

**A NONPARAMETRIC APPROACH
TO NONLINEAR TIME SERIES ANALYSIS:
ESTIMATION AND SIMULATION**

By

A. Ronald Gallant

and

George Tauchen

IMA Preprint Series # 705

September 1990

A Nonparametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation ¹

A. Ronald Gallant
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203 USA

George Tauchen
Department of Economics
Duke University
Durham, NC 27706 USA

June 1990

¹Research 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 a method of nonlinear time series analysis suitable for nonlinear, stationary, multivariate processes whose one-step-ahead conditional density depends on a finite number of lags. Such a density can be represented as a Hermite expansion. Certain parameters of the expansion can be set to imply sharp restrictions on the process such as a pure VAR, a pure ARCH, a nonlinear process with homogeneous innovations, etc. The model is fitted using maximum likelihood procedures on a truncated expansion together with a model selection strategy that determines the truncation point. The estimator is consistent for the true density with respect to a strong norm. The norm is strong enough to imply consistency of evaluation functionals and moments of the conditional density. We describe a method of simulating from the density. Simulation can be used for a great variety of applications. In this paper, we give special attention to using simulations to set sup-norm confidence bands. Fortran code is available via ftp anonymous at [ccvr1.cc.ncsu.edu](ftp://ccvr1.cc.ncsu.edu) (128.109.212.20) in directory `pub/arg/snp`; alternatively, it is available from the authors in the form of a DOS formatted diskette. The code is provided at no charge for research purposes without warranty. An appendix to this paper describes its use.

1 Introduction

1.1 SNP Models

Empirical modeling of time series data primarily concerns making inferences about the intertemporal law of motion of the observed process. For a strictly stationary, multivariate process, the law of motion is the one-step-ahead conditional density of the process given its past. Under stationarity, the one-step-ahead conditional density is time invariant and embodies all probabilistic information about the process. The conditional density is thus naturally viewed as the fundamental statistical object of interest.

This paper describes a nonparametric-type method, based on a modified Hermite series expansion, for estimation of the conditional density of a general nonlinear stationary process. The method was first proposed by Gallant and Tauchen (1989) in connection with an asset pricing application, and has since undergone a number of refinements. Estimation entails using maximum likelihood procedures on a truncated expansion together with a model selection strategy that determines the truncation point. Under reasonable regularity conditions, the estimator is consistent for the true density under a norm that is strong enough to imply consistency of evaluation functionals and conditional moments.

The method is termed seminonparametric, or SNP, to suggest that it lies halfway between parametric and nonparametric procedures. The leading term of the series expansion is an established parametric model known to give a reasonable approximation to the process; higher order terms capture departures from that model. With this structure, the SNP approach does not suffer from the curse of dimensionality to the same extent as kernels and splines. In regions where data are sparse, the leading term helps to fill in smoothly between data points. Where data are plentiful, the higher order terms accommodate deviations from the leading term and fits are comparable to the kernel estimates proposed by Robinson (1983).

For time series data, an expansion based on Hermite series is particularly attractive on the basis of both modeling and computational considerations. In terms of modeling, the Gaussian component makes it easy to subsume into the leading term familiar time series models, including vector autoregressive models and ARCH models (Engle, 1982). These models are

generally considered to give excellent first approximations in a wide variety of applications. In terms of computation, a Hermite density is easy to evaluate and differentiate. Also, its moments are easy to evaluate because they correspond to higher moments of the normal, which are computable using standard recursions. Finally, as described below, simulation from the fitted SNP model is quite practicable, which is a capability with numerous applications such as bootstrapping, assessing issues of long-term dependence, etc.

1.2 Refinements and Extensions

A sequence of empirical applications, beginning with Gallant and Tauchen (1989), has stimulated extensions and refinements of the SNP methodology. The original asset-pricing application was a limited information maximum likelihood situation where both the likelihood (which is the product of one-step-ahead conditional densities) and the Euler conditions (structural equations) had to have nonparametric properties and be mathematically convenient. This naturally led to a series expansion type of approach so that standard algorithms could be used to optimize the likelihood subject to the Euler conditions.

In his dissertation work, Hussey (1989a) developed the upward fitting and perturbation strategy that we use. In related work, Hussey (1989b) used SNP methods in conjunction with kernel methods to document the nonlinear structure of industrial employment data; Hussey (1989c) developed an approach for using an SNP estimate as a standard against which to compare the predictions of simulated nonlinear structural models. Brunner (1989) adapted the SNP method for the purpose of examining asymmetries in U.S. business cycles.

Extensions to better adapt the method to markedly conditionally heteroskedastic processes such as exchange rate data were developed by Gallant, Hsieh and Tauchen (1989). Further extensions to robustify the methodology against extremely heavy tailed processes such as real interest rate data are reported Gallant, Hansen and Tauchen (1990). Processes such as bivariate stock price and volume series can require a high degree Hermite polynomial to fit them which generates a plethora of irrelevant interactions. Gallant, Rossi and Tauchen (1990) described filters to remove them.

Our description of the SNP nonlinear time series methodology in Section 2 incorporates the above refinements.

1.3 Inference

Strategies based on series expansions are also used for the estimation of systems of demand equations. In this context, the data are presumed i.i.d. The estimation is thereby simpler, and the statistical theory is correspondingly more advanced: Elbadawi, Gallant and Souza (1983); Andrews (1989); Eastwood and Gallant(1987), and Gallant and Souza (1989). To the extent that the results of this work carry over to the time series context, then standard finite-parameter inference procedures should be asymptotically valid for SNP models, despite the use data driven rules for choosing the number of terms of the series to include in the expansion. That means that inferences based on the Wald test, the likelihood ratio test, and the Lagrange multiplier test should achieve their nominal p-values asymptotically.

At the same time, given that the asymptotic justification for applying standard inference procedures is conjectural with details still to be worked out, a parametric bootstrap off a fitted model appears to be a reasonable complementary strategy for guiding statistical inference. Also, some computations such as sup-norm confidence bounds are easier to compute using a parametric bootstrap. Using the bootstrap requires a means to sample from an SNP one-step-ahead conditional density.

In Section 3 we develop an algorithm for sampling an SNP density that is surprisingly efficient. This is the original contribution of the paper. In Section 4 we illustrate the SNP methodology and compare with kernel methods by estimating the leverage function which is the variance of the one-step-ahead conditional density for stock returns considered as a function of the most recent return, holding other conditioning variables fixed at their unconditional means. We compute estimates of the leverage function from data on stock returns alone, a bivariate stock returns and volume series, and a bivariate stock returns and bond returns series. We display volatility scatter plots which one can use to form subjective estimates for comparison and compute sup-norm confidence limits on the leverage function using the algorithm developed in Section 3:

Fortran code implementing the SNP methodology is available via ftp anonymous at ccvrl.cc.ncsu.edu (128.109.212.20) in directory pub/arg/snp; alternatively, it is available from the authors in the form of a 5-1/4 inch, 2S/2D, DOS formatted diskette. In the Appendix, we walk the user through an application using this code. Data, code, and output

for this application are included. The code is provided at no charge for research purposes without warranty.

2 Estimation and Model Selection

In this section, we describe an estimation strategy for nonlinear time series analysis proposed by Gallant and Tauchen (1989) and its extensions. These extensions are: an ARCH leading term, which better adapts the method to markedly conditionally heteroskedastic processes, proposed by Gallant, Hsieh and Tauchen (1989); a logistic transformation, which robustifies the methodology against extremely heavy tailed processes, proposed by Gallant, Hansen and Tauchen (1990); and, filters, which remove the high order interactions in fits to multiple time series, proposed by Gallant, Rossi and Tauchen (1990).

The derivation of the SNP model that we present here provides a fundamental understanding of the model so that one can easily appreciate the implications of the tuning parameters. It does not provide the mathematical connection with the results of Gallant and Nychka (1987) that is required for theoretical work. For this, see Gallant, Hsieh and Tauchen (1989).

2.1 Estimation

As stated above, the SNP method is based on the notion that a Hermite expansion can be used as a general purpose approximation to a density function. Letting z denote an M -vector, the particular Hermite expansion employed has the form $h(z) \propto [P(z)]^2 \phi(z)$ where $P(z)$ denotes a multivariate polynomial of degree K_z and $\phi(z)$ denotes the density function of the (multivariate) Gaussian distribution with mean zero and the identity matrix as its variance-covariance matrix. The constant of proportionality is $1 / \int [P(s)]^2 \phi(s) ds$ which makes $h(z)$ integrate to one. Because $[P(z)]^2 / \int [P(s)]^2 \phi(s) ds$ is a homogeneous function of the coefficients of the polynomial $P(z)$, the coefficients can only be determined to within a scalar multiple. To achieve a unique representation, the constant term of the polynomial part is put to one.

A change of variables using the location-scale shift $y = Rz + \mu$ where R is an upper

triangular matrix and μ is an M -vector, gives a density that is easier to interpret

$$f(y|\theta) \propto \{P[R^{-1}(y - \mu)]\}^2 \{\phi[R^{-1}(y - \mu)]/|\det(R)|\}$$

The constant of proportionality is the same as above, $1/\int [P(s)]^2 \phi(s) ds$. Because $\{\phi[R^{-1}(y - \mu)]/|\det(R)|\}$ is the density function of the M -dimensional, multivariate, Gaussian distribution with mean μ and variance-covariance matrix $\Sigma = RR'$, and because the leading term of the polynomial part is one, the leading term of the entire expansion is proportional to the multivariate, Gaussian density function. Denote the Gaussian density of dimension M with mean vector μ and variance-covariance matrix Σ by $n_M(y|\mu, \Sigma)$ and write

$$f(y|\theta) \propto [P(z)]^2 n_M(y|\mu, \Sigma)$$

where $z = R^{-1}(y - \mu)$ for the density above.

When K_z is put to zero, one gets $f(y|\theta) = n_M(y|\mu, \Sigma)$ exactly. When K_z is positive, one gets a Gaussian density whose shape is modified due to multiplication by a polynomial in $z = R^{-1}(y - \mu)$. The shape modifications thus achieved are rich enough to accurately approximate densities from a large class that includes densities with fat, t-like tails, densities with tails that are thinner than Gaussian, and skewed densities (Gallant and Nychka, 1987).

The parameters θ of $f(y|\theta)$ are made up of the coefficients of the polynomial $P(z)$ plus μ and R and are estimated by maximum likelihood. Equivalent to maximum likelihood but more stable numerically is to estimate θ in a sample of size n by minimizing $s_n(\theta) = (-1/n) \sum_{t=1}^n \ln[f(y_t|\theta)]$. As mentioned above, if the number of parameters p_θ grows with the sample size n , the true density and various features of it such as derivatives and moments are estimated consistently (Gallant and Nychka, 1987).

This basic approach can be adapted to the estimation of the conditional density of a multiple time series $\{y_t\}$ that has a Markovian structure. Here, the term Markovian structure is taken to mean that the conditional density of the M -vector y_t given the entire past y_{t-1}, y_{t-2}, \dots depends only on L lags from the past. For notational convenience, we collect these lags together in a single vector denoted as x_{t-1} which has length $M \cdot L$

$$x_{t-1} = (y'_{t-L}, \dots, y'_{t-2}, y'_{t-1})'$$

Note particularly that the serial order of the data is preserved in writing x_{t-1} . (Preserving order allows x_{t-1} to be passed to a subroutine as a pointer into an array.)

To approximate the density of $\{y_t\}$ using the ideas above, begin with a sequence of innovations $\{z_t\}$. First consider the case of homogeneous innovations; that is, the distribution of z_t does not depend on x_{t-1} . Then, as above, the density of z_t can be approximated by $h(z) \propto [P(z)]^2 \phi(z)$ where $P(z)$ is a polynomial of degree K_z . Follow with the location-scale shift $y_t = Rz_t + \mu_x$ where μ_x is the linear function

$$\mu_x = b_0 + Bx_{t-1}$$

The density that results is

$$f(y|x, \theta) \propto [P(z)]^2 n_M(y|\mu_x, \Sigma)$$

where $z = R^{-1}(y - \mu_x)$. The constant of proportionality is as above, $1/\int [P(s)]^2 \phi(s) ds$. The leading term of the expansion is $n_M(y|\mu_x, \Sigma)$ which is a Gaussian vector autoregression or Gaussian VAR.

When K_z is put to zero, one gets $n_M(y|\mu_x, \Sigma)$ exactly. When K_z is positive, one gets a density that can approximate over a large class whose shape is constant with respect to variation in x ; that is, a class which is conditionally homogeneous. Only the first moment of the density depends upon x .

To approximate conditionally heterogeneous processes, proceed as above but let each coefficient of the polynomial $P(z)$ be a polynomial of degree K_x in x . A polynomial in z of degree K_z whose coefficients are polynomials of degree K_x in x is, of course, a polynomial in (z, x) of degree $K_z + K_x$ (with some of the coefficients put to zero). Denote it by $P(z, x)$. The form of the density with this modification is

$$f(y|x, \theta) \propto [P(z, x)]^2 n_M(y|\mu_x, \Sigma)$$

where $z = R^{-1}(y - \mu_x)$. The constant of proportionality is $1/\int [P(s, x)]^2 \phi(s) ds$. When K_x is zero, the density reverts to the density above. When K_x is positive, the shape of the density will depend upon x . Thus, all moments can depend upon x and the density can, in principal, approximate any form of conditional heterogeneity. (Gallant and Tauchen, 1989; Gallant, Hsieh, and Tauchen, 1989).

In practice, especially in applications to data from financial markets, the second moment can exhibit marked dependence upon x . In an attempt to track the second moment, K_x can get quite large. To keep K_x small when data are markedly conditionally heteroskedastic, the leading term of the expansion can be put to a Gaussian ARCH rather than a Gaussian VAR. This is done by letting R be a linear function of the absolute values of the elements of the vectors y_{t-L_r} through y_{t-1} after y_{t-L_r} through y_{t-1} have been centered and scaled to have mean zero and identity variance-covariance matrix.

The centering and scaling is easiest to accomplish by (1) computing estimates of the unconditional mean and variance

$$\bar{y} = (1/n) \sum_{t=1}^n \tilde{y}_t$$

$$S = (1/n) \sum_{t=1}^n (\tilde{y}_t - \bar{y})(\tilde{y}_t - \bar{y})'$$

where \tilde{y}_t denotes the raw data, and (2) applying the methods above to

$$y_t = S^{-1/2}(\tilde{y}_t - \bar{y})$$

where $S^{-1/2}$ denotes the Cholesky factor of the inverse of S . That is, just replace the raw data $\{\tilde{y}_t\}$ by the centered and scaled data $\{y_t\}$ throughout. Because of the location-scale shift $y = Rz + \mu$ the consistency results cited above are not affected by the transformation from \tilde{y}_t to y_t .

With this done, the variance-covariance matrix can be written

$$\Sigma_x = R_x R_x'$$

$$\text{vech}(R_x) = p_0 + P|x|$$

where $\text{vech}(R)$ denotes a vector of length $M(M+1)/2$ containing the elements of the upper triangle of R and $|x|$ denotes x with its elements replaced by their absolute values.

The classical ARCH (Engle, 1982) has Σ_x depending on a linear function of squared lagged residuals. The SNP version of ARCH is more akin to the suggestions of Nelson (1989) and Davidian and Carroll (1987). Denoting the function in L_r lags of y_t by R_x and letting $\Sigma_x = R_x R_x'$, the form of the conditional density becomes

$$f(y|x, \theta) \propto [P(z, x)]^2 n_M(y|\mu_x, \Sigma_x)$$

where $z = R_x^{-1}(y - \mu_x)$. The constant of proportionality is as above, $1/\int [P(s, x)]^2 \phi(s) ds$. The parameter vector θ denotes the coefficients of the polynomial $P(z, x)$ and the parameters of the Gaussian ARCH $n_M(y|\mu_x, \Sigma_x)$ collected together. The parameters are estimated by minimizing

$$s_n(\theta) = (-1/n) \sum_{t=1}^n \ln[f(y_t|x_{t-1}, \theta)]$$

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

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

In summary, L_r and L_p determine the location-scale shift $y = R_x z_t + \mu_x$ and hence determine the nature of the leading term of the expansion. The number of lags in the location shift μ_x is the overall lag length L which is the maximum of L_r and L_p . The number of lags in the scale shift R_x is L_r . The number of lags that go into the x part of the polynomial $P(z, x)$ is L_p . The parameters K_z and K_x determine the degree of $P(z, x)$ and hence the nature of the innovation process $\{z_t\}$. I_z and I_x determine filters that suppress interactions when set to positive values.

Putting certain of the tuning parameters to zero implies sharp restrictions on the process $\{y_t\}$, the more interesting of which are displayed in Table 1.

Time series data often contains extreme or outlying observations, particularly data from financial markets. This is not a particular problem when the extreme value is considered as a y_t because it just fattens up the tails of the estimated conditional density. However, once it becomes a lag or x_{t-1} , and one has an ARCH leading term in the expansion, the optimization algorithm used to minimize $s_n(\theta)$ can use an extreme value in x_{t-1} to fit an element of y_t nearly exactly, reduce the corresponding conditional variance to near zero, and

Parameter setting	Characterization of $\{y_t\}$
$L_r = 0, L_p = 0, K_z = 0, K_x = 0$	iid Gaussian
$L_r = 0, L_p > 0, K_z = 0, K_x = 0$	Gaussian VAR
$L_r = 0, L_p > 0, K_z > 0, K_x = 0$	non-Gaussian VAR, homogeneous innovations
$L_r > 0, L_p = 0, K_z = 0, K_x = 0$	Gaussian ARCH
$L_r > 0, L_p > 0, K_z > 0, K_x = 0$	non-Gaussian ARCH, homogeneous innovations

Table 1: Restrictions Implied by Settings of the Tuning Parameters

inflate the likelihood. This problem is endemic to procedures that adjust variance on the basis of observed explanatory variables.

One can compensate for this effect by an additional transformation

$$\hat{x}_i = (4/c) \exp(cx_i) / [1 + \exp(cx_i)] - 2/c \quad i = 1, \dots, M \cdot L$$

with $c=1/2$ (x_i denotes the elements of x_{t-1}). This is a one-to-one (logistic) transformation that has a negligible effect on values of x_i between -3.5 and 3.5 but progressively compresses values that exceed ± 3.5 so they are bounded by ± 4 . The inverse transformation is $x = (1/c) \ln[(2 + c\hat{x}) / (2 - c\hat{x})]$. This transformation is roughly equivalent to variable bandwidth selection in kernel density estimation. Because it affects only x , and not y , the asymptotic properties of SNP estimators discussed above are unaltered. Also, note the order in which the transformations are to be applied

$$\begin{array}{cccccc}
\tilde{y}_t & \rightarrow & y_t & \rightarrow & x_{t-1} & \rightarrow & \hat{x}_{t-1} & \rightarrow & \mu_x, \Sigma_x \\
\text{raw} & \rightarrow & \text{centered,} & \rightarrow & \text{lagged} & \rightarrow & \text{logistic} & \rightarrow & \text{mean,} \\
\text{data} & & \text{scaled data} & & \text{data} & & \text{data} & & \text{variance}
\end{array}$$

In the sequel, we shall use x to mean either x or \hat{x} ; in applications, substitute \hat{x} for x in any formula involving x if the logistic transform is employed.

2.2 Model Selection

A model selection strategy that seems to work well is described and illustrated below. It is based on experience acquired in the sequence of applications: Gallant and Tauchen (1989);

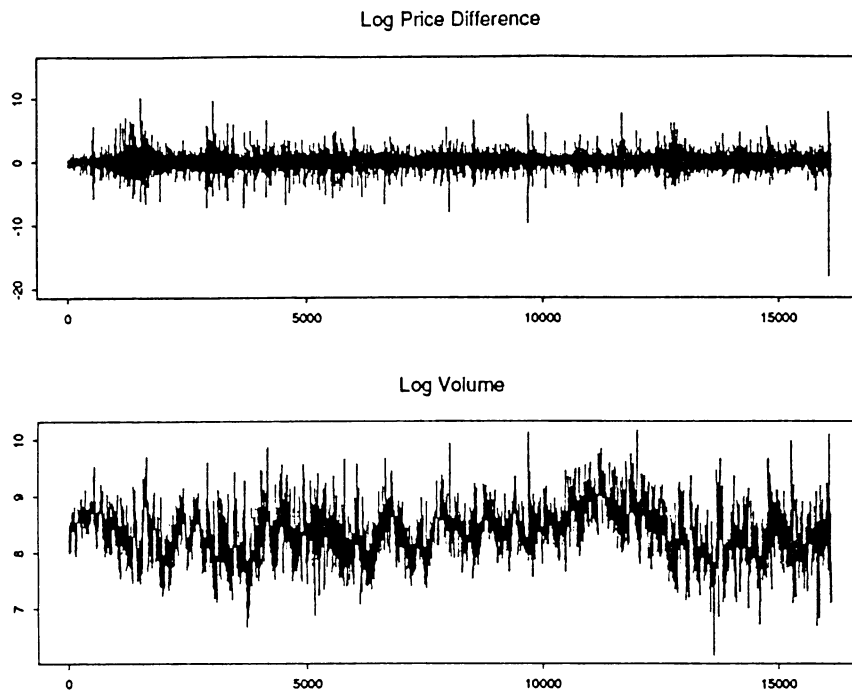


Figure 1: Daily NYSE S&P, 1928-1987, Adjusted for Calendar Effects

Hussey and Tauchen (1989); Gallant, Hsieh, and Tauchen (1989); Gallant, Hansen, and Tauchen (1990), Gallant; Rossi and Tauchen (1990). These papers also recount experience with number of alternative model selection procedures that did not work well, notably Akaike's (1969) criterion, the Brock, Dechert, and Scheinkman (1987) statistic, and upward likelihood ratio testing.

The data used for illustration is a bivariate series of log differences of daily closing values of the Standard and Poor's composite stock index (denoted as y_{1t} or Δp_t) and (2) the log daily volume of shares traded on the New York Stock Exchange (denoted as y_{2t} or v_t) from 1928 to 1987 ($t=1, \dots, 16127$) after adjustment for calendar effects. The data were taken from Gallant, Rossi and Tauchen (1990) which lists sources and describes the adjustment process. The data are plotted in Figure 1. The most notable feature of the data is the marked conditional heteroskedasticity as indicated by alternating periods of quiescence and volatility, especially in the price series.

The model selection procedure is straightforward. We first use the Schwarz criterion (Schwarz, 1978; Potscher, 1989) to move along an upward expansion path until an adequate model is determined. This Schwarz-preferred model is then subjected to a battery of specification tests designed to check for features of the data that the model fails to adequately

L_r	L_p	K_z	I_z	K_x	I_x	p_θ	Obj	Schwarz	Specification tests										
									Annual dummies				Ten lag cubic						
									Mean		Variance		Mean		Variance				
									W	p-val	W	p-val	W	p-val	W	p-val			
Data																			
2	2	0	0	0	0	25	2.1083	2.1158	.617	.0000	.577	.0000	.244	.0000	.210	.0000			
3	3	0	0	0	0	35	2.0707	2.0812	.974	.0001	.928	.0000	.900	.0000	.930	.0000			
4	4	0	0	0	0	45	2.0519	2.0655	.984	.0001	.942	.0000	.950	.0000	.960	.0001			
4	4	4	0	0	0	59	1.9187	1.9365	.987	.0001	.942	.0000	.941	.0000	.958	.0000			
4	4	4	0	1	0	179	1.8624	1.9162	.985	.0001	.950	.0000	.963	.0001	.977	.0001			
6	4	0	0	0	0	65	2.0293	2.0489	.987	.0001	.949	.0001	.963	.0001	.971	.0001			
8	4	0	0	0	0	85	2.0177	2.0432	.988	.0001	.950	.0001	.969	.0001	.971	.0001			
10	4	0	0	0	0	105	2.0085	2.0400	.989	.0001	.951	.0001	.975	.0001	.971	.0001			
12	4	0	0	0	0	125	2.0022	2.0397	.990	.0023	.952	.0001	.976	.0001	.971	.0001			
12	4	4	2	0	0	134	1.9110	1.9512	.978	.0001	.948	.0001	.958	.0001	.953	.0001			
12	4	4	2	1	0	214	1.8634	1.9277	.985	.0001	.954	.0001	.983	.0001	.981	.0001			
12	4	4	1	1	0	232	1.8260	1.8957	.989	.0005	.954	.0001	.983	.0001	.985	.0001			
12	4	4	0	1	0	259	1.8180	1.8958	.989	.0005	.955	.0001	.983	.0001	.986	.0001			
12	4	4	2	2	1	294	1.8358	1.9241	.985	.0001	.958	.0001	.991	.0962	.987	.0001			
12	6	4	2	1	0	254	1.8516	1.9279	.986	.0001	.956	.0001	.984	.0001	.976	.0001			
16	4	0	0	0	0	165	1.9950	2.0445	.990	.0109	.954	.0001	.976	.0001	.971	.0001			
16	4	4	2	0	0	174	1.9058	1.9581	.979	.0001	.950	.0001	.957	.0001	.953	.0001			
16	4	4	2	1	0	254	1.8577	1.9340	.985	.0001	.954	.0001	.984	.0001	.978	.0001			
16	4	4	2	2	1	334	1.8275	1.9278	.985	.0001	.959	.0001	.991	.0961	.988	.0001			
16	4	4	1	2	1	368	1.7848	1.8953	.988	.0001	.961	.0001	.992	.1674	.990	.0027			
16	4	4	0	2	1	419	1.7735	1.8994	.989	.0006	.960	.0001	.992	.1620	.989	.0005			

Table 2: Optimized Likelihood and Residual Diagnostics

approximate.

The Schwarz criterion is computed as

$$s_n(\hat{\theta}) + (1/2)(p_\theta/n) \ln(n)$$

with small values of the criterion preferred. The criterion rewards good fits as represented by small $s_n(\hat{\theta})$ but uses the term $(1/2)(p_\theta/n) \ln(n)$ to penalize good fits gotten by means of excessively rich parameterizations. The criterion is conservative in that it selects sparser parameterizations than the Akaike information criterion which uses the penalty term p_θ/n in place of $(1/2)(p_\theta/n) \ln(n)$. Schwarz is also conservative in the sense that it is at the high end of the permissible range of penalty terms in certain model selection settings (Potscher, 1989).

To illustrate, for each of the specifications considered, the settings of the tuning parameters L_r , L_p , K_z , I_z , K_x , I_x , the number of parameters p_θ that they imply, the value of the minimized objective function $s_n(\hat{\theta})$, Schwarz's criterion, and the battery of specification tests (discussed below) are reported in Table 2 for the bivariate price and volume series $y_t = (\Delta p_t, v_t)$. All reported values are comparable as the same number of leading observations (27) were set aside to provide the initial lags in every fit. The net sample size is 16,100

observations.

Of the models in Table 2, the Schwarz preferred model has $L_r = 16$, $L_p = 4$, $K_z = 4$, $I_z = 1$, $K_x = 2$, $I_x = 1$ with $p_\theta = 368$ at a saturation ratio of $(2 \cdot 16100)/368 = 87.5$ observations per parameter.

As mentioned above, to guard against the conservative nature of the Schwarz criterion, specification tests are conducted for each fit as follows.

First, we get scaled residuals by computing analytically the moments of the estimated conditional density and using them to compute the estimated conditional mean $\hat{E}(y|x_{t-1})$ and variance $\hat{V}ar(y|x_{t-1})$ at each $x_{t-1} = (y_{t-L}, \dots, y_{t-1})$ in the sample. Using these, a scaled residual is computed as $\hat{u}_t = [\hat{V}ar(y|x_{t-1})]^{-1/2}[y_t - \hat{E}(y|x_{t-1})]$ where $[\hat{V}ar(y|x_{t-1})]^{-1/2}$ denotes the inverse of the Cholesky factor of the conditional variance.

Next, we conduct diagnostic tests for predictability in both the scaled residuals and the squares of the scaled residuals. Predictability of the scaled residuals would suggest inadequacies in the conditional mean estimate implied by the fitted density, and thus such tests are termed mean tests. Similarly, predictability of the squared scaled residuals would suggest inadequacies in the implied estimate of the conditional variance, and thus such tests are termed variance tests. For both mean and variance, we conduct two types of tests for predictability, one of which is sensitive to short-term misspecification while the other is sensitive to long-term misspecification.

For the conditional mean, the short-term diagnostic test is a test for the significance of a regression of scaled residuals \hat{u}_t on linear, quadratic, and cubic terms of lagged values of y_t . The long-term test is a test for the significance of a regression of scaled residuals on annual dummies to check for a failure to capture long-term trends. For the conditional variance, the tests are the same with the squares of the scaled residuals as the dependent variable in these regressions. The significance test is the F-test when the residuals are from a univariate series and is the Wilk's test when the residuals are from a multivariate series such as the bivariate price and volume series that we are using for illustration. It should be noted that because of the "Durbin effects" of pre-fitting discussed in Newey (1985) and Tauchen (1985), the p-values could be somewhat inaccurate, even asymptotically.

As seen from Table 2, the Schwarz preferred model does reasonable well with respect to

the short term diagnostics but apparently fails to capture long term heterogeneity of some sort. In the next section we will use simulation to try to determine what features of the series the model doesn't approximate well. However, moving the truncation point beyond the Schwarz preferred model does not appear to help much, so we will stick with the Schwarz preferred model. This outcome is somewhat unusual in our experience. Usually, increasing the tuning parameters one or two notches beyond the Schwarz preferred model will reduce the diagnostic to insignificance.

3 Simulation

In this section, we describe the algorithms used to simulate from an SNP density. This material, and a few of the applications in the next section, are the original contributions of the paper.

3.1 Rejection Methods

The rejection method for sampling from a (multivariate) density $h(z)$ depends on finding a positive, integrable function $b(z)$ that dominates $h(z)$, viz.

$$0 \leq h(z) \leq b(z)$$

The dominating function $b(z)$ is called an upper envelope for $h(z)$ or majorizing function. Derive a density $g(v)$ from $b(z)$ by putting

$$g(v) = b(v) / \int b(s) ds$$

Using $b(z)$ and $g(v)$, a sample from $h(z)$ is generated as follows.

Generate the pair (u, v) by generating v from $g(v)$ and u from the uniform distribution on $[0, 1]$. If

$$u > h(v)/b(v)$$

reject the pair (u, v) and try again. If

$$u \leq h(v)/b(v)$$

accept $z = v$ as a sample from $h(z)$.

The method works because

$$\begin{aligned}
P(z \leq t) &= P[\{(u, v) : v \leq t\} | \{(u, v) : u \leq h(v)/b(v)\}] \\
&= \frac{\int_{v \leq t} \int_0^{h(v)/b(v)} g(v) \, du \, dv}{\int_{v \leq \infty} \int_0^{h(v)/b(v)} g(v) \, du \, dv} \\
&= \frac{\int_{v \leq t} h(v) / [\int b(s) \, ds] \, dv}{\int_{v \leq \infty} h(v) / [\int b(s) \, ds] \, dv} \\
&= \int_{v \leq t} h(v) \, dv
\end{aligned}$$

Above z , s , v , and t are vectors and inequalities such as $z \leq t$ are interpreted to mean inequality element by element. Also, it probably goes without saying but we mention it to be safe, the draws (u, v) are independent of each other and u is drawn independently of v within each pair.

For this approach to work well, $h(v)/b(v)$ must be near one over regions where $g(v)$ assigns high probability and $g(v)$ must be easy to sample. The $g(v)$ that we construct in the next section can be sampled using standard algorithms and the upper envelope $b(z)$ from which it is derived achieves a hit rate of about 50%.

See Kennedy and Gentle (1980, Section 6.4.3) for a additional details on rejection methods and a literature review.

3.2 Upper Envelope

In this subsection, we drive an upper envelope for the conditional density of a conditionally heterogeneous innovation.

This density is written as

$$\begin{aligned}
h(z|x) &= [P(z, x)]^2 \phi(z) / c(x) \\
c(x) &= \int [P(s, x)]^2 \phi(s) \, ds
\end{aligned}$$

The polynomial $P(z, x)$ that appears in the conditional density is the polynomial in z

$$\sum_{|\alpha|=0}^{K_z} a_\alpha(x) z^\alpha$$

with coefficients depending on a polynomial in x

$$a_\alpha(x) = \sum_{|\beta|=0}^{K_x} a_{\alpha\beta} x^\beta$$

where

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)' \quad \beta = (\beta_1, \beta_2, \dots, \beta_{ML})'$$

are multi-indices (vectors with integer elements), and

$$|\alpha| = \sum_{i=1}^M |\alpha_i| \quad |\beta| = \sum_{i=1}^{ML} |\beta_i|$$

$$y^\alpha = \prod_{i=1}^M y_i^{\alpha_i} \quad x^\beta = \prod_{i=1}^{ML} x_i^{\beta_i}$$

Some of the $a_{\alpha\beta}$ will have been put to zero to reflect the filtering out of interactions when I_z or I_x are positive.

Since

$$\sum_{|\alpha|=0}^{K_z} a_\alpha(x) z^\alpha \leq \sum_{|\alpha|=0}^{K_z} |a_\alpha(x)| |z|^\alpha$$

where $|z|$ denotes the vector z with each element replaced by its absolute value,

$$b(z|x) = \left[\sum_{|\alpha|=0}^{K_z} |a_\alpha(x)| / \sqrt{c(x)} |z|^\alpha \right]^2 \phi(z)$$

is an upper envelope for $h(z|x)$.

3.3 Simulating an Innovation

In this subsection we develop an algorithm to simulate from $g(v|x)$ which is embedded in an algorithm to sample $h(z|x)$ by rejection.

To obtain the density $g(v|x)$, note that $b(z|x)$ is weighted sum of Chi density functions

$$\chi(s, \nu) = \frac{2^{1-(\nu/2)}}{\Gamma(\nu/2)} s^{\nu-1} e^{-(s^2/2)} \quad s > 0$$

by writing $b(z|x)$ as

$$b(z|x) = \sum_{|\alpha|=0}^{K_z} \sum_{|\gamma|=0}^{K_z} \frac{|a_\alpha(x) a_\gamma(x)|}{c(x)} \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2] \prod_{i=1}^M \chi(|z_i|, \alpha_i + \gamma_i + 1)$$

Having observed that $b(z|x)$ is a weighted sum of Chi density functions, an algorithm to sample from $g(v|x)$ can be developed without having to derive an explicit expression for $g(v|x)$ as follows.

Normalize the weights to sum to one, viz.

$$w_{\alpha\gamma} = \frac{|a_\alpha(x)a_\gamma(x)| \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2]}{\sum_{|\alpha|=0}^{K_z} \sum_{|\gamma|=0}^{K_z} \{|a_\alpha(x)a_\gamma(x)| \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2]\}}$$

These weights define a distribution $F(\alpha, \gamma)$ over $\{(\alpha, \gamma) : 0 \leq |\alpha| \leq K_z, 0 \leq |\gamma| \leq K_z\}$. To have a convenient way to traverse this set, let the elements of $\{(\alpha, \gamma) : 0 \leq |\alpha| \leq K_z, 0 \leq |\gamma| \leq K_z\}$ be ordered in some arbitrary way so that they can be indexed by the sequence $(\alpha, \gamma)_\tau$ where $\tau = 1, \dots, N$. In forming this sequence, one may wish to delete those (α, γ) for which $w_{\alpha\gamma} = 0$ for all x to keep N small when $I_z > 0$. Put $w_\tau = w_{\alpha\gamma}$ where $(\alpha, \gamma) = (\alpha, \gamma)_\tau$.

The algorithm for sampling from $h(z|x)$, and $g(v|x)$ as an intermediate step, is as follows:

1. Sample (α, γ) from the distribution $F(\alpha, \gamma)$. This is easiest to do by generating u from the uniform distribution on $[0, 1]$, finding the largest T such that $\sum_{\tau=1}^T w_\tau \leq u$, and taking $(\alpha, \gamma) = (\alpha, \gamma)_T$ as a sample from $F(\alpha, \gamma)$.
2. Generate the sequence s_1, \dots, s_M by independently sampling from $\chi(s_1, \alpha_1 + \gamma_1 + 1), \dots, \chi(s_M, \alpha_M + \gamma_M + 1)$.
3. Randomly change the sign of s_i with probability $1/2$ and assign the result to v_i for $i = 1, \dots, M$. That is, draw u from the uniform distribution on $[0, 1]$ and put $v_i = s_i$ if $u > 1/2$ and $v_i = -s_i$ if $u \leq 1/2$. Put $v = (v_1, v_2, \dots, v_M)'$; v is a sample from $g(v|x)$.
4. Generate u from the uniform distribution on $[0, 1]$. If $u > h(v|x)/b(v|x)$ reject the pair (u, v) and return to 1. If $u \leq h(v|x)/b(v|x)$ accept $z = v$ as a sample from $h(z|x)$.

We use the algorithm by Schrage (1979) to sample from the uniform and the algorithm by Monahan (1987) to sample from the Chi.

3.4 Simulating Data

Given a sample z from $h(z|x)$, a sample from $f(y|x, \theta)$ is obtained from the location-scale shift

$$\begin{aligned} y &= R_x z + \mu_x \\ \mu_x &= b_0 + Bx \\ \text{vech}(R_x) &= p_0 + P|x| \end{aligned}$$

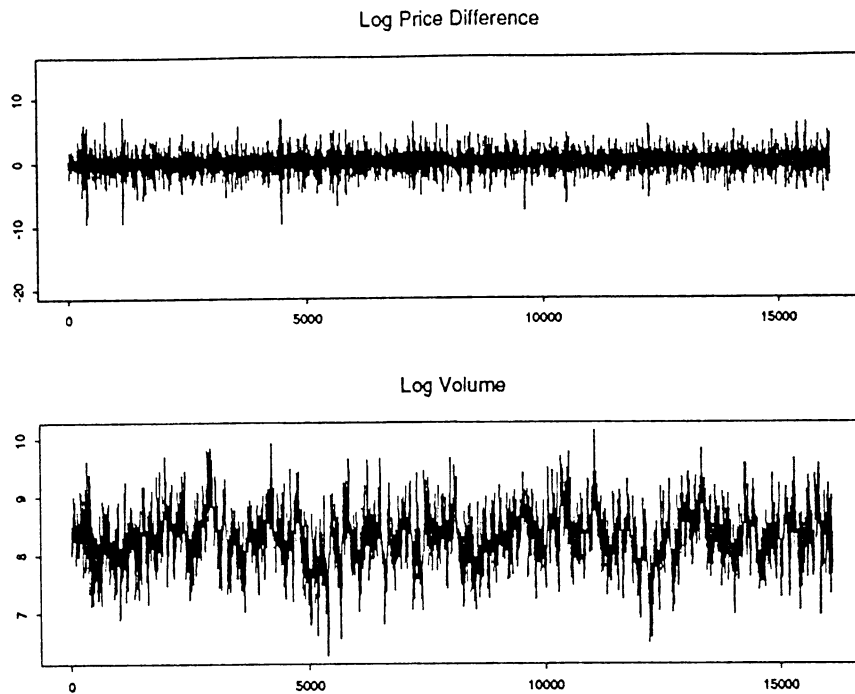


Figure 2: Daily NYSE S&P, Simulated Data

Recall that if the logistic transformation is employed, \hat{x} replaces x in these formulas.

Figure 2 is a simulated sample path from the model fitted in Section 2 of the same length $n = 16,127$ as the data, started using the first 27 lags of the data.

Comparing with Figure 1, the fitted model captures the qualitative features of the data quite well. Most interestingly, the simulation contains several market crashes, as does the data. However, it does appear that the fitted model has some trouble capturing the duration of spurts of volatility. But, if the time axis is magnified in Figures 1 and 2, this discrepancy seems to disappear. It is clear from the plots and the tests reported in Table 2 that there is some aspect of sample paths having to do with duration of episodic events that the model does not capture but it is hard to characterize it exactly.

4 Conditional Variance Function

In this section, we shall focus on estimation of a particular functional of the conditional density, the conditional variance function

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

where $f_\infty(y|x)$ denotes the true, one-step-ahead, conditional density for the process $\{y_t\}$. It is estimated by

$$\begin{aligned}\hat{V}ar(y|x) &= \int [y - \mathcal{E}(y|x)][y - \mathcal{E}(y|x)]' f_K(y|x, \hat{\theta}) dy \\ \hat{\mathcal{E}}(y|x) &= \int y f_K(y|x, \hat{\theta}) dy\end{aligned}$$

where $f_K(y|x, \hat{\theta})$ is the estimate using the truncated expansion of $f_\infty(y|x)$ described in Section 2.

More specifically, we are interested in the symmetry of the leverage function $L(\Delta p_{t-1})$ which is the conditional variance function above with every entry y_{t-j} in $x_{t-1} = (y_{t-L}, \dots, y_{t-1})$ put to its sample mean

$$\bar{y} = \frac{1}{n} \sum_{t=1}^{16,127} y_t$$

except y_{t-1} which is put to $(\Delta p_{t-1}, \bar{y}_2)$ in the case of the bivariate price-volume series and to Δp_{t-1} when considering the univariate price series obtained by deleting the volume component of the bivariate series.

The leverage function has been actively studied, see Nelson (1989a, 1989b), Schwert(1989), and Pagan and Schwert (1989) and their references. Of particular interest is the fact that estimates show larger conditional variance following down-ticks than up-ticks. That is, the estimated models predict more market volatility following a crash than a sharp rise of the same magnitude. This fact is of some interest to agencies charged with regulation of financial markets. Also of interest is the fact that the asymmetry is markedly reduced when other variables such as volume, as in this paper, or interest rates, as in Glosten, Jagannathan, and Runkle, (1989), are added to the conditioning set.

Estimates of the leverage function for large values of Δp_{t-1} are not very precise in the sense that estimates are quite different depending on estimation method and model. One can see why from Figure 3. In the upper panel of Figure 3 is a scatter plot of Δp_t^2 against Δp_{t-1} for all $n=16,127$ points (less one for the lag). In the lower panel is a scatter plot of Δp_t^2 against Δp_{t-1} after points where $|\Delta p_{t-j} - \bar{y}_j| > (1.5)StdDev$ for $j = 2, 3$, or 4 have been deleted; there are 5749 points in the lower panel. Basically, estimating a leverage function is equivalent to trying to put a (nonparametric) regression line through the scatter plots shown in Figure 3. These fits will be heavily influenced by the few outlying points at the extreme

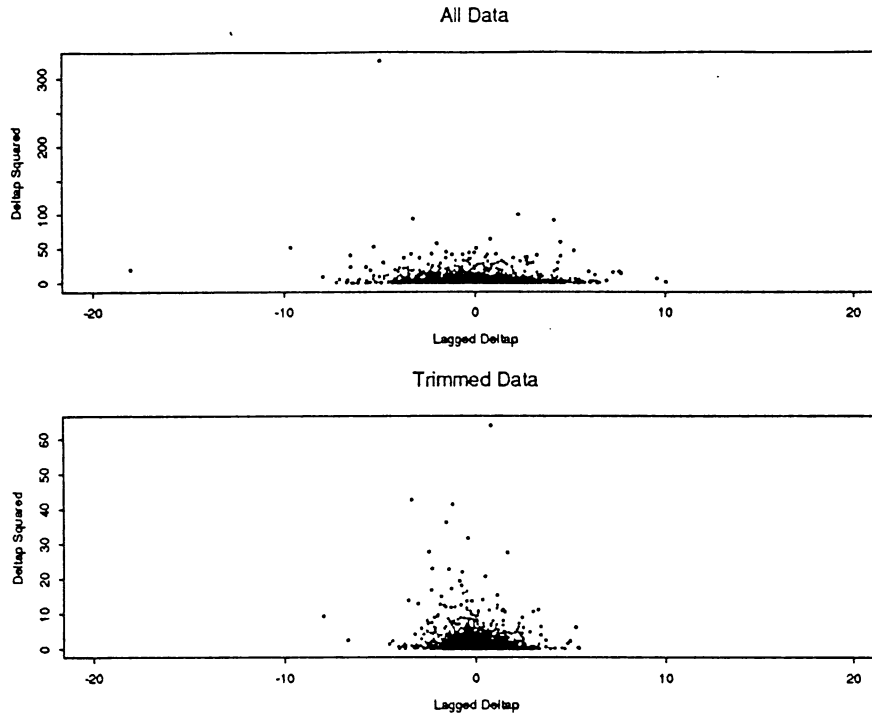


Figure 3: Scatter Plot of Squared Price on Lagged Price

left of the plots. (The point at about $(-5, 320)$ in the upper panel is Friday, October 16, 1987, and the point at about $(-15, 20)$ is Monday, October 19, 1987.)

One can see the effect that this has on estimates in Figures 4, 5, and 6. The kernel estimate was fitted by methods suggested for time series analysis by Robinson (1983) with the bandwidth parameter selected by visual inspection of plots such as Figure 4. The univariate SNP fit was fitted by the methods described in Section 2 resulting in a model with $L_r = 16$, $L_p = 6$, $K_z = 4$, $I_z = 0$, $K_x = 1$, $I_x = 0$ with $p_\theta = 34$ at a saturation ratio of 277.6 observations per parameter. The leverage function was computed analytically from the SNP estimate using the standard recursions for the moments of the normal distribution; a moment of an SNP density is just some higher moment of a normal. Figure 6 is the same SNP specification robustified using the $x_{t-1} \rightarrow \hat{x}_{t-1}$ logistic transformation.

The influence of extreme points on the fit decreases progressively in Figures 4 through 6. Kernel estimates are most sensitive. As mentioned above, the leading term of an SNP expansion tends to fill in where data is sparse and give more realistic estimates over sparse regions than kernel estimates. The logistic transformation has the intended robustifying effect and gives estimates that accord better with a visual fit to the scatter plots in Figure 3.

Figures 7 and 8 are kernel and SNP estimates of the leverage function using bivariate data instead of univariate data. Again, the bandwidth parameter for the kernel estimate was

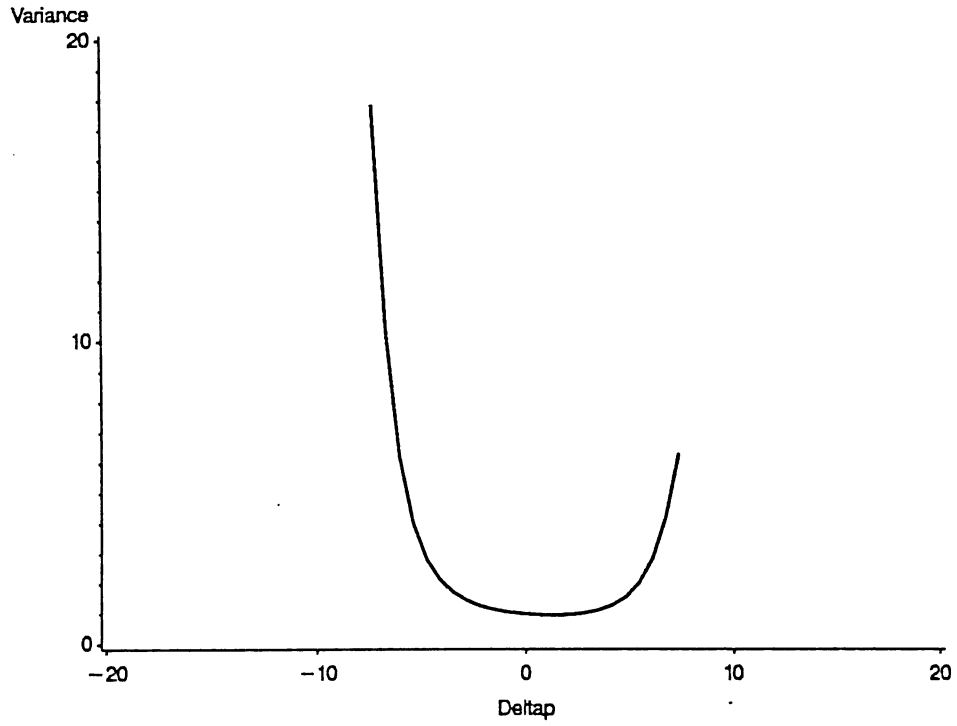


Figure 4: Kernel Estimate of the Leverage Function: Univariate S&P Data

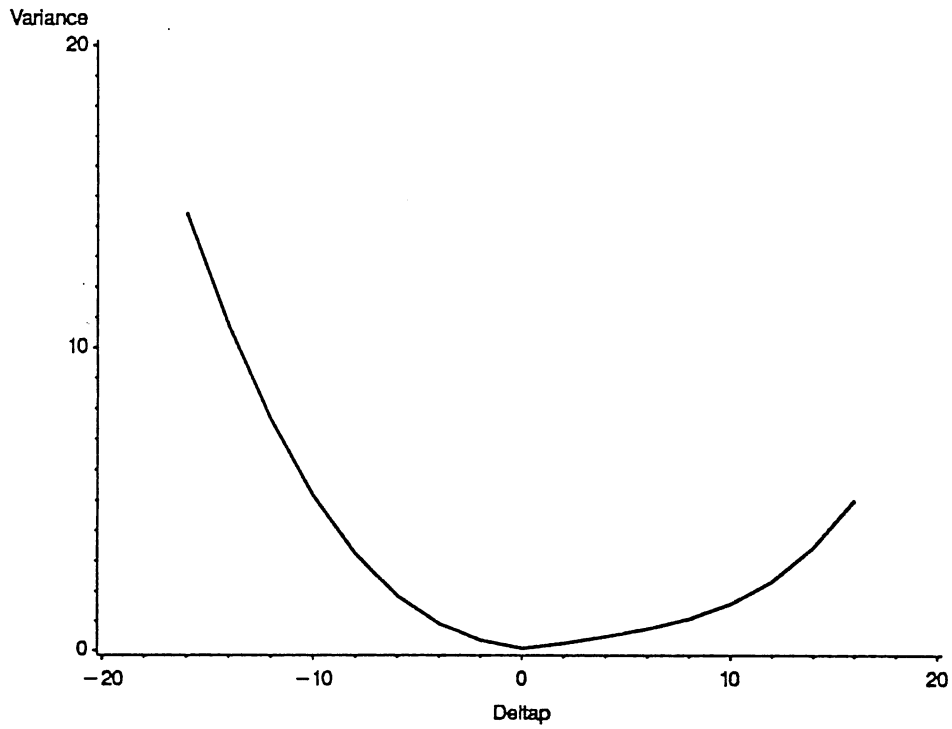


Figure 5: SNP Estimate of the Leverage Function: Univariate S&P Data

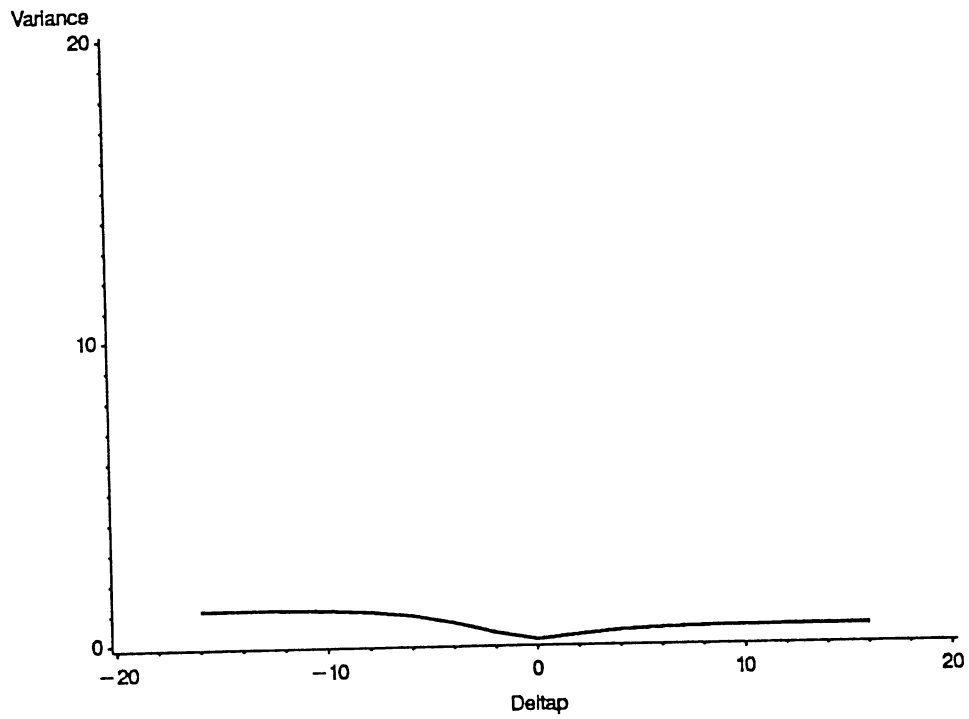


Figure 6: Robust SNP Estimate of the Leverage Function: Univariate S&P Data

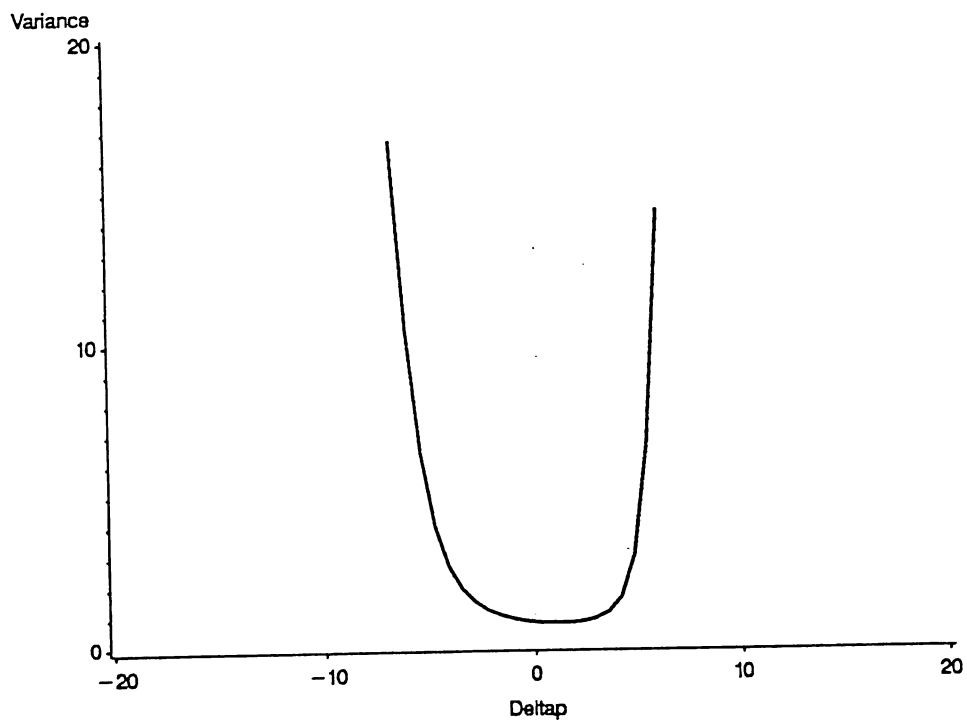


Figure 7: Kernel Estimate of the Leverage Function: Bivariate S&P Data

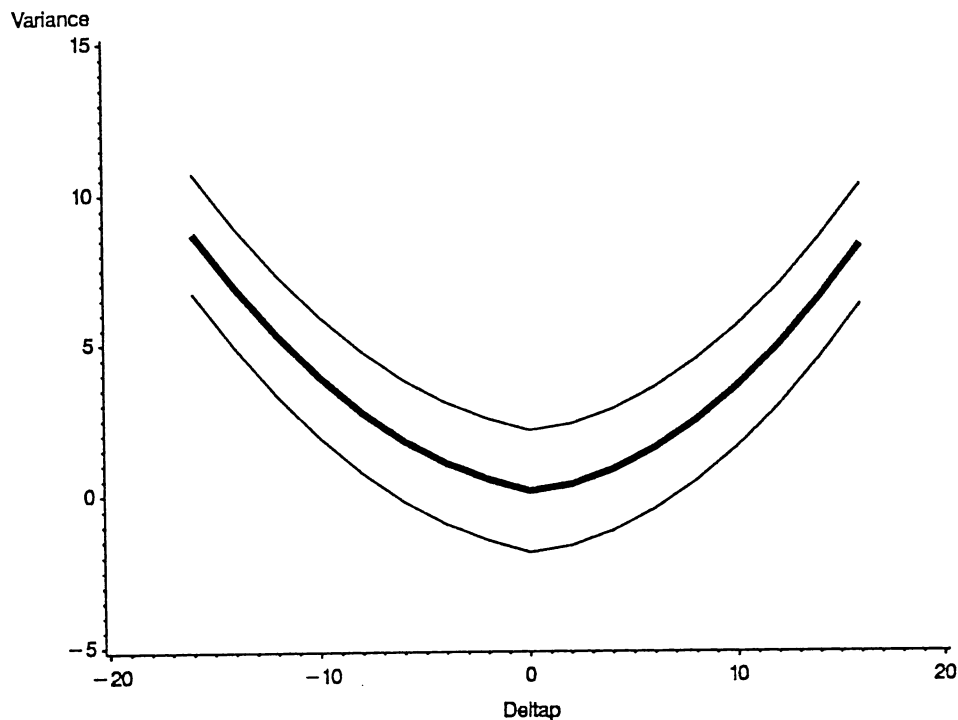


Figure 8: SNP Estimate of the Leverage Function: Bivariate S&P Data

selected visually. The bivariate SNP fit is that described in Section 2; sup norm confidence bands are superimposed on the estimate using methods discussed below.

The most striking feature of Figures 7 and 8 is the substantial lessening of the asymmetry by comparison with Figures 4, 5, and 6. Our preference is to ascribe the asymmetry to influential points in the data whose impact is lessened when another explanatory variable is entered in the model and to view seeking an explanation by appeal to economic theory as unwarranted. Supporting this view are both the fact that Glosten, Jagannathan, and Runkle (1989) observe the same effect when interest rates replace volume and also the instabilities in Figures 4, 5, and 6.

The sup norm confidence bands were constructed from the bivariate SNP fit by a parametric bootstrap: Using the methods discussed in Section 3, 250 independent sample paths similar to Figure 2 were generated. The bivariate SNP specification $L_r = 16$, $L_p = 4$, $K_z = 4$, $I_z = 1$, $K_x = 2$, $I_x = 1$ was fit to each; that is, there was no specification search. A leverage function was computed for each fit. The bands shown in Figure 8 are just wide enough to contain 95% of them.

In this connection, it is of interest to check that SNP methods confirm Glosten, Jagannathan, and Runkle's findings. We used a real, monthly, bivariate stock and bond returns

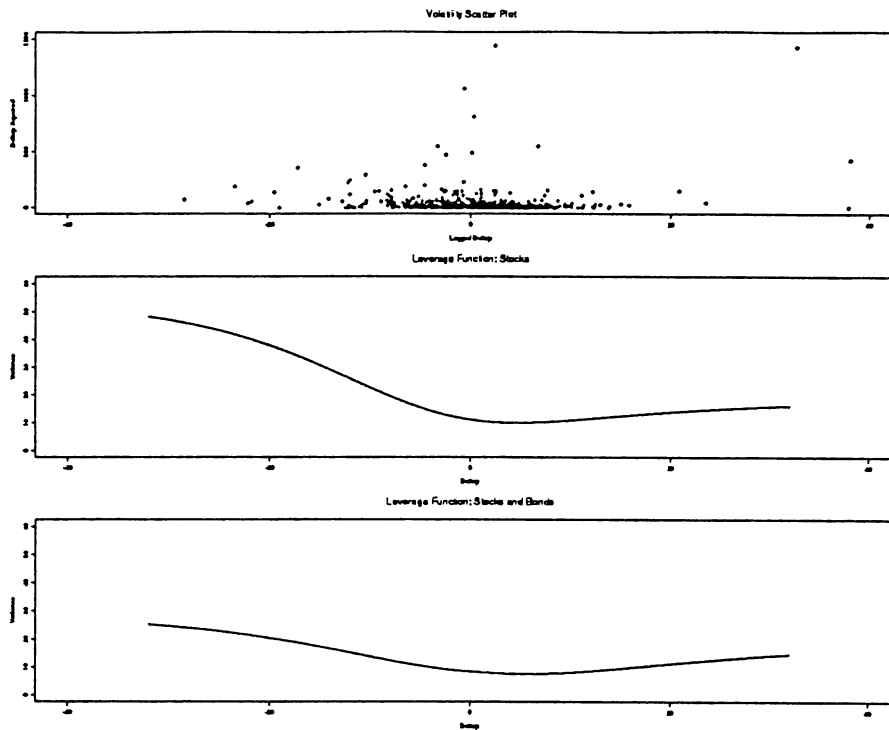


Figure 9: SNP Estimate of the Leverage Function: Stock and Bond Returns

series taken from Gallant, Hansen, and Tauchen (1990). The results are reported in Figure 9. The upper panel is a scatter plot of Δp_t^2 against Δp_{t-1} for $n = 737$ points; the second panel is the leverage function computed from analytical moments from an SNP fit to price returns alone; and the third panel is the leverage function for the bivariate price and bond returns fit. Our computations confirm their results: the leverage function conditional on both past stock and bond returns is more symmetric than the leverage function conditional on past stock returns alone.

5 References

- Akaike, H. (1969), "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics* 21, 243-247.
- Andrews, Donald W. K. (1989), "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Models," Cowles Foundation Discussion Paper No. 874R, Yale University.
- Brock, W. A., Dechert, W. D., and J. A. Scheinkman (1987), "A Test for Independence Based on the Correlation Dimension," Department of Economics, University

of Chicago, working paper.

Brunner, Allan D. (1989), *An Empirical Investigation of U. S. Business Cycles Using Non-linear Models*, Ph.D. Dissertation, Duke University.

Davidian, M. and R. J. Carroll (1987), "Variance Function Estimation," *Journal of the American Statistical Association* 82, 1079-1091.

Eastwood, Brian J., and A. Ronald Gallant (1987), "Adaptive Truncation Rules for Semi-nonparametric Estimators That Achieve Asymptotic Normality," H. B. Alexander Foundation Paper 87-53, Graduate School of Business, University of Chicago.

Elbadawi, Ibrahim, A. Ronald Gallant, and Geraldo Souza (1983), "An Elasticity Can be Estimated Consistently Without A Priori Knowledge of Functional Form," *Econometrica* 51, 1731-1752.

Gallant, A. Ronald, and Douglas W. Nychka (1987) "Seminonparametric Maximum Likelihood Estimation" *Econometrica* 55, 363-390.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, 987-1007.

Gallant, A. Ronald, Lars Peter Hansen, and George Tauchen (1990), "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution," *Journal of Econometrics*, forthcoming.

Gallant, A. Ronald, David A. Hsieh, and George E. Tauchen (1989), "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-83," in William A. Barnett, James Powell, George E. Tauchen, eds. *Nonparametric and Semiparametric Methods in Econometrics and Statistics, Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, Cambridge: Cambridge University Press, forthcoming.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen (1990), "Stock Prices and Volume," Graduate School of Business, University of Chicago, working paper.

- Gallant, A. Ronald, and Douglas W. Nychka (1987) "Seminonparametric Maximum Likelihood Estimation" *Econometrica* 55, 363-390.
- Gallant, A. Ronald, and Geraldo Souza (1989), "On the Asymptotic Normality of Fourier Flexible Form Estimates," Institute of Statistics Mimeograph Series No. 1957, North Carolina State University.
- Gallant, A. Ronald, and George Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica* 57, 1091-1120.
- Gill, Phillip E., Walter Murray, Michael A. Saunders, and Margaret H. White (1983), "User's Guide for SOL/NPSOL: A Fortran Package for Nonlinear Programming," Technical Report Number SOL 83-12, Systems Optimization Laboratory, Department of Operations Research, Stanford University.
- Glosten, Lawrence R., Ravi Jagannathan, and David Runkle (1989), "Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," Banking Research Center Working Paper No. 166, Kellogg Graduate School of Management, Northwestern University.
- Hussey, Robert M. (1989a), *Seminonparametric Modeling of Aggregate Employment Dynamics*, Ph.D. Dissertation, Duke University.
- Hussey, Robert M. (1989b), "Nonparametric Evidence on Asymmetry in Business Cycles Using Aggregate Employment Time Series," The Jesuits at Wernersville, Pennsylvania 19565, working paper.
- Hussey, Robert M. (1989c), "Asymmetric Adjustment Costs in Labor Demand Models," The Jesuits at Wernersville, Pennsylvania 19565, working paper.
- Kennedy, William J. Jr., and James E. Gentle (1980), New York, Marcel Dekker.
- Monahan, John F. (1987) "An Algorithm for Generating Chi Random Variables, Algorithm 651" *ACM Transactions on Mathematical Software* 13, 168-172. (1988) Corrigendum, *ACM Transactions on Mathematical Software* 14, 111.

- Nelson, D. (1989a), "Conditional Heteroskedasticity in Asset Returns: A New Approach," Graduate School of Business, University of Chicago, working paper.
- Nelson, D. (1989b), "Modelling Stock Market Volatility Changes," Graduate School of Business, University of Chicago, working paper.
- Newey, W. K., 1985, Maximum Likelihood Specification Testing and Conditional Moment Tests, *Econometrica* 53, 1047-1071.
- Pagan, Adrian and G. William Schwert (1989) "Alternative Models for Conditional Stock Volatility," Department of Economics, University of Rochester, working paper.
- Potscher, B. M. (1989), "Model selection under nonstationarity: Autoregressive models and stochastic linear regression models," *The Annals of Statistics* 17, 1257-1274.
- Robinson, P. M. (1983), "Nonparametric Estimators for Time Series," *Journal of Time Series Analysis* 4, 185-207.
- Schwert, G. William (1989), "Why Does Stock Market Volatility Change Over Time?," *The Journal of Finance* XLIV, 1115-1153.
- Schwarz, G., 1978, "Estimating the Dimension of a Model," *Annals of Statistics* 6, 461-464.
- Schrage, L. (1979), "A More Portable FORTRAN Random Number Generator," *ACM Transactions on Mathematical Software* 5, 2.
- Tauchen, George (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models," *Journal of Econometrics* 30, 415-443.

6 Appendix: Fitting SNP Models

Fortran code that implements the SNP methodology is available from the authors. Details on installing and using the code are available in Subsection 6.2 below.

In addition to facilitating model estimation, the code makes it easy to retrieve residuals, predicted conditional means, and predicted conditional variances. These statistics useful for diagnostic testing, model evaluation, forecasting, and related purposes. In addition, the

code provides the ordinates of the SNP conditional density over a rectangular grid of points, which is useful for plotting purposes and for performing numerical integration against the SNP conditional density. Finally, it can generate Monte Carlo simulated realizations of arbitrary length from the SNP density, a capability with a variety of applications.

6.1 Fitting Strategy

As discussed in Section 2, the model selection strategy entails moving upward along an expansion path. The fitted SNP models becomes more richly parameterized at each level along the path. The expansion tentatively stops when the best model under the Schwarz criterion is obtained. The Schwarz-preferred model is then subjected to a battery of specification tests on the conditional first and second moments. Often, but not always, some further expansion of the model is needed in order to achieve satisfactory performance on the diagnostics.

Experience suggests that care is needed in fitting the SNP model at any particular level along the expansion path. Estimates at one level provide start values for the next, so the user should be cautious of hitting a local optimum at any level. Among other things, a false optimum could adversely affect computations at all subsequent levels in the path.

The software is thus designed to facilitate the process of checking for a local optimum, so that the user can have reasonable confidence in the computations before proceeding to the next level. On a single run, the program is capable of performing a wave of optimizations with start values and perturbation factors read in from various files. In typical practice, between one and three waves of runs, with five to ten optimizations within each wave, might be performed to compute the SNP model at a particular level before proceeding to the next. In making the decision as to whether to accept the computations at one level and proceed to the next, the user should look for convergence to the same overall optimum from several start values. This agreement can sometimes be difficult to obtain for exceedingly large models, and near the end of the expansion path the user might simply have to accept the best computed objective function value out of a wave of fits. In numerical experiments, we have found that, near the end of the path, this probably does little harm as the various optima differ only in their implications for extreme tail behavior.

Table 2 provides an example of a computed expansion path. Each level, i.e. row in

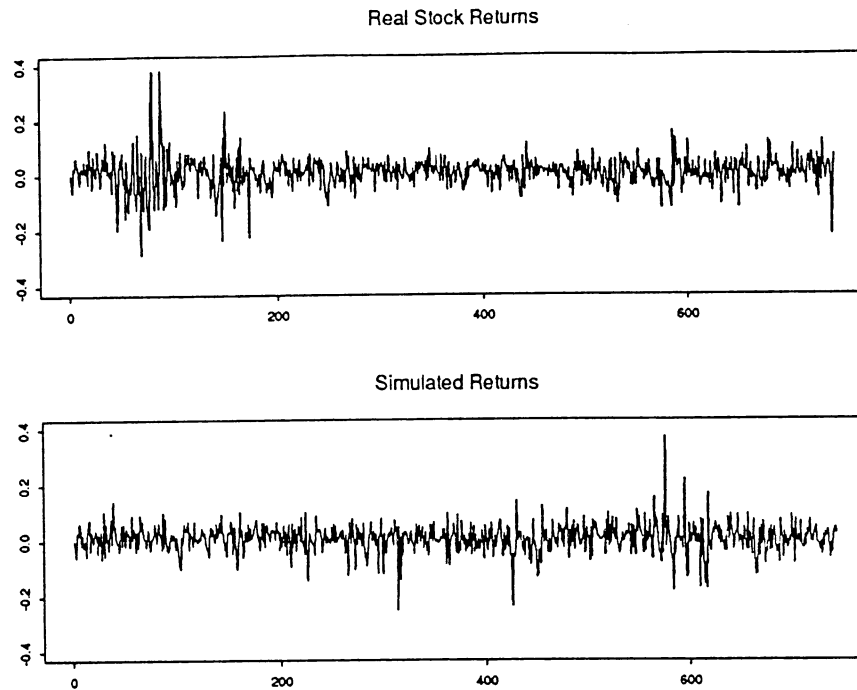


Figure 10: Monthly NYSE Returns, 1926-1987, Not Adjusted for Calendar Effects

the table, was computed using the software in the above-described manner. In the next subsection we walk the reader through the steps of computing a similar table.

6.2 Using the Program

Fortran code is available via ftp anonymous at [ccvrl.cc.ncsu.edu](ftp://ccvrl.cc.ncsu.edu) (128.109.212.20) in directory `pub/arg/snp`; alternatively, it is available from the authors in the form of a 5-1/4 inch, 2S/2D, DOS formatted diskette. The code is provided at no charge for research purposes without warranty. It has run on Sun workstations, IBM 3090 MVS mainframes, 386 DOS PC's, and DEC VAX minicomputers. It relies on a fairly widely distributed optimization routine by Gill, Murray, Saunders, and Wright (1983) which is available as NPSOL from the Office of Technology Licensing, Stanford University, 350 Cambridge Avenue, Suite 250, Palo Alto, CA 94306 and as E04UCF in the Nag Library which is available from Numerical Algorithms Group, 1400 Opus Place, Suite 200, Downers Grove, IL 60515-5702. The SNP code is written so that it is easy to substitute another optimization routine. However, we have experimented with several alternatives but found nothing that works as well as NPSOL. The SNP code is documented in comments to the main. The purpose of this appendix is to walk the user through an application under the assumption that ambiguities can be resolved by reference to the documentation. The data that we use for illustration is a monthly, real,

value weighted stock series for 1926 through 1987 taken from Gallant, Hansen, and Tauchen (1990). The data are plotted in Figure 10. Data, code, and output for this application are on the distribution diskette.

Program control is through a subroutine written by the user to read the data, getdat, and three text files. The names assigned to these files are determined by the subroutines in the file snpopen.f on the distribution diskette, which will probably have to be edited to reflect the idiosyncrasies of the Fortran compiler and operating system under which it will be compiled and run. An example of snpopen.f, as edited for SunOS, that will define the file names in the subsequent discussion is:

snpopen.f

```

SUBROUTINE OPENFL
OPEN(UNIT=3 ,FILE='detail.dat',status='old',form='formatted')
OPEN(UNIT=11,FILE='pointer.dat',status='old',form='formatted')
OPEN(UNIT=12,FILE='control.dat',status='old',form='formatted')
OPEN(UNIT=14,FILE='stocks.dat',status='old',form='formatted')
OPEN(UNIT=27,FILE='scratch.dat',status='old',form='formatted')
RETURN
END
SUBROUTINE OPEN89(FILE8,FILE9)
CHARACTER*13 FILE8,FILE9
OPEN(UNIT=10,FILE='summary.dat',status='old',form='formatted')
OPEN(UNIT=8,FILE=FILE8,status='old',form='formatted')
OPEN(UNIT=9,FILE=FILE9,status='old',form='formatted')
END

```

The purpose of getdat is straightforward. It fills an M by N matrix with the data $\{\tilde{y}_t\}_{t=1}^n$ where M is the dimension of \tilde{y}_t and N is the number of observations. An example says it all.

getdat.f

```

SUBROUTINE GETDAT(DATA,M,N)
IMPLICIT REAL*8 (A-H,O-Z)
REAL*8 DATA(M,N)
DO 10 I=1,N
READ(14,14001) DATA(1,I)
10 CONTINUE
RETURN
14001 FORMAT(1X,D18.12)
END

```

The file control.dat assigns file names to FILE8 and FILE9 and sets the perturbation to be applied to the incremental coefficients in the polynomial part of the Hermite expansion on

a restart from a lower order polynomial. The format is FORMAT(2A13,F10.0). For our example, we shall use the following file names.

control.dat

s300000.in0 s300000.out 0.0

The file pointer.dat gives the number of lines of control.dat that are to be skipped. The format is FORMAT(I4):

pointer.dat

0

The parameter file s300000.in0 sets all program parameters and options. The easiest way to deal with a parameter file is to start by describing a fit that is to commence with a VAR start because the amount of information required for a VAR start is minimal. A new parameter file will be written to s300000.out. This new file can then be edited to produce additional fits or choose various options. The file s300000.in0 is as follows.

s300000.in0

Stock		FORMAT(A5,66X,F3.1)	7.5
1	M	FORMAT(I5)	
744	N	FORMAT(I5)	
6	IDROP	FORMAT(I5)	
3	0 LRX, LPOLY	FORMAT(I5,1X,I3)	
0	0 IDEGY, LINTY	FORMAT(I5,1X,I3)	
0	0 IDEGX, LINTX	FORMAT(I5,1X,I3)	
0	ISTART	FORMAT(I5)	
400	ITMAX	FORMAT(I5)	
1.0D-05	TOLER	FORMAT(D9.1)	
0	0 NLOG, IDIAG	FORMAT(I5,1X,I3)	
3.0	SFAC	FORMAT(D9.1)	
50	NGRD	FORMAT(I5)	
0	IUNIT9	FORMAT(I5)	

In the first line, Stock is a user chosen label and the 7.5 in column 72 is obligatory. M is the dimension of y_t , N is the number of observations, IDROP is the number of observations at the beginning of the series to skip to produce initial lags; this should be larger than any value of L_r or L_p envisaged in any fit. LRX is L_r , LPOLY is L_p , IDEGY is K_z , LINTY is I_z , IDEGX is K_x , LINTX is I_x , setting ISTART to 0 means a VAR start, ITMAX is an iteration limit that is passed to NPSOL, and TOLER is a convergence tolerance that is passed to NPSOL. Setting NLOG to 1 means suppress the logistic transform, setting IDIAG

to 1 puts the off diagonal elements of S to zero prior to the transformation $\tilde{y}_t \rightarrow y_t$. SFAC determines the plotting increment; 3.0 is usually about right. NGRD is either the number of plot points for a graphic or the seed for a simulation.

IUNIT9 states what is to be written to unit 9 after estimation. If IUNIT9=0, a new parameter file is written to unit 9 that can be used to move from a sparse parameterization to a richer one or to summarize an estimation for subsequent simulation, plotting, or moment computations. The other choices are: IUNIT9=1, the residuals used for diagnostics are written to UNIT 9; IUNIT9=2, the mean of the one-step-ahead conditional density is computed at each x_{t-1} in the sample and written to unit 9; IUNIT9=3, the upper triangle of the variance-covariance matrix of the one-step-ahead conditional density is computed at each x_{t-1} in the sample and written to unit 9; IUNIT9=4, plot data is written to unit 9; and IUNIT9=5, a simulation is written to UNIT 9. To suppress a re-estimation when IUNIT9=2, 3, 4, or 5 put ITMAX to 0.

Upon running the program one gets:

pointer.dat

1

summary.dat

TWK= 0.0000 ITR= 7/ 400 INF= 0 VAR Stock300000 8 1.31286 s300000.out

s300000.out

```

Stock      s300000.out   from s300000.in0   with TWEAK = 0.0000 by SNP 7.5
 1         M
744       N
 6        IDROP
 3 0 LRX,   LPOLY
 0 0 IDEGY, LINTY
 0 0 IDEGX, LINTX
 1        ISTART (0 VAR START, 1 PARMFILE START)
400       ITMAX
1.0D-05   TOLER (D9.1)
 0 0 NLOG, IDIAG
 3.0 SFAC (F9.1)
50        NGRD
 0        IUNIT9 (0 PARMFILE, 1 RES, 2 MEAN, 3 VAR, 4 PLOT, 5 SIM)
 3 LRXO    0      LPOLYO  1      M
 0 IDEGYO  1      KYO     0      LINTYO
 0 IDEGXO  1      KXO     0      LINTXO
0.1000000000000000D+01  0.0000000000000000D+00  THETA( 1)
-0.1397253183944389D-01  0.3441934196989738D-01  THETA( 2)

```

```

-0.8323603380299298D-02  0.4762451533260579D-01  THETA( 3)
-0.9344250653346974D-02  0.5477343529603911D-01  THETA( 4)
 0.8056675316624475D-01  0.4302966945632105D-01  THETA( 5)
 0.5089763104139401D+00  0.1777389057763482D-01  THETA( 6)
 0.2169310766190776D+00  0.3310732852507569D-01  THETA( 7)
 0.3069599594169708D+00  0.2748515884740973D-01  THETA( 8)
 0.1193061671637246D+00  0.3180382071492907D-01  THETA( 9)
 0.6751395602658273D-02  PLOT POINTS: LAG  3, MUSTAT( 1)
 0.6751395602658273D-02  PLOT POINTS: LAG  2, MUSTAT( 1)
 0.6751395602658273D-02  PLOT POINTS: LAG  1, MUSTAT( 1)

```

At this point, we have the wherewithal to make some progress. We shall now move upward to an $L_r = 3$, $L_p = 2$, $K_z = 4$, $I_z = 0$, $K_x = 1$, $I_x = 0$ parameterization with a perturbation of -0.0001 applied to the incremental coefficients of the polynomial part of the model by editing this file to read as follows.

s324010.in0

Stock s300000.out from s300000.in0 with TWEAK = 0.0000 by SNP 7.5

```

 1 M
744 N
 6 IDROP
 3 2 LRX, LPOLY
 4 0 IDEGY, LINTY
 1 0 IDEGX, LINTX
 1 ISTART (0 VAR START, 1 PARMFILE START)
400 ITMAX
1.0D-05 TOLER (D9.1)
 0 0 NLOG, IDIAG
 3.0 SFAC (F9.1)
50 NGRD
 0 IUNIT9 (0 PARMFILE, 1 RES, 2 MEAN, 3 VAR, 4 PLOT, 5 SIM)
 3 LRXO 0 LPOLYO 1 M
 0 IDEGYO 1 KYO 0 LINTYO
 0 IDEGXO 1 KXO 0 LINTXO
 0.1000000000000000D+01  0.0000000000000000D+00  THETA( 1)
-0.1397253183944389D-01  0.3441934196989738D-01  THETA( 2)
-0.8323603380299298D-02  0.4762451533260579D-01  THETA( 3)
-0.9344250653346974D-02  0.5477343529603911D-01  THETA( 4)
 0.8056675316624475D-01  0.4302966945632105D-01  THETA( 5)
 0.5089763104139401D+00  0.1777389057763482D-01  THETA( 6)
 0.2169310766190776D+00  0.3310732852507569D-01  THETA( 7)
 0.3069599594169708D+00  0.2748515884740973D-01  THETA( 8)
 0.1193061671637246D+00  0.3180382071492907D-01  THETA( 9)
 0.6751395602658273D-02  PLOT POINTS: LAG  3, MUSTAT( 1)
 0.6751395602658273D-02  PLOT POINTS: LAG  2, MUSTAT( 1)
 0.6751395602658273D-02  PLOT POINTS: LAG  1, MUSTAT( 1)

```

and editing control.dat to read:

control.dat

```

s300000.in0 s300000.out 0.0
s324010.in0 s324010.out -.0001

```

Running SNP we get:

summary.dat

```
TWK= 0.0000 ITR= 7/ 400 INF= 0 VAR Stock300000 8 1.31286 s300000.out
TWK= -0.0001 ITR= 32/ 400 INF= 0 S300000 Stock324010 22 1.22099 s324010.out
```

For purposes of illustration, we shall accept s324010.out as our final fit. In practice this is ill advised. We have found it much better to move slowly upward in small increments from a VAR and construct a table like Table 2. Before moving upward, a complete set of perturbations from, say, ± 0.00001 through ± 0.1 should be tried and the best outcome of the set should be selected as the .out file.

Continuing, we edit s324010.out by putting ITMAX to 0 and IUNIT9 to 1, 2, 3, 4, 5 to get s324010.in1, s324010.in2, s324010.in3, s324010.in4, s324010.in5 respectively. As an example, here is s324010.in4.

s324010.in4

```
Stock      s324010.out   from s324010.in0   with TWEAK = -0.0001 by SNP 7.5
  1         M
 744       N
  6        IDROP
  3  2 LRX,   LPOLY
  4  0 IDEGY, LINTY
  1  0 IDEGX, LINTX
  1        ISTART (0 VAR START, 1 PARMFILE START)
  0        ITMAX
1.0D-05 TOLER (D9.1)
  0  0 NLOG, IDIAG
  3.0 SFAC (F9.1)
 50        NGRD
  4        IUNIT9 (0 PARMFILE, 1 RES, 2 MEAN, 3 VAR, 4 PLOT, 5 SIM)
  3        LRXO   2        LPOLYO   1        M
  4        IDEGYO 5        KYO      0        LINTYO
  1        IDEGXO 3        KXO      0        LINTXO
0.1000000000000000D+01  0.0000000000000000D+00  THETA( 1)
0.1832444438423234D+00  0.1088300010005072D+00  THETA( 2)
-0.8086331136253797D-01  0.1436681211316008D+00  THETA( 3)
0.5129142049707111D-02  0.9758806356144807D-01  THETA( 4)
-0.3976197974533604D-01  0.5532179802038235D-01  THETA( 5)
0.2861808833896622D-01  0.6184155416769808D-01  THETA( 6)
-0.2869080742497851D+00  0.4084437538333630D-01  THETA( 7)
-0.2018410408237735D+00  0.4206488919761711D-01  THETA( 8)
0.1683752734514798D-01  0.4735169559295457D-01  THETA( 9)
-0.1101964481284351D-01  0.1997621742563651D-01  THETA(10)
0.7932235094232168D-02  0.1355720827176593D-01  THETA(11)
-0.1034517096810649D-01  0.1187138122000073D-01  THETA(12)
0.2488169930252172D-01  0.5938401260026879D-02  THETA(13)
0.2020288668801402D-01  0.6176820920266395D-02  THETA(14)
-0.3191018761082451D-02  0.6964328124506381D-02  THETA(15)
0.1423966151358360D-01  0.1083215509080528D+00  THETA(16)
0.2261225785502205D-01  0.4284527354136315D-01  THETA(17)
```

```

0.1454904302528228D-01    0.9416190148597978D-01    THETA( 18)
0.6304842943102193D-01    0.6745747275466799D-01    THETA( 19)
0.7869221447263167D+00    0.6817635810986175D-01    THETA( 20)
0.2688792679247046D+00    0.6511947113209941D-01    THETA( 21)
0.3497565740644625D+00    0.7579948874685423D-01    THETA( 22)
0.2023487967577421D-02    0.6433383283212217D-01    THETA( 23)
0.6751395602658273D-02    PLOT POINTS: LAG 3, MUSTAT( 1)
0.6751395602658273D-02    PLOT POINTS: LAG 2, MUSTAT( 1)
0.6751395602658273D-02    PLOT POINTS: LAG 1, MUSTAT( 1)

```

We then edit control.dat to read as follows.

control.dat

```

s300000.in0  s300000.out  0.0
s324010.in0  s324010.out  -.0001
s324010.in1  s324010.res  0.0
s324010.in2  s324010.mu  0.0
s324010.in3  s324010.sig  0.0
s324010.in4  s324010.plt  0.0
s324010.in5  s324010.sim  0.0

```

Running SNP we get

summary.dat

```

TWK= 0.0000 ITR= 7/ 400 INF= 0 VAR Stock300000 8 1.31286 s300000.out
TWK= -0.0001 ITR= 32/ 400 INF= 0 S300000 Stock324010 22 1.22099 s324010.out
TWK= 0.0000 ITR= 0/ 0 INF= 4 S324010 Stock324010 22 1.22099 s324010.res
TWK= 0.0000 ITR= 0/ 0 INF= 4 S324010 Stock324010 22 1.22099 s324010.mu
TWK= 0.0000 ITR= 0/ 0 INF= 4 S324010 Stock324010 22 1.22099 s324010.sig
TWK= 0.0000 ITR= 0/ 0 INF= 4 S324010 Stock324010 22 1.22099 s324010.plt
TWK= 0.0000 ITR= 0/ 0 INF= 4 S324010 Stock324010 22 1.22099 s324010.sim

```

The information in summary.dat is, reading across a row, the perturbation applied, the number of iterations, ITMAX, the NPSOL return code, a description of the starting model in the format

$$s L_r L_p K_z I_z K_x I_x$$

the fitted model in the same format with the user determined label replacing the leading s , p_θ , $s_n(\hat{\theta})$, and the file name of the file written to unit 9.

The file s324010.sim contains a simulation of length $N=744$ from the fit using $\text{NGRD}=50$ as the seed and the segment x_{t-1} of the series from 1 to $\text{IDROP}=6$ to start the simulation off. This simulation is plotted as the second panel of Figure 10. The units in .sim, as well as .plt, .sig, and .mu, are in the original units of the data; that is, in the same units as $\{\tilde{y}_t\}$.

The file s324010.plt contains data with which to plot the one-step-ahead density conditional on the values $\tilde{y}_{t-L}, \dots, \tilde{y}_{t-2}, \tilde{y}_{t-1}$ that are appended to the end of s324010.in4. The

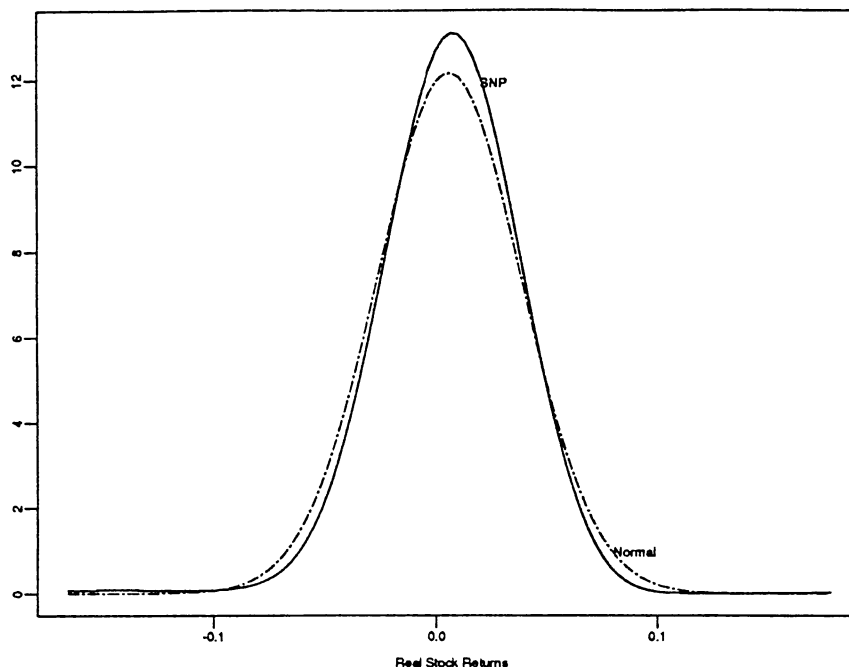


Figure 11: Monthly NYSE Returns, 1926-1987, Not Adjusted for Calendar Effects

program itself will append $\tilde{y}_{t-L}, \dots, \tilde{y}_{t-2}, \tilde{y}_{t-1}$ set to the mean \bar{y} and in this case we merely accept them:

```
0.6751395602658273D-02  PLOT POINTS: LAG 3, MUSTAT( 1)
0.6751395602658273D-02  PLOT POINTS: LAG 2, MUSTAT( 1)
0.6751395602658273D-02  PLOT POINTS: LAG 1, MUSTAT( 1)
```

To plot conditional on a different point, edit these values appropriately. These plot points are in the original units of the data, \tilde{y}_t , the program takes care of the transformation $\tilde{y}_t \rightarrow y_t \rightarrow x_{t-1}$ (and $\rightarrow \hat{x}_{t-1}$ if necessary).

The output file s324010.plt contains: first, $\hat{\mathcal{E}}(\tilde{y}_t|x_{t-1})$, M values; second, the upper triangle of $\hat{V}ar(\tilde{y}_t|x_{t-1})$ stored columnwise, $M*(M+1)/2$ values; third, the grid increment, M values; thereafter, \tilde{y}_t and $f_K(\tilde{y}_t|x_{t-1}, \hat{\theta})$ written end to end, there are $(M+1)*(2*NGRD+1)**M$ of these. Total file length is $M+M*(M+1)/2+M+(M+1)*(2*NGRD+1)**M$. The reason for pre-pending $\hat{\mathcal{E}}(\tilde{y}_t|x_{t-1})$ and $\hat{V}ar(\tilde{y}_t|x_{t-1})$ is that one often wants to compare the plot with the normal distribution at the same mean and variance as in Figure 11 which is a plot of the data in s324010.plt. The density $f_K(\tilde{y}_t|x_{t-1}, \hat{\theta})$ displays the typical shape for data from financial markets: peaked with fatter tails than the normal with a bit of asymmetry.

The options described above do not cover every contingency. For instance, to estimate a leverage function one would have to code numerous .in4 files and strip the conditional

variance out of the corresponding .plt files. The flexibility to perform these sorts of computations is provided by isolating the plot and simulation drivers so that they can be recoded to handle special situations. All control parameters and data that we could anticipate that one might need are passed through the argument lists of the subroutines in the packet of drivers. Nearly all other program control parameters and data reside in labelled common so that in cases where we anticipated incorrectly, inserting a labeled common statement in snpopen.f or one of the drivers will provide the necessary controls.

We illustrate with the modification to pltdrv.f that produced the data to plot the leverage function shown in the second panel of Figure 9:

pltdrv.f

```

SUBROUTINE PLTDEF(N,M,L,NGRD,SFAC,NPLOT,ISW)
IMPLICIT REAL*8 (A-H,O-Z)
NPLOT=100
NGRD=1
RETURN
END
SUBROUTINE PLTGET(X,MUSTAT,CSTAT,DATA,IDROP,N,M,L,IPLLOT,NGRD,SFAC)
IMPLICIT REAL*8 (A-H,O-Z)
REAL*8 X(M*L),DATA(M*(IDROP+N)),MUSTAT(M),CSTAT(M;M)
X(1)=0.6751395602658273D-02
X(2)=0.6751395602658273D-02
X(3)=-0.4d0+(DFLOAT(IPLLOT)/100.DO)*0.8D0
WRITE(9,9000) x(3)
RETURN
9000 FORMAT(' ',D25.16)
END
SUBROUTINE PLTPUT(D,IPLLOT,IPT)
IMPLICIT REAL*8 (A-H,O-Z)
IF (IPT.EQ.2) WRITE(9,9000) D
RETURN
9000 FORMAT(' ',D25.16)
END

```

Recent IMA Preprints

#	Author/s	Title
621	J. David Logan,	Wave propagation in a qualitative model of combustion under equilibrium conditions
622	M.L. Zeeman,	Hopf bifurcations in competitive three-dimensional Lotka-Volterra Systems
623	Allan P. Fordy,	Isospectral flows: their Hamiltonian structures, Miura maps and master symmetries
624	Daniel D. Joseph, John Nelson, Michael Renardy, and Yuriko Renardy,	Two-Dimensional cusped interfaces
625	Avner Friedman and Bei Hu,	A free boundary problem arising in electrophotography
626	Hamid Bellout, Avner Friedman and Victor Isakov,	Stability for an inverse problem in potential theory
627	Barbara Lee Keyfitz,	Shocks near the sonic line: A comparison between steady and unsteady models for change of type
628	Barbara Lee Keyfitz and Gerald G. Warnecke,	The existence of viscous profiles and admissibility for transonic shocks
629	P. Szmolyan,	Transversal heteroclinic and homoclinic orbits in singular perturbation problems
630	Philip Boyland,	Rotation sets and monotone periodic orbits for annulus homeomorphisms
631	Kenneth R. Meyer,	Apollonius coordinates, the N-body problem and continuation of periodic solutions
632	Chjan C. Lim,	On the Poincaré–Whitney circuitspace and other properties of an incidence matrix for binary trees
633	K.L. Cooke and I. Györi,	Numerical approximation of the solutions of delay differential equations on an infinite interval using piecewise constant arguments
634	Stanley Minkowitz and Matthew Witten,	Periodicity in cell proliferation using an asynchronous cell population
635	M. Chipot and G. Dal Maso,	Relaxed shape optimization: The case of nonnegative data for the Dirichlet problem
636	Jeffery M. Franke and Harlan W. Stech,	Extensions of an algorithm for the analysis of nongeneric Hopf bifurcations, with applications to delay-difference equations
637	Xinfu Chen,	Generation and propagation of the interface for reaction–diffusion equations
638	Philip Korman,	Dynamics of the Lotka–Volterra systems with diffusion
639	Harlan W. Stech,	Generic Hopf bifurcation in a class of integro-differential equations
640	Stephane Laederich,	Periodic solutions of non linear differential difference equations
641	Peter J. Olver,	Canonical Forms and Integrability of BiHamiltonian Systems
642	S.A. van Gils, M.P. Krupa and W.F. Langford,	Hopf bifurcation with nonsemisimple 1:1 Resonance
643	R.D. James and D. Kinderlehrer,	Frustration in ferromagnetic materials
644	Carlos Rocha,	Properties of the attractor of a scalar parabolic P.D.E.
645	Debra Lewis,	Lagrangian block diagonalization
646	Richard C. Churchill and David L. Rod,	On the determination of Ziglin monodromy groups
647	Xinfu Chen and Avner Friedman,	A nonlocal diffusion equation arising in terminally attached polymer chains
648	Peter Gritzmann and Victor Klee,	Inner and outer j - Rradii of convex bodies in finite-dimensional normed spaces
649	P. Szmolyan,	Analysis of a singularly perturbed traveling wave problem
650	Stanley Reiter and Carl P. Simon,	Decentralized dynamic processes for finding equilibrium
651	Fernando Reitich,	Singular solutions of a transmission problem in plane linear elasticity for wedge-shaped regions
652	Russell A. Johnson,	Cantor spectrum for the quasi-periodic Schrödinger equation
653	Wenxiong Liu,	Singular solutions for a convection diffusion equation with absorption
654	Deborah Brandon and William J. Hrusa,	Global existence of smooth shearing motions of a nonlinear viscoelastic fluid
655	James F. Reineck,	The connection matrix in Morse–Smale flows II
656	Claude Baesens, John Guckenheimer, Seunghwan Kim and Robert Mackay,	Simple resonance regions of torus diffeomorphisms
657	Willard Miller, Jr.,	Lecture notes in radar/sonar: Topics in Harmonic analysis with applications to radar and sonar
658	Calvin H. Wilcox,	Lecture notes in radar/sonar: Sonar and Radar Echo Structure
659	Richard E. Blahut,	Lecture notes in radar/sonar: Theory of remote surveillance algorithms
660	D.V. Anosov,	Hilbert’s 21st problem (according to Bolibruch)
661	Stephane Laederich,	Ray–Singer torsion for complex manifolds and the adiabatic limit

- 662 **Geneviève Raugel and George R. Sell**, Navier-Stokes equations in thin 3d domains: Global regularity of solutions I
- 663 **Emanuel Parzen**, Time series, statistics, and information
- 664 **Andrew Majda and Kevin Lamb**, Simplified equations for low Mach number combustion with strong heat release
- 665 **Ju. S. Il'yashenko**, Global analysis of the phase portrait for the Kuramoto–Sivashinsky equation
- 666 **James F. Reineck**, Continuation to gradient flows
- 667 **Mohamed Sami Elbially**, Simultaneous binary collisions in the collinear N–body problem
- 668 **John A. Jacquez and Carl P. Simon**, Aids: The epidemiological significance of two different mean rates of partner-change
- 669 **Carl P. Simon and John A. Jacquez**, Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations
- 670 **Matthew Stafford**, Markov partitions for expanding maps of the circle
- 671 **Ciprian Foias and Edriss S. Titi**, Determining nodes, finite difference schemes and inertial manifolds
- 672 **M.W. Smiley**, Global attractors and approximate inertial manifolds for abstract dissipative equations
- 673 **M.W. Smiley**, On the existence of smooth breathers for nonlinear wave equations
- 674 **Hitay Özbay and Janos Turi**, Robust stabilization of systems governed by singular integro-differential equations
- 675 **Mary Silber and Edgar Knobloch**, Hopf bifurcation on a square lattice
- 676 **Christophe Golé**, Ghost circles for twist maps
- 677 **Christophe Golé**, Ghost tori for monotone maps
- 678 **Christophe Golé**, Monotone maps of $T^m \times R^n$ and their periodic orbits
- 679 **E.G. Kalnins and W. Miller, Jr.**, Hypergeometric expansions of Heun polynomials
- 680 **Victor A. Pliss and George R. Sell**, Perturbations of attractors of differential equations
- 681 **Avner Friedman and Peter Knabner**, A transport model with micro- and macro-structure
- 682 **E.G. Kalnins and W. Miller, Jr.**, A note on group contractions and radar ambiguity functions
- 683 **George R. Sell**, References on dynamical systems
- 684 **Shui-Nee Chow, Kening Lu and George R. Sell**, Smoothness of inertial manifolds
- 685 **Shui-Nee Chow, Xiao-Biao Lin and Kening Lu**, Smooth invariant foliations in infinite dimensional spaces
- 686 **Kening Lu**, A Hartman–Grobman theorem for scalar reaction-diffusion equations
- 687 **Christophe Golé and Glen R. Hall**, Poincaré’s proof of Poincaré’s last geometric theorem
- 688 **Mario Taboada**, Approximate inertial manifolds for parabolic evolutionary equations via Yosida approximations
- 689 **Peter Rejto and Mario Taboada**, Weighted resolvent estimates for Volterra operators on unbounded intervals
- 690 **Joel D. Avrin**, Some examples of temperature bounds and concentration decay for a model of solid fuel combustion
- 691 **Susan Friedlander and Misha M. Vishik**, Lax pair formulation for the Euler equation
- 692 **H. Scott Dumas**, Ergodization rates for linear flow on the torus
- 693 **A. Eden, A.J. Milani and B. Nicolaenko**, Finite dimensional exponential attractors for semilinear wave equations with damping
- 694 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Inertial sets for dissipative evolution equations
- 695 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Hölder continuity for the inverse of Mañé’s projection
- 696 **Michel Chipot and Charles Collins**, Numerical approximations in variational problems with potential wells
- 697 **Huanan Yang**, Nonlinear wave analysis and convergence of MUSCL schemes
- 698 **László Gerencsér and Zsuzsanna Vágó**, A strong approximation theorem for estimator processes in continuous time
- 699 **László Gerencsér**, Multiple integrals with respect to L -mixing processes
- 700 **David Kinderlehrer and Pablo Pedregal**, Weak convergence of integrands and the Young measure representation
- 701 **Bo Deng**, Symbolic dynamics for chaotic systems
- 702 **P. Galdi, D.D. Joseph, L. Preziosi, S. Rionero**, Mathematical problems for miscible, incompressible fluids with Korteweg stresses
- 703 **Charles Collins and Mitchell Luskin**, Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem
- 704 **Peter Gritzmann and Victor Klee**, Computational complexity of inner and outer j -radii of polytopes in finite-dimensional normed spaces
- 705 **A. Ronald Gallant and George Tauchen**, A nonparametric approach to nonlinear time series analysis: estimation and simulation
- 706 **H.S. Dumas, J.A. Ellison and A.W. Sáenz**, Axial channeling in perfect crystals, the continuum model and the method of averaging