JOURNAL OF APPLIED ECONOMETRICS, VOL. 10, 109–125 (1995)

A NONLINEAR APPROACH TO US GNP

SIMON M. POTTER
Department of Economics, University of California, Los Angeles, 405 Hilgard Ave, Los Angeles, CA 90024-1477, USA

SUMMARY

A univariate nonlinear model is estimated for US GNP that on many criteria outperforms standard linear models. The estimated model is of the threshold autoregressive type and contains evidence of asymmetric effects of shocks over the business cycle. In particular the nonlinear model suggests that the post-1945 US economy is significantly more stable than the pre-1945 US economy.

1. INTRODUCTION

The modern study of economic fluctuations is built on the foundation that the economy's temporal behaviour can be well represented by random impulses being propagated through time by an invariant linear structure. This foundation is sometimes known as the extrinsic view or Frisch/Slutsky paradigm. An alternative earlier approach, often given the title intrinsic, places more emphasis on a nonlinear deterministic mechanism and little emphasis on random shocks for the generation of cycles. The subject of this paper is a synthesis of the propagation/impulse characterization of the extrinsic paradigm with the nonlinear structure of the intrinsic approach applied to the modeling of US GNP. I find that an important aspect of US GNP's time-series properties is hidden by the use of linear methods: the response of output to shocks at different stages of the business cycle is asymmetric. Starting with the seminal contribution of Hamilton (1989) a number of other researchers have also recently estimated nonlinear time-series models (Beaudry and Koop, 1993; Brunner, 1992; Mittnik, 1991; Teräsvirta and Anderson, 1992; Tiao and Tsay, 1991). As well as an abrupt (and asymmetric in the probabilistic sense) movement from expansion into recession and vice versa found in Hamilton's model, these researchers have also found similar additional asymmetries to those examined in this paper.

I find that the form of the asymmetry has very interesting economic content. Nonlinear models for post-1945 US GNP suggest that even if the economy was hit by negative shocks similar to the Great Depression era output would return to 'trend' quickly. Linear models for post-1945 US GNP show no evidence of increased stability with output remaining approximately below 'trend' for many years if hit by shocks of the magnitude experienced during the Great Depression. The extra stability found is based on post-1945 data and is not vulnerable to the suspicion of measurement error raised by the work of Romer for the pre-1945 data (for example, Romer, 1986).

Since US GNP is perhaps the most examined univariate time series in modern macroeconomics it is important to understand why previous studies based on linear models have not found the extra stability contained in the nonlinear model. The Wold Representation tells us that any purely nondeterministic covariance stationary time series has a representation as a linear mechanism

CCC 0883–7252/95/020109–17

Received February 1993
Revised May 1994

propagating *uncorrelated* impulses. In order to produce the linear model the Wold Representation exclusively utilizes the information contained in the autocovariance function of the time series which imposes strong symmetry properties on the dynamics of the estimated model.

The nonlinear approach used in this paper analyses models whose impulses are not predictable from their own past (martingale difference sequences). This is a stronger condition than lack of unconditional correlation and requires one to use moments of the data in addition to the autocovariance function to estimate the model. Thus, in some sense there is no information in the previous linear models about the form that the nonlinear model should take. Alternatively, the nonlinear model estimated in this paper has a Wold Representation that is the same (up to estimation error) as the Wold Representation of a linear model for US GNP but the nonlinear model contains additional information on dynamic properties of GNP that do not impose a symmetry restriction.

The plan of the paper is as follows. Section 2 describes the nonlinear model, Self-Exciting Threshold Autoregression, used in this paper and relates it to the other nonlinear models that have been estimated. It also reviews the statistical evidence in favour of US GNP containing nonlinearities. Section 3 introduces the concept of a nonlinear impulse response function as a means of illustrating asymmetries and provides examples of the asymmetries produced by the estimated nonlinear propagation mechanism. Section 4 provides an illustration of the estimated stabilizing property of post-1945 US output by attempting to 'recreate' the Great Depression. Section 5 introduces a test statistic for testing the null hypothesis of a symmetric propagation mechanism. I find that allowing for parameter uncertainty does not change the conclusion of asymmetries in the propagation mechanism. Section 6 offers conclusions.

2. SETAR MODELS: IDENTIFICATION AND ESTIMATION TECHNIQUES

2.1 Overview of Nonlinear Time-series Models

Self-Exciting Threshold Autoregressions (SETAR) and many of the other recent models estimated on economic time series are special cases of nonlinear models with a single index restriction (see Tong, 1990). Let Y_t represent the observed univariate time series, in our case it will be the first difference of US GNP and Z_t an unobserved time series. Let H_t denote the single index, which is assumed to be a continuous map from the history of $\{Y_t, Z_t\}$ to the line. Let $F(\cdot)$ be a function from the line to the unit interval with at most a finite number of discontinuities. Then a univariate first-order Single Index Generalized Multivariate Autoregressive (SIGMA) model would be:[1]

$$Y_t = \alpha_1 + \alpha_2 \mathrm{F}(H_t) + \{\phi_1 + \phi_2 \mathrm{F}(H_t)\} Y_{t-1} + \{\phi_1 + \phi_2 \mathrm{F}(H_t)\} V_t$$

where V_t are Independent and Identically Distributed (IID) with mean zero and unit variances. Some special cases are:

(1) If $\alpha_1 = \phi_2 = \psi_2 = 0$, we have an AR(1) model.
(2) If $\mathrm{F}(H_t) = Z_t$ and Z_t is a two-state Markov Chain then we have the regime-switching model of Hamilton (1990). Such models are different from the GNP model in Hamilton (1989) where Y_t is composed of two unobserved processes: $Y_t = Z_t + X_t$, where Z_t is a two-state Markov Chain and X_t is a Gaussian autoregression. The fact that both Z_t and X_t are unobservable makes the estimation of this model particularly difficult.

[1] See Potter (1990) for a more complete description.

(3) If $F(H_t) = 1(Y_{t-d} > r)$ then we have a SETAR $(1, d, r)$ model, where $1(A)$ is the indicator function equal to one if the event A occurs and zero otherwise, and d is known as the delay parameter and r the threshold parameter. In contrast to the Markov switching models, in the SETAR models the nonlinearity is defined by the directly observable history of the time series. This greatly simplifies estimation. Further, the probabilities of switching between regimes in the future are determined endogenously by the underlying model, whereas in the Markov switching model the probabilities are fixed. This gives the SETAR model much greater flexibility in fitting the observed data and a greater range of dynamic response.

SETAR models are also fitted to US GNP by Tiao and Tsay (1991). The single index I use is the two-quarter lagged growth rate of GNP with a threshold value of zero whereas they refine it to the following four-regime model:

Regime 1 if $Y_{t-2} \leqslant 0$, and $Y_{t-2} > Y_{t-1}$, a worsening recession
Regime 2 if $Y_{t-2} \leqslant 0$, and $Y_{t-2} < Y_{t-1}$, an improving recession
Regime 3 if $Y_{t-2} \geqslant 0$, and $Y_{t-2} > Y_{t-1}$, a contraction with negative growth
Regime 4 if $Y_{t-2} \geqslant 0$, and $Y_{t-2} < Y_{t-1}$, an expansion with increasing growth

(4) If $F([Y_{t-1} - r]/\gamma])$ is a cumulative distribution function then we have the Smooth Transition Autoregression (STAR) model of Chan and Tong (1985). Note that in the limit as $\gamma \to 0$ the SETAR model and STAR model are observationally equivalent. Teräsvirta and Anderson (1992) make extensive use of the logistic distribution function in their analysis of OECD Industrial Production indices.

(5) Beaudry and Koop (1993) define a single index on the logarithm of level of GNP: $H_t = X_t - \max\{X_t, X_{t-1}, X_{t-2}, \dots)$ and $Y_t = (1 - L)X_t$. They consider two possible functional forms: $F(\cdot)$ is the identity function and $F(\cdot)$ as an indicator function equal to one when H_t is greater than zero. The index used by Beaudry and Koop has the attractive characteristic that it comes into effect when output falls below its previous peak.

2.2 Review of Tests for Nonlinearity in US GNP

There are numerous conceptual and practical difficulties in providing definitive tests of a null hypothesis of linearity in a time series (see Brock and Potter, 1993, for an extended discussion). Some of the difficulties are shared with the controversy over testing for unit roots, others are new. With the exception of a test known as the Bispectrum the econometrician must commit to a particular linear model (i.e. the number of lags in the autoregressive and moving average parts of the model) in order to test for linearity. Furthermore, most tests (the main exceptions are the Bispectrum again and the BDS test[2]) require the econometrician to specify a direction to look in for evidence of nonlinearity. This causes problems in deciding the overall significance of a test if a number of directions are examined or moral hazard problems if the result of only one direction is reported (Leamer, 1978). Even if a specific direction can be derived by theoretical arguments one must still deal with the difficulties produced by nuisance parameters present under the alternative hypothesis but not under the null. In the SETAR model the nuisance parameters are the threshold, r, and the delay, d. Recently the issue has been given considerable attention in the econometric literature by Andrews and Ploberger (1993) and Hansen (1993) building on the work of Davies (1977) and other statisticians summarized in Tong (1990). An

[2] The Bispectrum test is based on the ratio of the third-order polyspectra of the time series to its second-order spectra which is constant for a linear time series. The BDS test is based on the a scaling property of the correlation integral for IID stochastic process. Further descriptions can be found in Brock and Potter (1993).

application of Hansen's approach to the problem is discussed below after the estimation of the SETAR model.

A number of researchers have tried a large number of nonlinearity tests on US GNP. A summary of the results is as follows:[3]

(1) The Bispectrum and BDS tests cannot reject the null hypothesis of linearity, and this result is supported by other researchers (e.g. Brock and Sayers, 1988).
(2) Polynomial type tests (i.e. testing the orthogonality of the residuals from the linear model against polynomial functions of the observed history (see Tsay, 1986) have also been applied. For certain choices of the polynomial function it is possible to reject linearity. However, the overall significance of these rejections is uncertain since one is basically minimizing over probability values;
(3) Tsay's (1989, 1991) Recursive polynomial tests find strong evidence of nonlinearity using the two-quarter-lagged-growth rate as an index to define the ordering of the recursion. It was this type of test that led to my original adoption of the second-quarter lagged growth rate as the index. Tiao and Tsay (1991) quote a probability value of 0·026 for a similar test. However, in Potter (1990) I show that if one takes account of the fact that a number of directions are examined in a similar test then the overall significance level is well above 5%.

2.3 A SETAR Model for US GNP

The observations are post-Second World War quarterly, seasonally adjusted, real US GNP drawn from the Citibase data bank. The sample used is 1947Q1 to 1990Q4.[4] Prior to analysis logarithms and first differences of the data were taken (the result was multiplied by 100 and called Y_t). Table I reports results for a fifth-order autoregressive specification. The number shown for AIC is the value of Akaike's Information Criterion which is a weighting of the residual variance by the number of parameters estimated. Smaller values indicate better-fitting models.[5]

A SETAR $(p, d, r_1, \dots r_k)$ has the following form:

$$Y_t = \alpha_i + \phi_i(L)Y_{t-1} + e_{it} \quad \text{if} \quad Y_{t-d} \in A_i, \qquad i = 1, \dots k$$

where $\phi_i(L) = \phi_{1i} + \phi_{2i}L + \cdots + \phi_{pi}L^{p-1}$, $LY_t = Y_{t-1}$ and $A_i = [r_{i-1}, r_i)$.

Tong (1983, 1990) in his authoritative books on threshold models suggested a grid search method for estimation of the structural parameters d, $\{r_i\}$. Potter (1990) contains an extensive

Table I. Linear model results 1948Q3 to 1990Q4

Coefficient	Estimate	Standard error
Intercept	0·540	0·122
AR1	0·330	0·078
AR2	0·193	0·082
AR3	−0·105	0·083
AR4	−0·092	0·082
AR5	−0·024	0·078
S.E. of regression 0·99415		
Number of observations 170		
AIC 8·00		

[3] I have carried out all the tests below myself and give references to other researchers who have used the same tests.
[4] Drawn in the first quarter of 1991, with the 1982 index year.
[5] See Tong (1983) for a description of AIC and its use in nonlinear time series.

discussion of a graphical and testing approach to 'estimating' d, $\{r_i\}$ that is developed from earlier work by Tsay (1991). This method has the advantage over the grid search method of not restricting attention to SETAR models in the initial steps but the disadvantage of introducing a subjective element on the part of the individual researcher that is hard to quantify. These testing and graphical techniques produced estimates of $d=2$ and $r=0$ (Potter, 1990). A similar set of techniques were also used by Tiao and Tsay on US GNP with identical results. Hansen (1993) used a grid search-based method on the same data that produced an estimate of $d=2$ and an estimate of $r=0.1$.

Given the single index lag 2 and threshold of zero, two linear least squares estimation techniques are available:

(1) One can split the data into two groups and run a least squares regression for each regime separately. Thus, the estimated residual variance for each regime will be different (the estimation approach used below).
(2) One can run a single regression with indicator functions given by the single index multiplying the lags of the time series. Thus, the estimated residual variance is restricted to be constant across the regimes.

The second method is useful if one wishes to restrict certain estimated coefficients to be the same across regimes or an exogenous variable is introduced whose regression coefficient does not change with the single index.[6] Both methods give consistent estimates for the intercept and slope coefficients in each regime (see Tong, 1990), conditional on the correct choice of r and d.

It is useful to label the cases where $Y_{t-2} \leq 0$ regime 1 or the contractionary regime, and those where $Y_{t-2} > 0$ regime 2 or the expansionary regime. Table II contains estimation results based on the use of $1(Y_{t-2} > 0)$ as the relevant nonlinear function.

The coefficients on the AR3 and AR4 terms in Table II are not significantly different from zero at the 5% level and have similar values in both regimes. Therefore, the nonlinear model

Table II. SETAR without restrictions 1948Q3 to 1990Q4

	Regime 1 estimate (S.E.) $Y_{t-2} \leq 0$	Regime 2 estimate (S.E.) $Y_{t-2} > 0$
Intercept	−0·705 (0·480)	0·545 (0·161)
AR1	0·510 (0·192)	0·312 (0·081)
AR2	−0·849 (0·416)	0·245 (0·113)
AR3	−0·048 (0·223)	−0·104 (0·084)
AR4	−0·123 (0·275)	−0·057 (0·077)
AR5	0·398 (0·240)	−0·094 (0·076)
$\hat{\sigma}^2$	1·59	0·758
Observations	37	133
Standard error of regression = 0·95948, AIC = 4·22		

[6] Both approaches are easy to implement in a standard regression package. Assuming the package has a sign function or similar one would create a dummy variable for each regime and interact them with the intercept and the autoregressive lags. For the second approach the standard regression output would be correct. For the first one could interact the residuals with the dummy variables to form estimates of the residual variation in each regime. However, the generated standard errors for the regression coefficients would be incorrect. In practice the heteroscedastic standard errors produced by most regression packages are similar to those produced by using the separate regime residual variances.

was re-estimated with these terms restricted to be zero. The restriction was supported by the fall in the AIC and there was no evidence of autocorrelation in the residuals. Table III contains the results of the restricted estimation. The presence of the two AR5 terms may seem somewhat strange but they improve the fit of the model and are jointly significantly different from zero at the 10% level.[7] Tiao and Tsay (1991) in their model for US GNP decide to ignore the AR5 terms without affecting any qualitative conclusions.

Perhaps the most striking aspect of the results is the very large negative coefficient on the AR2 term in the contractionary regime (remember in this regime the lag 2 value must be negative implying a positive effect on growth when multiplied by the negative coefficient).[8] For both the impulse response and stability properties of the post-1945 propagation mechanism this will be a crucial term. Another striking aspect of the results is the size of the estimated discontinuity for $Y_{t-2}=0$. However, note that the residual variance doubles in the contractionary regime, smoothing the sample path effect of the discontinuity with additional noise.

The estimated nonlinearity was found to be robust to the use of subsamples. The qualitative results are also robust to the maintained assumption of an integrated specification for output. If one estimates the model in levels but uses the same single index defined on the growth rates r then the short-run dynamics discussed below are not affected.

The combined error variance indicates a reduction of 7% over the linear model. In order to assess whether this indicated a real improvement or just overfitting in-sample the nonlinear model and linear model were estimated recursively. The correlation of the recursive forecasts with actual output growth from 1960(i) to 1990(iv) was 0·23 for the linear model, 0·25 for the unrestricted SETAR and 0·35 for the restricted SETAR. A six-variable (output, consumption, investment, interest rates, money and prices) five-lag equation was also estimated recursively for output growth. Its recursive forecasts have a correlation of 0·44 (0·32 for the univariate model) with output growth from 1979(iv) to 1990(iv) compared to 0·47 and 0·52 for the

Table III. SETAR with restrictions 1948Q3 to 1990Q4

	Regime 1 estimate (S.E.) $Y_{t-2} \leqslant 0$	Regime 2 estimate (S.E.) $Y_{t-2} > 0$
Intercept	−0·808 (0·423)	0·517 (0·161)
AR1	0·516 (0·185)	0·299 (0·080)
AR2	−0·946 (0·353)	0·189 (0·107)
AR5	0·352 (0·216)	−1·143 (0·069)
$\hat{\sigma}^2$	1·50	0·763
Observations	37	133
Standard error of regression = 0·95597, AIC = −4·89		

[7] Potter (1990) estimates a SETAR model on seasonally unadjusted data and finds the large change in the intercept and AR2 coefficients is still present.

[8] Potter (1990) presents a Monte Carlo evaluation of the potential for small-sample bias to be causing the results. It is found that the choice of the second lag of the growth rate as the index does produce some downward bias in the first regime estimate of the AR2 coefficient but is incapable of simultaneously moving the intercept in the required direction in the first regime. Hansen (1992) also finds evidence using a Hamilton (1990) type model that Gaussian linear models fitted to US GNP can be rejected by a model that allows for perfect correlation between movements in the intercept and AR2 coefficient.

unrestricted and restricted SETAR models, respectively.[9] Tiao and Tsay (1991) obtain an even more dramatic forecast improvement for the refined four-regime SETAR model and present additional evidence on the forecasting superiority of the two-regime SETAR model over linear autoregressions.

2.4 Statistical Evidence in Favour of SETAR Nonlinearity in US GNP

Hansen (1993) directly evaluates the restricted SETAR model estimated above against a restricted linear AR(5) (that is, the AR3 and AR4 coefficients are restricted to be zero) with techniques that allow for presence of the nuisance parameters, r and d. This is achieved by simulating an empirical distribution, under the null hypothesis of the linear model, for functionals of the collection of Lagrange Multiplier (LM) test statistics produced by searching over a grid values for r and d. Specifically, Hansen allows the delay to take the values 1, 2 or 5 and the threshold to range from the 15th quantile to the 85th quantile of the marginal distribution of GNP growth rate, in the regression model

$$Y_t = \alpha_1 + \phi_1(L)Y_{t-1} + 1(Y_{t-d} < r)\{\alpha_2 + \phi_2(L)Y_{t-1}\} + V_t$$

where $\phi_i(L) = \phi_{i1}L + \phi_{i2}L^2 + \phi_{i5}L^5$, $i = 1$ and 2. The LM test is calculated for the null hypothesis that $\alpha_2 = \phi_2(L) = 0$ at each point in the grid values of r and d.

Hansen examines three functionals of the collection of LM statistics produced: the supremum (SupLM), the average (AveLM) and the exponential average, $\ln\{1/N \sum \exp(\mathrm{LM}_i/2\}$, (ExpLM). The latter two test statistics are discussed in Andrews and Ploberger (1993) where various optimality properties for them are derived. Hansen also calculates heteroscedastic robust versions of each of the test statistics (SupLMH, AveLMH and ExpLMH, respectively). Test statistics and asymptotic p-values for these tests are reported in Table IV. Hansen concludes from the six test statistics that the evidence does not support the rejection of the linear null model at conventional significance levels, with only the SupLM statistic finding evidence at the 5% level.

Table IV. Test results of linear model versus restricted SETAR[a]

TEST	Statistic	P-value	Size	Power	Size-adjusted P-value
SupLM	18·2	0·04	0·035	0·809	0·025
ExpLM	4·8	0·09	0·030	0·790	0·065
AveLM	4·6	0·29	0·026	0·470	0·241
SupLMH	14·1	0·17	0·014	0·231	0·087
ExpLMH	4·0	0·17	0·016	0·245	0·095
AveLMH	4·7	0·27	0·015	0·159	0·255

[a] See Hansen (1993) for a description of the derivation of the test statistics. The size and power numbers are based on 1000 bootstrap replications of the linear model and restricted SETAR model at a nominal size of 5%.

[9] To assess whether these results could be explained by parameter instability in less restricted models the *smallest* forecast errors of univariate linear models ranging from a univariate AR(1) to an AR(5) were compared to the forecast errors from the SETAR model for the 1979(iv) to 1990(iv) sample period. The minimization over the forecast errors of the linear models produced a root mean squared error of 0·821 (none of the individual models performed better than either nonlinear model) compared to 0·858 and 0·821 for the unrestricted the restricted models respectively (if one minimizes over the forecast errors of the unrestricted and restricted nonlinear model one obtains 0·800 for the RMSE).

A Monte Carlo experiment was performed to assess the finite sample properties of the asymptotic p-values for the test statistics. 1000 replications of the linear null model from Table I, and 1000 replications of the nonlinear restricted model from Table III were constructed by drawing with replacement from the estimated residuals and the test statistics were calculated for each replication. Table IV contains the size and power of the tests at a nomimal 5% significance level. It appears that the finite sample size of the tests is conservative especially for the heteroscedastic robust statistics. The power of the heteroscedastic robust tests at the nominal 5% level is very weak. The power of the SupLM and ExpLM is much higher. The last column of Table IV uses the information from the Monte Carlo experiment for finite sample size to adjust the asymptotic p-values for the test statistics generated by the observed sample. A review of the evidence in Table IV suggests that at the 10% level there is statistically significant evidence against the linear model.

Additional evidence in favour of the estimated model can be found in the similar models that have been estimated. A different type of supporting evidence can be found in studies of business cycle duration and turning points that estimate moments other than the first and second but do not estimate a particular model. Diebold and Rudebusch (1991) and Sichel (1992) find evidence of duration dependence during post-1945 recessions. The large negative coefficient on the second lag of the growth rate in the contractionary regime would tend to produce evidence of duration dependence, especially in conjunction with the very negative intercept term that ensures that if contractions persist they will get deeper and the stabilizing effect of the AR2 coefficient will come into play. Sichel (1994) presents evidence that the growth rate in recoveries tends to be higher than the average expansion growth rate. This is difficult to reconcile with linear models but is a potential property of the SETAR model estimated above with the caveat that the magnitude of the recovery should be positively correlated with the magnitude of the recession.

Pesaran and Potter (1991) apply the Cox Non-nested Testing methodology to distinguish between different nonlinear representations of the data. They concentrate on a comparison of Hamilton's (1989) model of GNP and the SETAR model above. If Hamilton's model is taken as the null model and the SETAR model as the alternative then it is possible to reject the Markov trend model. However, the SETAR cannot be rejected as the null model against the Hamilton model as the alternative. The test statistics were produced by simulation and there are potentially important conceptual difficulties produced by the nonlinear estimation required for Hamilton's model in this simulation procedure (see Hansen, 1992; Pesaran and Pesaran, 1993).[10]

3. USING NONLINEAR IMPULSE RESPONSE FUNCTIONS TO ASSESS ASYMMETRY

3.1 Definition of Nonlinear Impulse Response Functions

The estimated SETAR models contain two main asymmetric effects between the contraction and expansion regimes: the change in the intercept and the AR2 value. In order to illustrate and quantify the extent of the asymmetry I shall use a Nonlinear Impulse Response Function (NLIRF). NLIRFs are defined in a similar manner to standard impulse response functions except one replaces the linear predictor with a conditional expectation:

$$\mathrm{NLIRF}_n(v; y_t, y_{t-1}, \ldots) = \mathrm{E}[Y_{t+n} \mid Y_t = y_t + v, Y_{t-1}, \ldots] - \mathrm{E}[Y_{t+n} \mid Y_t = y_t, Y_{t-1} = y_{t-1}, \ldots]$$

where lower-case letters represent realized values and v is the postulated impulse.

[10] Note that Hansen's (1992) test of the Markov switching model is not correctly implemented (see Hansen, 1994).

For example, if $Y_t = \phi Y_{t-1} + V_t$, with V_t IID, then $\text{NLIRF}_n(v; y_t, y_{t-1}, \ldots) = \phi^n v$.

This result is identical to standard linear impulse response functions or transfer functions. Note that the response is independent of the history of the time series and the sign and magnitude of the postulated shock when suitably scaled.

Nonlinear models, in contrast, produce impulse response functions that are themselves functions of the history of the time series and the size and magnitude of the shock. Asymmetric response occurs in two main forms: (1) for any specific history the effect of shocks of varying magnitudes and signs is not a simple scaling of a unit shock; (2) for the same shock but different histories the response can differ markedly. Except for a few special cases analytical results are not available and Monte Carlo Integration techniques as described in Potter (1991) are required. As discussed in Potter (1991) and Gallant *et al.* (1993) there is a difficulty in summarizing the information contained in the NLIRFs produced by all the possible different histories and shocks. In this paper I present five examples that appear to represent qualitatively the possible dynamic responses from the estimated model. In Potter (1991) I calculate some measures of persistence for the estimated nonlinear GNP model. These measures suggest that the nonlinear model implies more persistence in GNP than previously found using linear models.

3.2 Comparison with Previous Approaches to Measuring Asymmetry

Neftci (1984) introduced the idea of modelling business cycle asymmetry by measuring differences in the retention probabilities for positive and negative growth in an economic time series. That is:[11]

$$P_{11} = \text{P}[Y_t > 0 \mid Y_{t-1} > 0]$$
$$P_{22} = \text{P}[Y_t \leq 0) \mid Y_{t-1} \leq 0]$$

It is clear that an econometrician could find statistically significant evidence that $P_{11} \neq P_{22}$ even if the underlying time-series process was linear (i.e. has a symmetric propagation mechanism) as long as the process is driven by asymmetric innovations. Similarly, the approach advocated by Sichel (1988) based on the skewness coefficient cannot distinguish between asymmetries in the propagation mechanism and asymmetries to the innovations. In addition, neither of these approaches is useful for describing the extent of the asymmetry.

Teräsvirta and Anderson (1992) suggest analysing the roots of the characteristic polynomial of the lag operator defined by specific values of the single index of a nonlinear model. Although directly concentrating on the propagation mechanism their approach ignores the effect of switching intercepts in the estimated equation and the future movement between regimes. Therefore, it is potentially very misleading. Other studies have used the 'zero innovation' impulse response function where all shocks after time t are set to zero. As discussed in Potter (1991), the zero innovation prediction can be very uninformative about the behaviour of the conditional expectation.

3.3 Nonlinear Impulse Response Functions for US GNP

Figures 1–5 give the possible effects of the positive and negative shocks of ±1% and ±2% on the logarithm of the level of output for various historical periods. The graphs start at the date before the shock. The size of the shock for each line can be found by looking at the effect at the second date on the x-axis (this matches the date in the title). Dates instead of conventional numbering for the horizon of the NLIRF are used to emphasize the history dependence.

(1) If the shock keeps the growth rate positive (Figure 1, 1984Q1) then the response is very similar to that obtained from a linear model. If one calculated an impulse response function

[11] Neftci actually examined a second-order case but it is easier to use a first-order example.

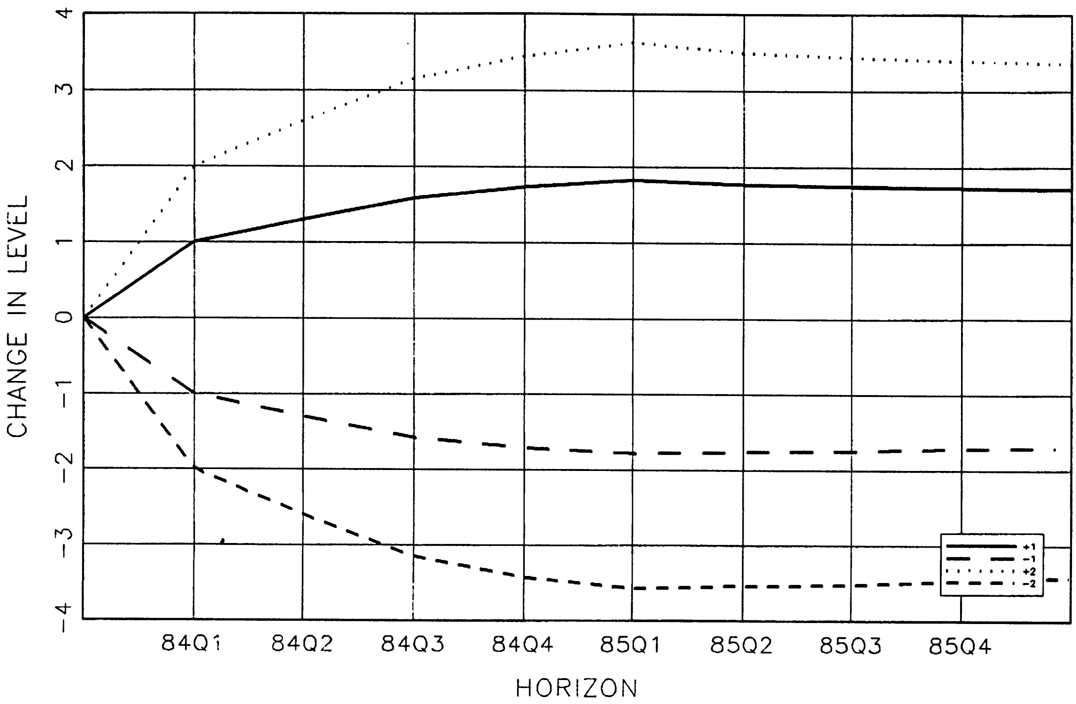

Figure 1. Nonlinear impulse response 1984Q1

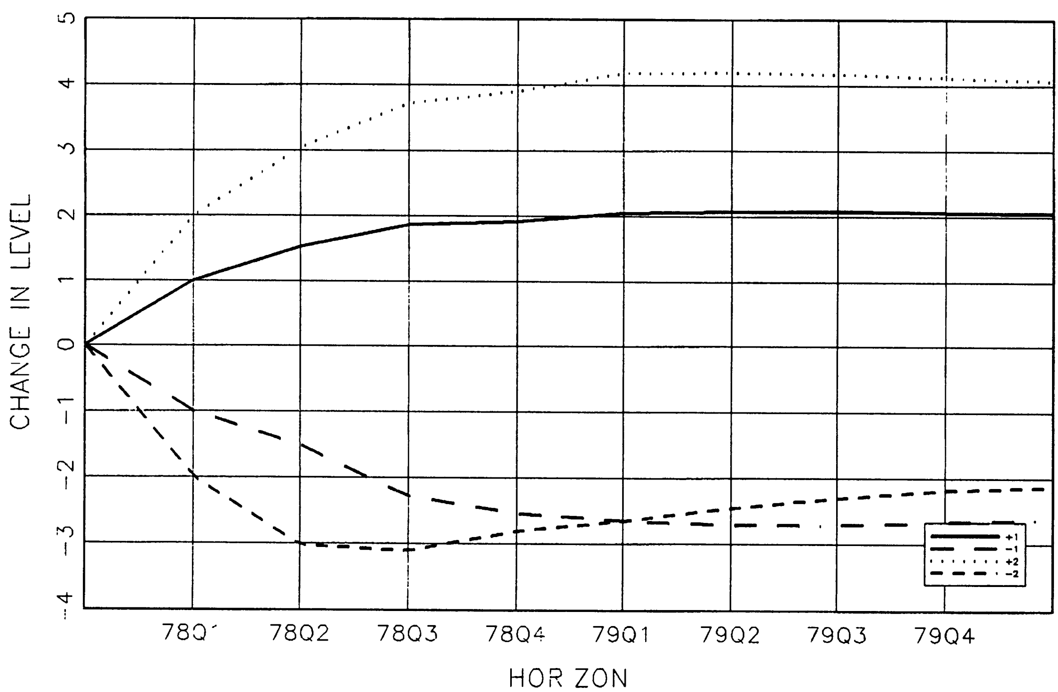

Figure 2. Nonlinear impulse response 1978Q1

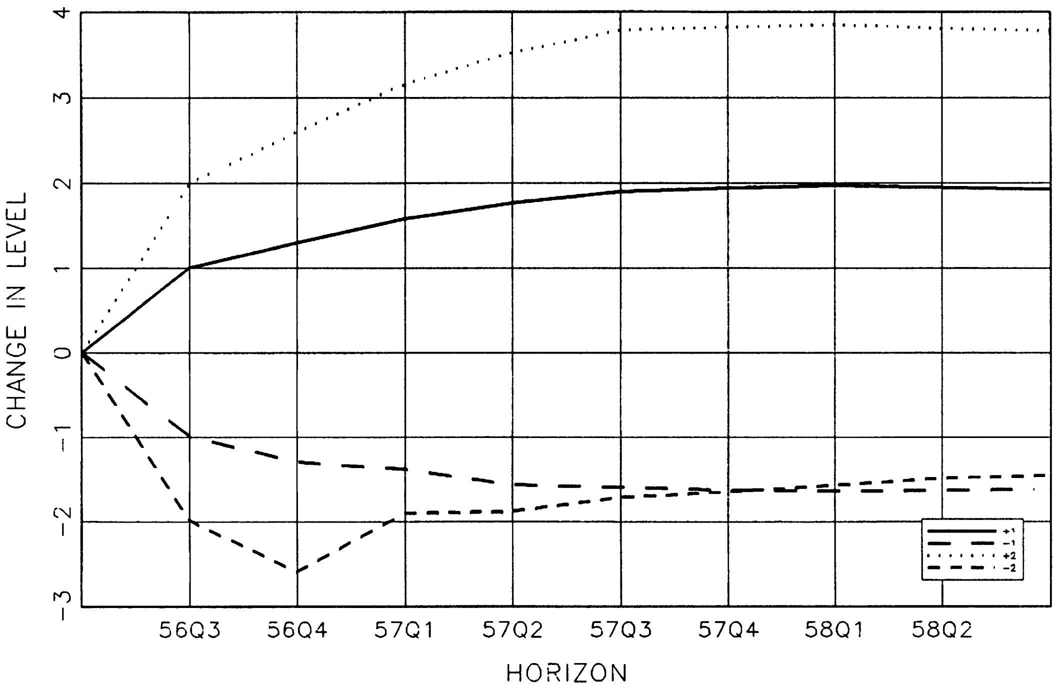

Figure 3. Nonlinear impulse response 1956Q3

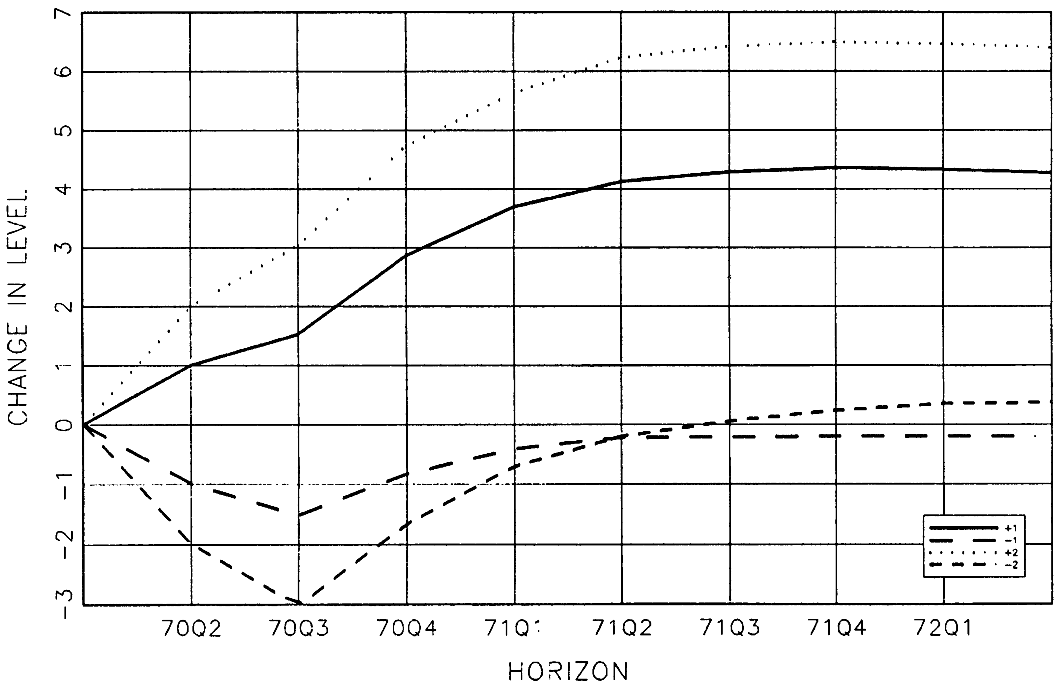

Figure 4. Nonlinear impulse response 1970Q2

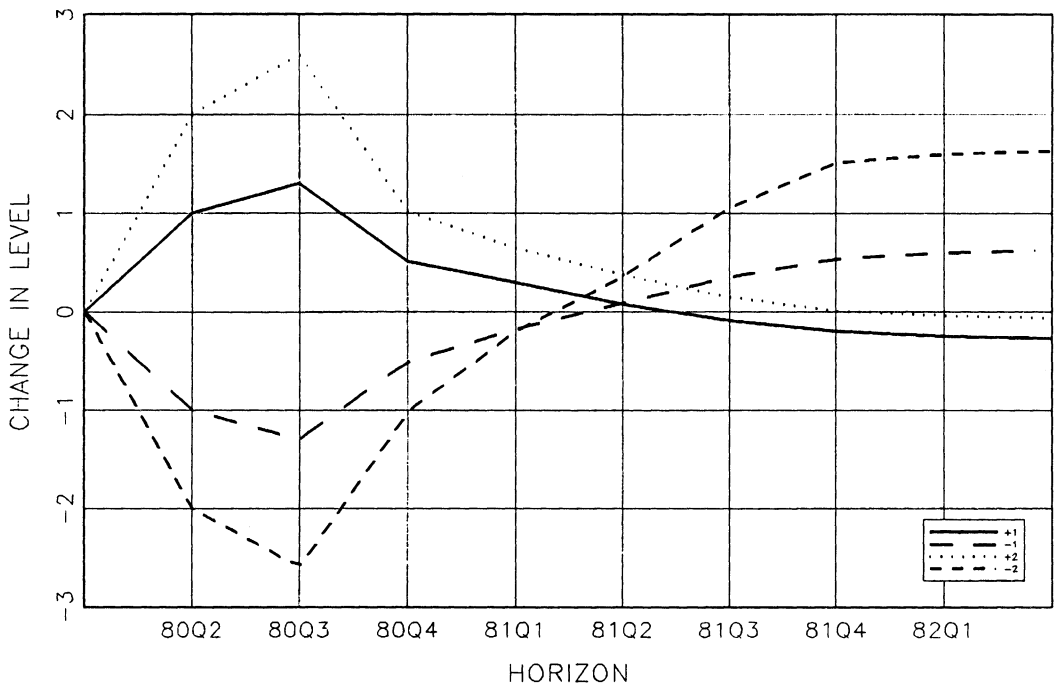

Figure 5. Nonlinear impulse response 1980Q2

for the linear AR(5) of Table I it would look very similar to Figure 1 for *all shocks and histories* of output growth.

(2) If the negative unit shock turns the growth rate negative (Figure 2, 1978Q1) then its effect will be magnified compared with Figure 1 by the switch in the intercept term. Magnification also occurs for the positive shock because the probability of a future contraction decreases. This is a similar effect to the abrupt switch between contraction and expansion in Hamilton's GNP model. However, for the negative two shock the stabilizing influence of the AR2 coefficient in the contractionary regime starts to take hold. Notice that the effect after two years of the negative two shock is smaller in absolute value than the negative unit shock. Hamilton's GNP model is unable to capture such an effect since it constrains the probability of movement out of contraction to be fixed no matter how large is the negative shock.

(3) If the value of growth perturbed in the starting values is only slightly greater than zero (Figure 3, 1956Q3) then for the positive shocks the effects are very similar to Figure 2. For negative shocks the effect after two years is now smaller than in Figure 2 because of the increased stabilizing effect of the AR2 coefficient.

(4) If the value of growth perturbed in the starting value is only slightly below zero (Figure 4, 1970Q2) then there is an approximate doubling of the effects of positive shocks compared to Figure 1. The main reason is the switch in the intercept values produced by the perturbation. For the negative shocks the stabilizing property is now more powerful with output returning to 'trend' after 8 quarters.

(5) Figure 5 shows another possibility for 1980Q2 where the negative growth of 2·3% produces a reaction to positive shocks similar to a trend stationary process but negative shocks tend to increase the logarithm of the level of output. That is, the effect of the positive shock is cancelled by the reduction in the strength of the stabilizing force, whereas the negative

shock adds to the power of the stabilizing force. I label this effect an intrinsic stabilizer. One can estimate the SETAR model in a trend stationary form (detrend the logarithmic level of GNP with a time trend then define regimes by whether the growth rate of GNP two quarters ago was positive or not) and obtain impulse response functions for 1980Q2 with very similar responses for the two-year horizon (see Potter, 1994).

4. PROPAGATING GREAT DEPRESSION TYPE IMPULSES: LINEAR AND NONLINEAR PERSPECTIVES

The reaction to negative shocks for periods of negative growth suggests that the post-1945 US economy contains an 'intrinsic' stabilizer. In this section I illustrate how covariance analysis might hide significant stability properties by conducting the following thought experiment. Imagine in the current quarter and for the next eighteen quarters the impulses to US GNP were of 'Great Depression' size and magnitude. What would be the effect on output if growth rates prior to the occurrence of the shocks were the same as in 1928 and 1929? I took Great Depression sized residuals[12] from simple linear and SETAR nonlinear models estimated on the pre-war GNP series reported in the Data Appendix of Gordon (1986). Both models were based on five lags of output growth rates, the nonlinear one having a threshold of zero and a delay of 1. I used data from 1875Q1 to 1933Q1 to estimate the models.

Figure 6 shows the effect of such shocks from linear and nonlinear perspectives and the actual realization of US GNP from 1929 to 1933. By linear perspective I mean using the residuals from the pre-1945 linear model to propagate the post-1945 linear model. For the

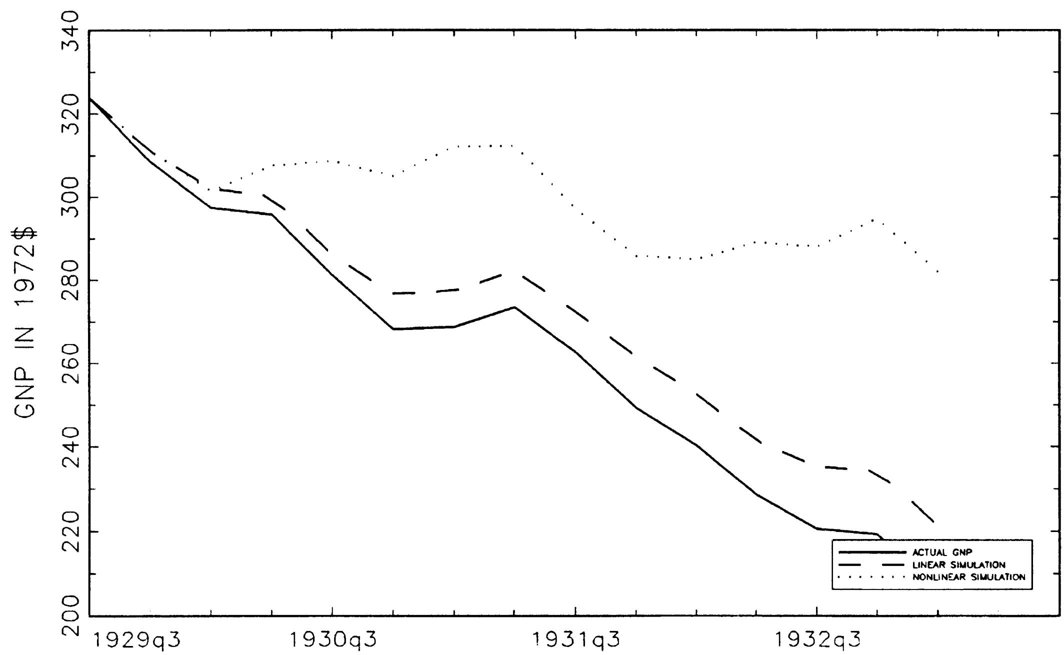

Figure 6. Simulation of effect of Great Depression type shocks on linear and nonlinear post-1945 propagation mechanisms

[12] The standard deviation of the residuals from the linear and nonlinear models are very similar. However, the linear residuals contain approximately an extra 2% drop in GNP.

nonlinear perspective I took residuals from the pre-1945 nonlinear model and used them to propagate the post-1945 restricted nonlinear model. Observe that the post-1945 linear model shows little improvement in the stability properties of the US economy (i.e. it mimics the path of the actual Great Depression) but the nonlinear model does (i.e. it does not produce a sample path that looks like the Great Depression). The asymmetric response to shocks suggests that since the Second World War the long-run effect of large negative shocks has been considerably diminished.

The quarterly data used are subject to important measurement errors as to the exact timing of the fluctuations but the total fall in GNP in the time period examined is similar to that found in more accurate annual data. Thus impulses of approximately the correct overall magnitude have been used. It is crucial to realize that covariance analysis, as represented by the linear perspective above, is incapable of uncovering such extra stability. I now turn to assessing the statistical significance of the extra stability in the post-1945 propagation mechanism conditional on the SETAR specification.

5. TESTING FOR ASYMMETRIC EFFECTS ALLOWING FOR PARAMETER UNCERTAINTY

Define a measure of asymmetry, ASYM, as follows:

$$\mathrm{ASYM}_k(v; y_t, y_{t-1}, \ldots) = \mathrm{NLIRF}_k(v; y_t, y_{t-1}, \ldots) + \mathrm{NLIRF}_k(-v; y_t, y_{t-1}, \ldots)$$

If Y_t is a linear time series then ASYM_k is identically equal to zero. For example, in the case of the linear AR(1) where the NLIRF is independent of the history of the process: $\mathrm{ASYM}_n(v) = \phi^n v + \phi^n(-v) = 0$, for all values of v, n. Hence, actual sample paths of a time series could be highly asymmetric due to asymmetries in the innovations but if the propagation mechanism is linear then ASYM_k will be identically zero.

One expects all nonlinear time-series models to contain asymmetries as defined above. The finding of no asymmetries would be as surprising as a linear impulse response function that was identically equal to zero after the initial impulse. It is standard to assess whether the impulse response function is significantly different from zero given the sampling variability. Similarly, a test is required to see whether the asymmetry is significant given the sampling variability in the estimated models. The reader should realize that such a test cannot take into account the uncertainty surrounding the choice of a SETAR specification to represent the nonlinearity.[13]

The asymmetries in the estimated SETAR model are only present for certain values of the shock and for particular histories of the growth rate. Hence, the statistical significance of the test below is conditional on the particular history and shock chosen. Since the economic significance of the stabilizing effect found is conditional on negative growth such a test is valid. Furthermore, we are only interested in a one-sided confidence interval of the negative of the response to a negative shock being smaller in magnitude than the response to a positive shock (i.e. the stabilizing effect), unlike the two-sided case for linear impulse response functions. The following simulation technique is used to form a sampling distribution for the NLIRFs:

(1) Pick a particular set of growth rates from the distribution of realized growth rates and a value for the shock to perturb the most recent growth rate in the initial condition in a positive and negative direction.

[13] A strength of the more nonparametric approach of Gallant *et al.* (1993) is that it allows for uncertainty surrounding the true model specification in evaluating asymmetries in impulse response functions.

(2) Use the estimated coefficients and their variance covariance matrix to generate M samples of the regime coefficients by drawing from their asymptotic normal variance covariance matrix.[14]
(3) From the M samples of the regime coefficients generate the set of asymmetry statistics, $\{ASYM_k\}$ by calculating the NLIRF for each draw.
(4) Find the 5th percentile of $\{ASYM_k\}$. Call it χ. If $\chi \leq 0$ then the symmetric response lies within a 95% confidence interval for the particular initial condition and shock. If $\chi > 0$ then the symmetric response lies outside the 95% confidence interval for the particular initial condition and shock.
(5) The effect of the variability introduced by Monte Carlo Integration is to make rejection of the null hypothesis of asymmetry more likely, hence the actual size of the test will be higher than 5%.

Results for an initial condition of 1990Q4 and shocks of ±2% using the coefficient estimates and variance covariance matrix from the unrestricted SETAR model for an eight-quarter horizon are shown in Figure 7.[15] The figure displays the mean and median values of the asymmetry measure over the draws from the variance covariance matrix, the 5th percentile and the asymmetry measure at the point estimates from Table III. It also shows that the asymmetric effect is statistically significant and that there is little asymmetry between the mean and the

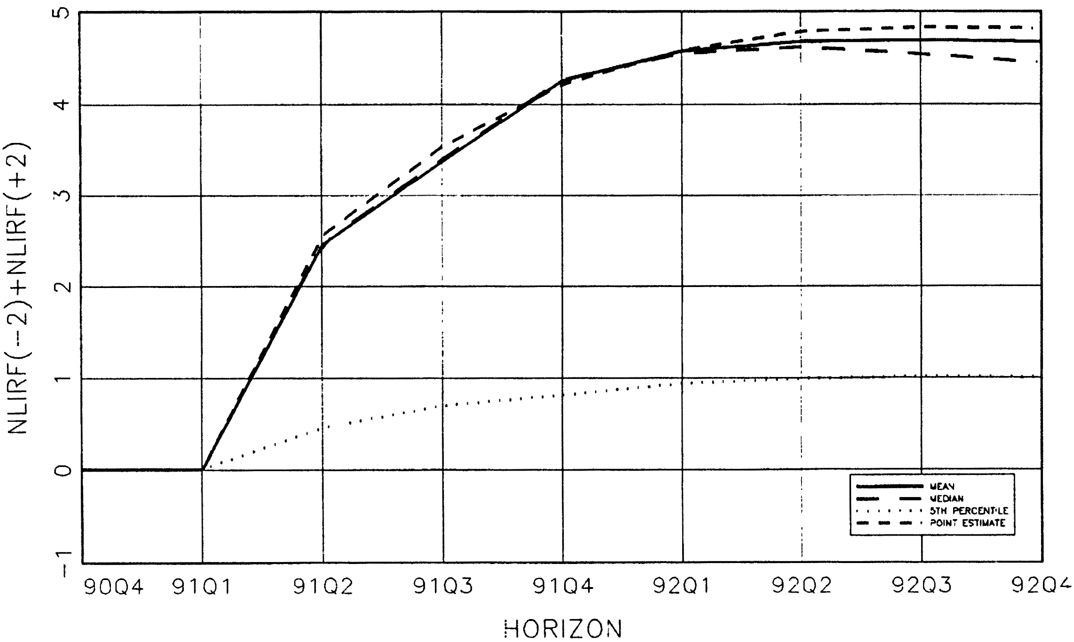

Figure 7. Asymmetry in nonlinear impulse response 1990Q4, shock = 2, −2

[14] An alternative method would be to simulate data from the estimated model then use the simulated data to estimate a new model whose nonlinear impulse response functions could then be calculated. I do not provide a formal justification for the approach used here but it is the analog of Bayesian techniques in the case of producing standard errors for impulse response functions from VARs.
[15] The results are produced by 1000 draws from the variance covariance matrix of the coefficients and 10000 replications for each impulse response function.

median, which are also both very close to response implied by the point estimate. Such findings can be replicated for other histories and shocks.

6. CONCLUSIONS

Macroeconomics has been dominated by the use of linear time-series methods since the Second World War, partly because these were the only statistical techniques available and partly because economic theories are usually tested in a linear form whether exact or as an approximation. In this paper I have shown that linear methods can hide much interesting economic structure in the most examined of all univariate time series, US GNP. It would be convenient if one could point to an economic theory that is consistent with the results I find or, more negatively, an economic theory that is incompatible with the results. However, that is not possible, since nothing I have found is incompatible with previous research that has concentrated on the *covariance properties* of time series. The results of the paper do suggest that, for example, Real Business Cycle simulations should examine more than covariance properties in their model evaluations.

ACKNOWLEDGEMENTS

This is a revised version of parts of Potter (1989). I would like to thank Buz Brock for introducing me to the question of nonlinearity in the economy and guiding my research. I have also received many useful comments and suggestions from Blake LeBaron, Roger Farmer, Bruce Hansen, Hashem Pesaran, James Powell, Julie Press, Ken West, and three anonymous referees. Financial support from the Graduate School of the University of Wisconsin and the Academic State, UCLA, is gratefully acknowledged. All errors are my own responsibility.

REFERENCES

Andrews, D. W. and W. Ploberger (1993), 'Optimal tests when a nuisance parameter is present only under the alternative', *Econometrica* (forthcoming).

Beaudry, P. and G. Koop (1993), 'Do recessions permanently change output?' *Journal of Monetary Economics*, **31**, 149–163.

Brock, W. A. and S. M. Potter (1993), 'Nonlinear time series and macroeconometrics', G. S. Maddala, C. R. Rao and H. D. Vinod (eds), *Handbook of Statistics*, Vol. II, Elsevier, Amsterdam.

Brock, W. A. and C. Sayers (1988), 'Is the business cycle characterized by deterministic chaos?' *Journal of Monetary Economics*, **22**, 71–90.

Brunner, A. D. (1992), 'Conditional asymmetries in real GNP: a semi-parametric approach', *Journal of Business and Economics Statistics*, **10**, 65–72.

Chan, K. S. and H. Tong (1985), 'On estimating thresholds in autoregressive models', *Journal of Time Series Analysis*, **7**, 179–190.

Davies, R. B. (1977), 'Hypothesis testing when a nuisance parameter is present only under the alternative', *Biometrika*, **64**, 247–254.

Diebold, F. and G. Rudebusch (1991), 'Have postwar economic fluctuations been stabilized?' mimeo, University of Pennsylvania.

Gallant, A. R., P. Rossi and G. Tauchen (1993), 'Nonlinear dynamic structures', *Econometrica*, **61**, 871–908.

Gordon, R. J. (1986), *American Business Cycle: Continuity and Change*, National Bureau of Economic Research and University of Chicago Press, Chicago.

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica*, **57**, 357–384.

Hamilton, J. D. (1990), 'Analysis of time series subject to changes in regime', *Journal of Econometrics*, **45**, 39–70.

Hansen, B. E. (1992), 'Testing the Markov trend model', *Journal of Applied Econometrics*, **7**, S61–S82.
Hansen, B. E. (1993), 'Influence when a nuisance parameter is not identified under the null hypothesis', mimeo, Department of Economics, University of Rochester.
Hansen, B. E. (1994), 'Erratum: the likelihood ratio test: under nonstandard conditions: testing the Markov switching model of GNP', mimeo, Department of Economics, Boston College.
Leamer, E. E. (1978), *Specification Searches*, John Wiley, New York.
Mittnik, S. (1991), 'State space modeling of nonlinear time series', paper presented at the UCLA and *Journal of Applied Econometrics* Conference on Nonlinear Dynamics and Econometrics.
Neftci, S. N. (1984), 'Are economic time series asymmetric over the business cycle?' *Journal of Political Economy*, **92**, 307–328.
Pesaran, B. and M. H. Pesaran (1993), 'A simulation approach to the problem of computing Cox's statistic for non-nested models', *Journal of Econometrics*, **00**, 000–000.
Pesaran, M. H. and S. M. Potter (1991), 'Application of Non-Nested Hypothesis Testing to Nonlinear Time Series', paper presented at winter 1992 meetings of Econometric Society, New Orleans.
Potter, S. M. (1989), 'A Nonlinear Approach to U.S. GNP',mimeo, University of Wisconsin-Madison, Department of Economics.
Potter, S. M. (1990), *Nonlinear Time Series and Economic Fluctuations*, Unpublished PhD thesis, University of Wisconsin.
Potter, S. M. (1991), 'Nonlinear impulse response functions', mimeo, UCLA.
Potter, S. M. (1994), 'Asymmetric economic propagation mechanisms', in W. Semmler (ed.), *Business Cycles: Theory and Empirical Methods*, Kluwer Academic Publishers, Norwell, MA, forthcoming.
Romer, C. D. (1986), 'Is the stabilization of the postwar economy a figment of the data?' *American Economic Review*, **76**, 314–334.
Sichel, D. (1988), *Essays in Business Cycle Asymmetry*, Unpublished PhD thesis, Department of Economics, Princeton University.
Sichel, D. (1992), 'Business cycle duration dependence: a parametric approach', *Review of Economics and Statistics*, forthcoming.
Sichel, D. (1994), 'Inventories and the three phases of the business cycle', *Journal of Business Economic Statistics*, **12**, 269–278.
Teräsvirta, T. and H. Anderson (1992), 'Modelling nonlinearities in business cycles using smooth transition autoregressive models', *Journal of Applied Econometrics*, **7**, S119–S136.
Tiao, G. and R. Tsay (1991), 'Some advances in nonlinear and adaptive modeling in time series analysis', University of Chicago Graduate School of Business Statistics Research Centre Report #118.
Tong, H. (1983), *Threshold Models in Non-linear Time Series Analysis*, Springer-Verlag, New York.
Tong, H. (1990), *Non-Linear Time Series: A Dynamical Systems Approach*, Oxford University Press, Oxford.
Tsay, R. S. (1986), 'Nonlinearity tests for time series', *Biometrika*, **73**, 461–466.
Tsay, R. S. (1989), 'Testing and modeling threshold autoregressive processes', *Journal of the American Statistical Association*, **84**, 231–240.
Tsay, R. S. (1991), 'Detecting and modeling nonlinearity in univariate time series', *Statistica Sinica*, **1**, 431–452.