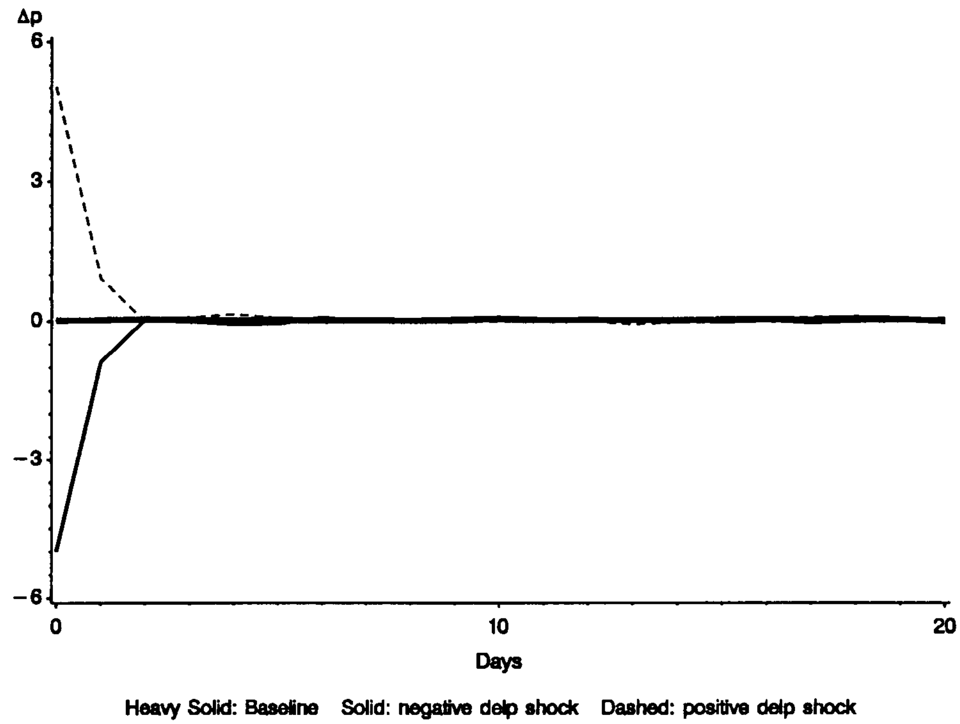


Figure 1: Impulse Response of Δp to Δp Shock in Univariate Fit

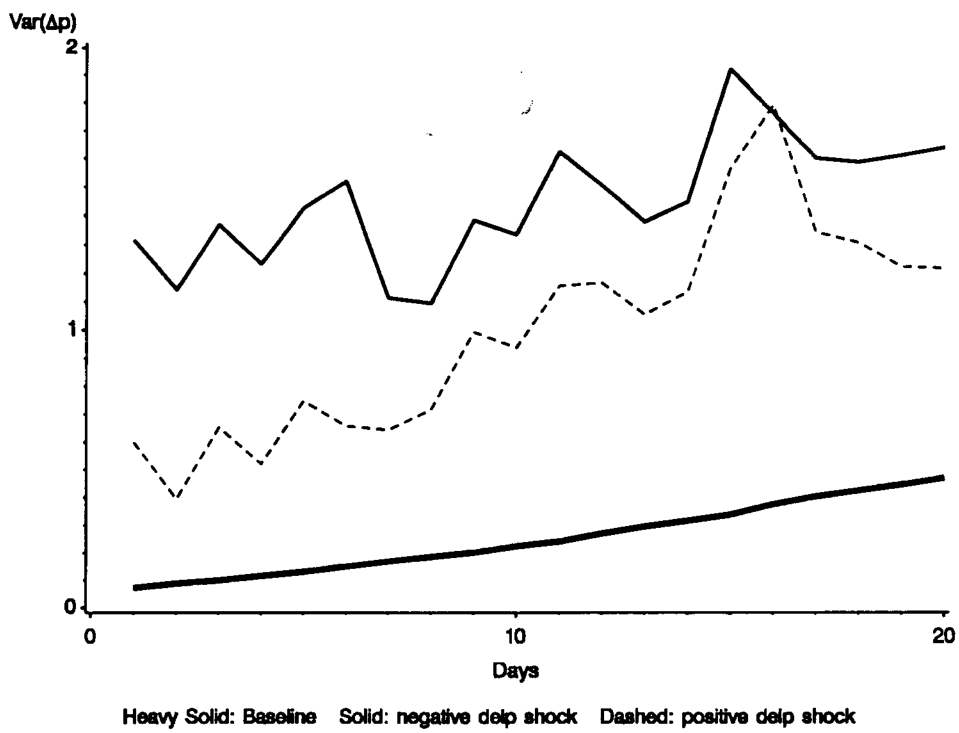


Figure 2: Impulse Response of Volatility to Δp Shock in Univariate Fit

by Monte Carlo integration. The initial conditions are

$$
\begin{aligned}
x^{+} &= (\mu_{\Delta p}, \mu_{\Delta p}, \cdots, \mu_{\Delta p})' + (0, 0, \cdots, 5.0)' \\
x^{0} &= (\mu_{\Delta p}, \mu_{\Delta p}, \cdots, \mu_{\Delta p})' \\
x^{-} &= (\mu_{\Delta p}, \mu_{\Delta p}, \cdots, \mu_{\Delta p})' + (0, 0, \cdots, -5.0)'
\end{aligned}
$$

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 Δp_{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 Δp_{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 suggests that the conditional mean of the $\{\Delta p_t\}$ series exhibits essentially no interesting higher order structure or serial dependence beyond lag one, and exhibits a mild linear dependence at lag one.

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

$$
\mathcal{V}_j(x) = \mathcal{E}[Var(y_{t+j}|x_{t+j-1})|x_t = x] \qquad j = 1, 2, \ldots, 20
$$

at each of the three initial x values, x^+, x^0, and x^-, defined above. $Var(y_{t+j}|x_{t+j-1})$ is evaluated analytically, $\mathcal{E}[\,\cdot\,|x_t = x]$ is evaluated by Monte Carlo integration. The impulse responses shown in Figure 2 indicate a clear 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 (1989). 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 $\Delta p_{t-j} = \mu_{\Delta p}$, $j \geq 1$, which is more quiescent that usual. Hence volatility must drift upwards from that point.

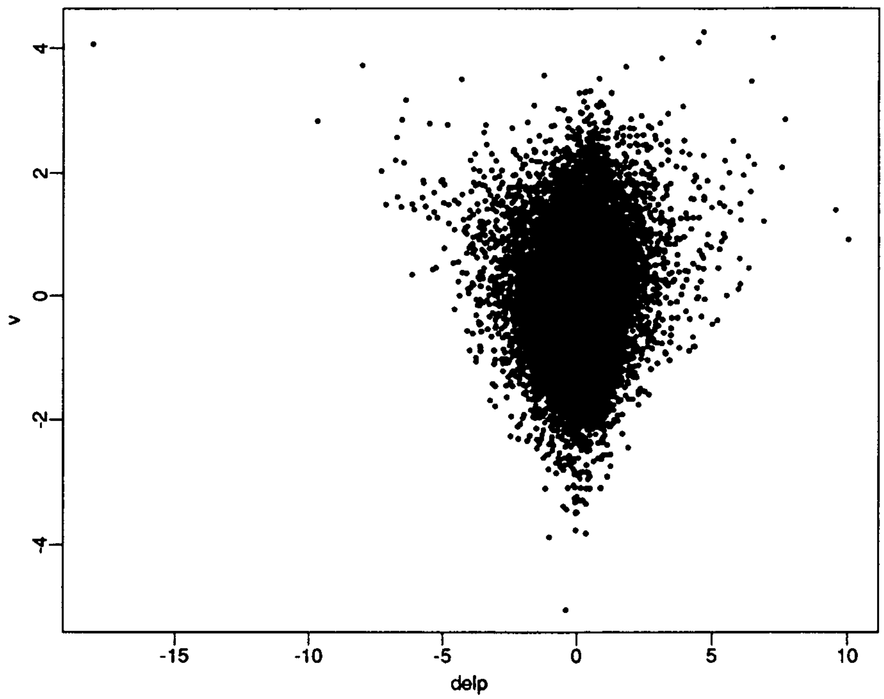


Figure 3: Scatter Plot of Adjusted $(\Delta p, v)$ Data

4.2 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 Δp_t is the adjusted log daily price change and v_t is the adjusted log daily volume. To ease interpretation, all results are reported with v_t expressed in units of unconditional standard deviations.

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, Δp_t, to the level of the other variable, v_t. There is essentially no relationship between Δp_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

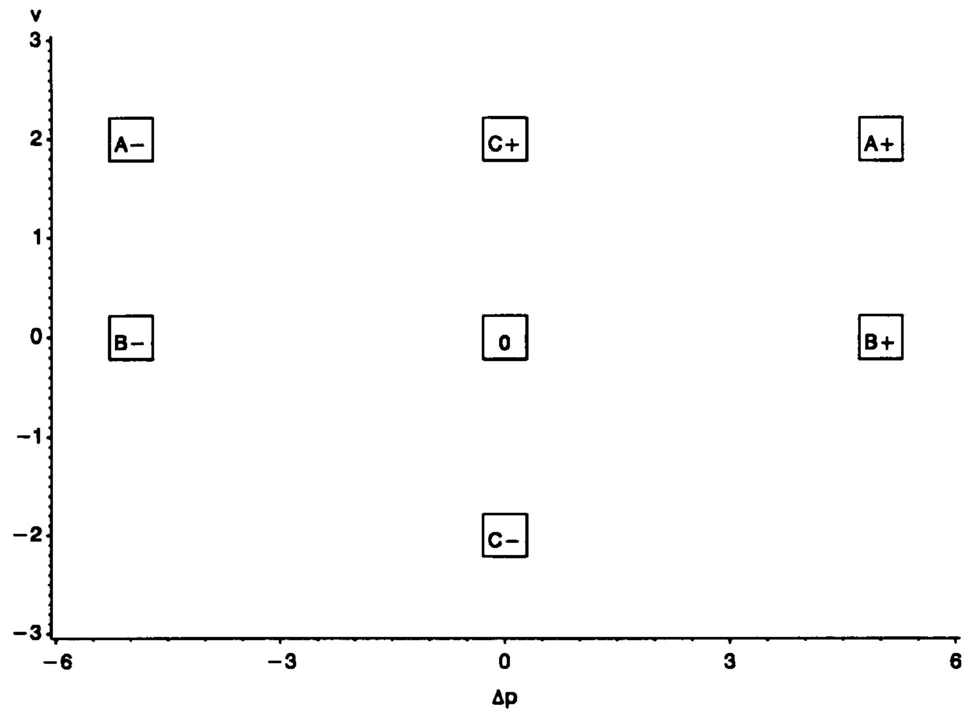


Figure 4: Experimental Design

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:

$$
\begin{aligned}
\delta y_A^+ &= (\;\;5.0,\;\; 2.0)' \\
\delta y_A^- &= (-5.0,\;\; 2.0)' \\
\delta y_B^+ &= (\;\;5.0,\;\; 0.0)' \\
\delta y_B^- &= (-5.0,\;\; 0.0)' \\
\delta y_C^+ &= (\;\;0.0,\;\; 2.0)' \\
\delta y_C^- &= (\;\;0.0, -2.0)'
\end{aligned}
$$

The A shocks are combined price-volume shocks where the price movements are ± 5.0 percent and volume is 2 standard deviations above its unconditional mean. The B shocks are pure price shocks of ± 5.0 percent with volume pinned at its mean. Finally, the C shocks are pure volume shocks of ± 2.0 standard deviations with no price movements. Figure 4 is a diagrammatic representation 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 analysis reported here concentrates exclusively on the effects of shocks on the one-step variance of Δp_t (volatility) and the one-step mean of v_t. The effects of price and volume shocks on forecasts of Δp_{t+j}, $j \geq 1$, are either very heavily damped, as in Figure 1 above, or negligible, and thus are not reported. Price and volume shocks do affect the one-step covariance of Δp_{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

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

evaluated x^+, x^0, and x^-. The volume responses, shown in the bottom panels, are computed as

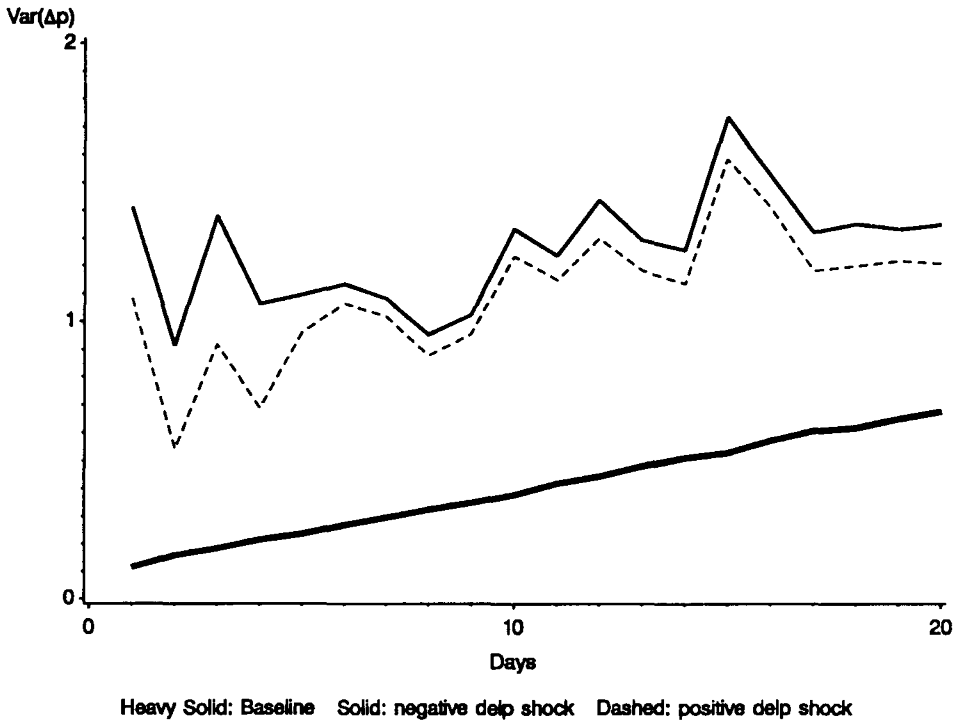
$$\hat{v}_j(x) = \mathcal{E}[\mathcal{E}(v_{t+j}|x_{t+j-1})|x_t = x] \qquad j = 0, 2, \cdots, 20$$

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

$$\begin{aligned} x_A^+ &= (\mu_y', \mu_y', \ldots, \mu_y')' + (0, 0, \cdots, \delta y_A^{+\prime})' \\ x_A^0 &= (\mu_y', \mu_y', \ldots, \mu_y')' \\ x_A^- &= (\mu_y', \mu_y', \ldots, \mu_y')' + (0, 0, \ldots, \delta y_A^{-\prime})' \end{aligned}$$

where μ_y' is the sample average of the $\{y_t\}$ process. For B and C shocks, the initial conditions are defined similarly with the appropriate δy in the rightmost place. Note that the shocks δy were determined from inspection of Figure 3, and are at the extreme edges of an unconditional point cloud. Relative to the conditional distribution of y_{t+1} given x_t at the means, these shocks are even that much more extreme.

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



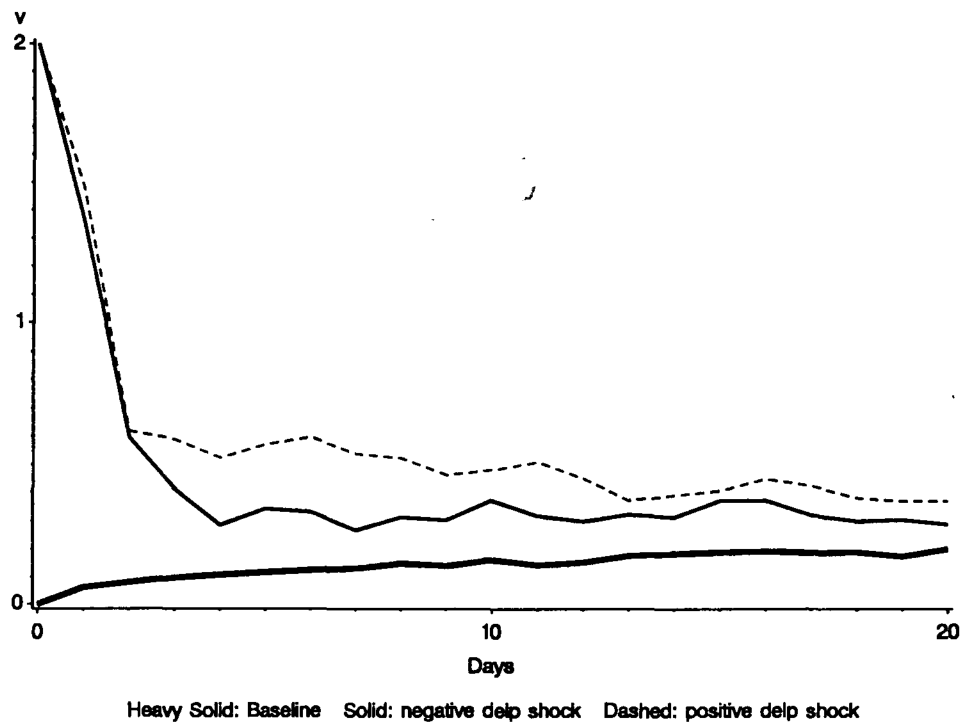
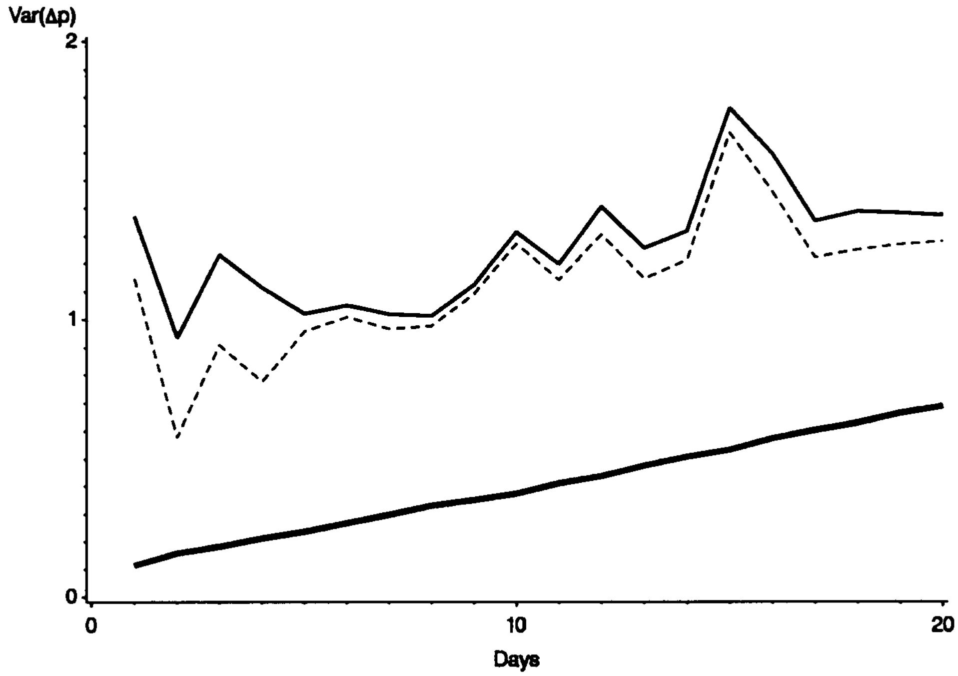
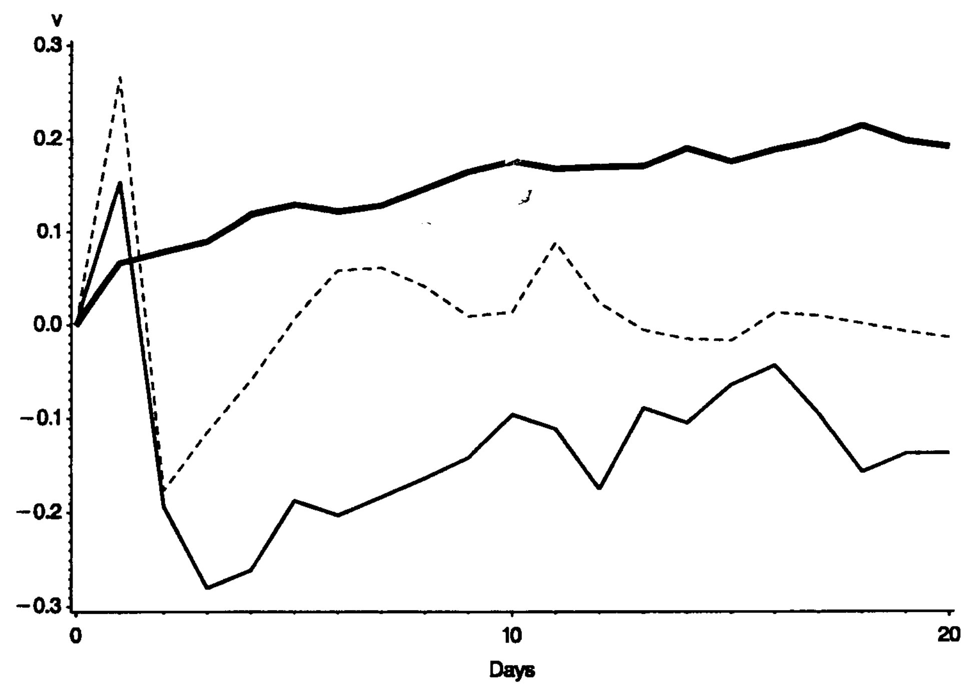


Figure 5: Impulse Responses of Volatility and Volume to Δp and v Shock (Type A)



Heavy Solid: Baseline   Solid: negative delp shock   Dashed: positive delp shock



Heavy Solid: Baseline   Solid: negative delp shock   Dashed: positive delp shock

Figure 6: Impulse Responses of Volatility and Volume to Pure Δp Shock (Type B)

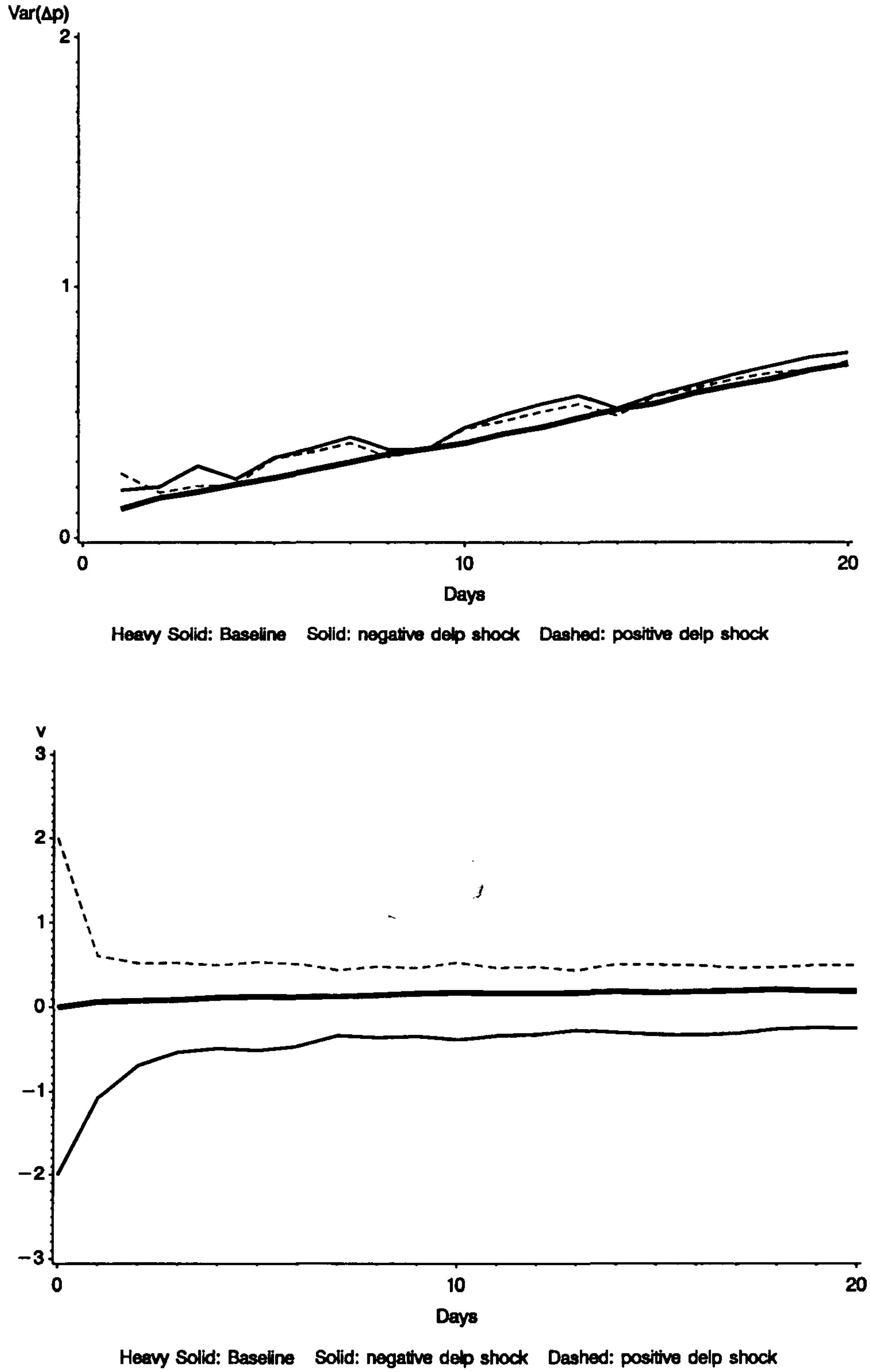


Figure 7: Impulse Responses of Volatility and Volume to Pure v Shock (Type C)

$Var(\Delta p_{t+j}|x_{t+j-1})$ are exact as a function of x_{t+j-1}, but their forecasts given $x_t = x$ must be done by Monte Carlo. This type of analysis differs from Hardouvelis (1990), who uses linear VAR analysis to examine the impulse responses of volatility and volume to changes in legal margin requirement changes.

Inspection of Figures 5, 6, and 7 reveals four characteristics of the behavior of volatility and volume dynamics. First, the asymmetry of volatility (leverage effect) is greatly attenuated in the bivariate system, as can be seen by comparing the top panels of Figures 5 and 6 to Figure 2. This extends to multi-step ahead volatility the finding of Gallant, Rossi, and Tauchen (1990), who detect attenuated leverage in one-step ahead volatility. Second, the impulse responses of volume to volume shocks are symmetric and extremely slowly damped, as can be seen in the bottom panels of Figures 5 and 7. Third, volume 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 R^2 calculations of Gallant, Rossi, and Tauchen (1990) and the findings of Schwert (1989). Fourth, pure price shocks, that is, type B shocks, increase volume in the very short run but decrease volume over the longer term, as 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 (1990) who examine one-step ahead volatility. The very slowly damped long-term negative effect is a new finding of this paper.

Of these four characteristics, the more interesting and novel 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. Figures 8 and 9 show 95 percent sup-norm confidence bands around quantities relevant for these findings, and thereby provide an indication of the statistical significance. Figure 8 shows a 95 percent confidence band around the estimates

$$\hat{\mathcal{V}}_{\Delta p,j}(x_A^+) - \hat{\mathcal{V}}_{\Delta p,j}(x_A^-) \qquad j = 1, 2, \cdots, 20$$

If the population volatility function is symmetric, that is, the leverage effect is absent, then the above differences should be jointly insignificant. That is, the confidence bands should include the horizontal axis, which is almost the case for the first few days after the shock and is certainly true beyond day five or so.

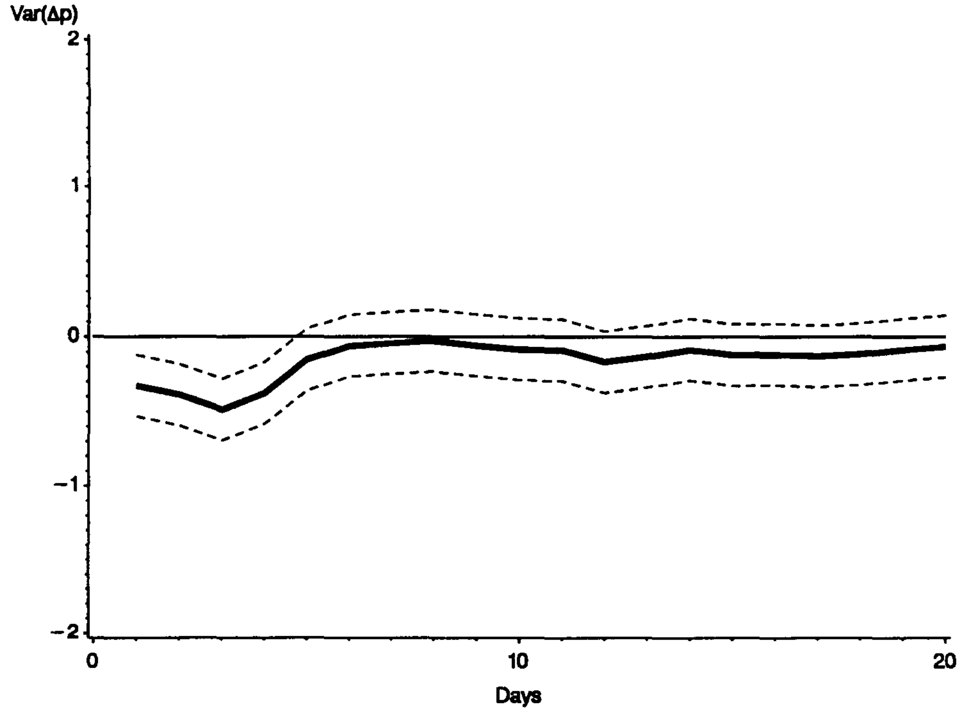


Figure 8: 95% Band for Differential Response of Volatility to Δp and v Shock (Type A)

The confidence bands are 95 percent sup-norm confidence bands obtained by bootstrapping the estimation as described in Subsection 2.6. Sup-norm bands have the appealing property that they contain the entire function of interest with 95 percent confidence, and are thereby more useful than bands obtained by plotting 95 percent pointwise confidence intervals. Furthermore, being based on a bootstrap procedure, the bands do not require the usual mean value (delta method) approximations of conventional asymptotics.

The top and bottom panels of Figure 9 show similarly constructed 95 percent confidence bands around the effects of pure price shocks on volume relative to baseline. The top panel shows

$$\hat{v}_j(x_B^+) - \hat{v}_j(x_B^0) \pm \Delta_+^{0.95} \qquad\qquad j = 1, 2, \cdots, 20$$

while the bottom panel shows

$$\hat{v}_j(x_B^-) - \hat{v}_j(x_B^0) \pm \Delta_-^{0.95} \qquad\qquad j = 1, 2, \cdots, 20$$

where the $\Delta's$ 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.

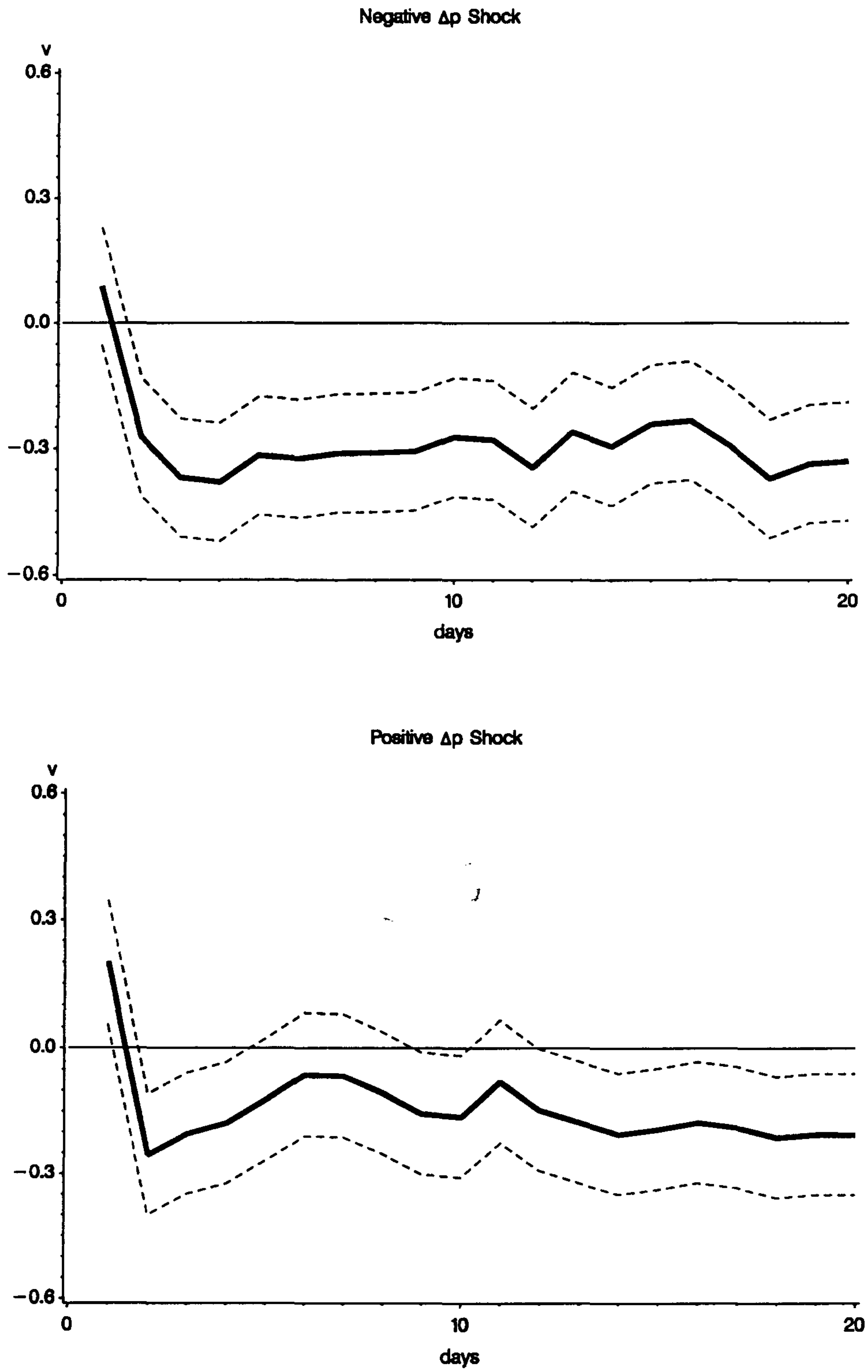


Figure 9: 95% Band for Differential Response of Volume to Pure Δp Shock (Type B)

4.3 Persistence in Variance

An important issue in empirical finance is the character of the persistence of shocks to volatility. As is well known (see Engle and Bollerslev, 1986), when a GARCH model is fitted to short-term financial price movements, the implied volatility process looks integrated, or very nearly so. Nelson (1990b) 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.

In Subsection 2.7 we described methods for studying persistence. We now implement these ideas for the SNP model fitted to the bivariate price and volume process.

The top panel of Figure 10 shows overplots of conditional mean profiles $\{\Delta\hat{p}_j(x)\}_{j=1}^{100}$, for $x \in \{x_t : t = 28, 156, 284, \ldots, 16,028\}$. Some subsetting of the 16,100 available x values is needed to keep the plots from becoming overly dense. The choice of every 128^{th} x_t, which generates 125 profiles, was determined by experimentation. The middle panel of Figure 10 shows overplots of similarly constructed profiles for the volume. For each t, the profile is computed out $J = 100$ steps.

The profiles for Δp and v in Figure 10 suggest that neither series is integrated. The Δp profiles are very strongly damped to the unconditional mean. The v profiles are much more slowly damped, but still they seem to show convergence setting in after 100 days.

The bottom panel of Figure 10 shows the conditional volatility profiles $\{\hat{\mathcal{V}}_{\Delta p,j}(x)\}_{j=1}^{100}$ for $x \in \{x_t : t = 28, 156, 284, \ldots, 16,028\}$ 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 seem to die out. The volatility shows little evidence of integration 100 steps out. While the bivariate fits do not exhibit much persistence in volatility, the conditional volatility profiles of the univariate fitted SNP are unstable and persistent. The conditional volatility profiles shown in Figure 2 for the univariate fit show no evidence of convergence to the baseline. This evidence from the univariate fit is consistent with the findings of researchers working with GARCH and E-GARCH models who find near unit-roots in the variance equation for models fitted to asset returns series (see Bollerslev, Chou, Jayaraman, and Kroner (1990) for an excellent survey of this work).

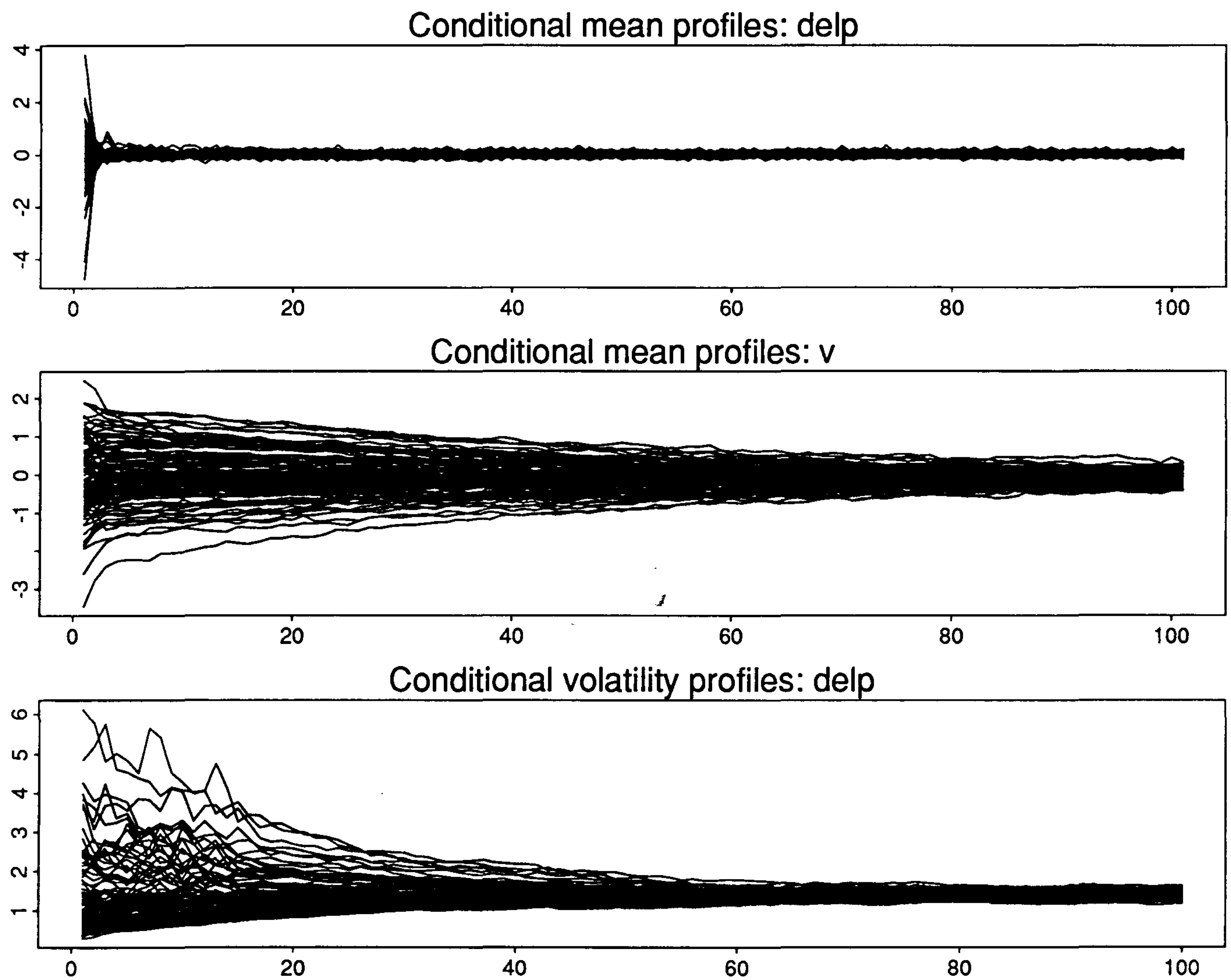


Figure 10: Profiles: $\{\Delta\hat{p}_j(x_t)\}_{j=1}^{100}$, $\{\hat{v}_j(x_t)\}_{j=1}^{100}$, $\{\hat{\mathcal{V}}_{\Delta p,j}(x_t)\}_{j=1}^{100}$, $t = 28, 156, \ldots, 16028$

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 (1990) suggest that the fitted bivariate SNP model fails to capture some of the movements of volatility at a yearly frequency. However, evidence in Gallant, Rossi, and Tauchen (1990) shows that volatility computed from the fitted bivariate SNP closely conforms to movements in volatility on a monthly and shorter term basis. The bias against long-run persistence in the fitted model may have some effect on the profiles computed for 100 days, although this bias is likely to be quite small.

5 Conclusion

This paper considers strategies for examining the dynamics of a nonlinear time series process as represented by a nonparametric estimate of the one-step ahead conditional density. For Markovian models, the one-step ahead conditional density of the series given finite lags of the past embodies all the information about the dynamics of the process. Three strategies are examined in detail: The first is an extension of the error shock analysis applied in conventional linear time series models. Multi-step ahead conditional moment profiles are examined by comparing a profile to a baseline profile. The second assesses the significance of a profile by comparing its sup-norm confidence band to a null profile. The third examines profile bundles for evidence of damping or persistence. Experimental designs for choosing an appropriate set of impulses are discussed. These strategies depend on efficient simulation of long realizations of the process by drawing from the conditional density.

We apply these strategies to explore the multi-period dynamics of an SNP density fitted to the bivariate process of the change in the Standard and Poor's stock price index and the transactions volume on the NYSE. One finding is that the multi-step ahead volatility of price changes responds symmetrically to positive and negative price shocks even though there is some evidence that the one-step ahead response is asymmetric. Another is that the effects of price shocks on the level of volume are different in the near and longer term. A price shock induces a one-step ahead increase in volume followed by a long run decline in volume.

Following Bollerslev and Engle (1989), persistence in variance is defined as the extreme sensitivity of volatility to changes in initial conditions. Conditional volatility profiles up to 100 days ahead reveal that the effect of initial shocks damps off suggesting that volatility changes are not very persistent in the fitted bivariate SNP model. This finding contrasts markedly with the findings for both SNP and conventional parametric ARCH models fitted to the univariate price change series.

6 References

Doan, Thomas, Robert Litterman, and Christopher Sims (1984), "Forecasting and Conditional Projection using Realistic Prior Distributions," *Econometric Reviews* 3, 1–100.

Bollerslev, Tim, Ray Chou, Narayanan Jayaraman, and Kenneth Kroner (1990), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence, with Suggestions for Future Research," working paper, Northwestern University.

Bollerslev, Tim, and Robert Engle (1989), "Common Persistence in Conditional Variances," working paper, Northwestern University.

Engle, Robert (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflations," *Econometrica* 50, 987–1008.

Engle, Robert, and Tim Bollerslev (1986), "Modeling the Persistence of Conditional Variance," *Econometric Reviews* 5, 1–50.

Engle, Robert, 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.

Gallant, A. Ronald, and George Tauchen (1990), "A Nonparametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation," working paper, Duke University.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen (1990), "Stock Prices and Volume," working paper, North Carolina State University.

Hardouvelis, Gikas (1990), "Margin Requirements, Volatility, and the Transitory Component of Stock Prices," *American Economic Review* 80, 736–762.

Karpoff, J. (1987), "The Relation between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis* 22, 109–126.

Litterman, R. and L. Weiss (1985), "Money, Interest Rates and Output," *Econometrica* 53, 129–156.

Liu, Regina (1990), "On a Notion of Data Depth Based on Random Simplexes," *Annals of Statistics* 18, 405–414.

Nelson, Daniel (1990a), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," working paper, University of Chicago.

Nelson, Daniel (1990b), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, forthcoming.

Schwert, G. William (1989), "Why Does Stock Volatility Change Over Time?" *Journal of Finance* 44, 1115–1154.

Sims, Christopher (1980), "Macroeconomics and Reality," *Econometrica* 48, 1–48.

Sims, Christopher (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.