A General Dependence Test and Applications[*]

David Johnson[**]
Robert McClelland[***]

Division of Price and Index Number Research
Bureau of Labor Statistics
2 Massachusetts Ave., NE, Room 3105
Washington, D.C. 20212

[**](202) 606-6579 Ext 599; JohnsoDS@oplc.psb.bls.gov
[***](202) 606-6579 Ext 607; McClelland_R@bls.gov

Abstract:

We describe a test, based on the correlation integral, for the independence of a variable and a vector that can be used to detect model misspecification in serially dependent data. In Monte Carlo simulations this test performs nearly as well or better than the BDS test in univariate time series and complements the BDS test in distributed lag models. Finally, we apply our test to detect misspecification in models of U.S. unemployment data.

Recent work on nonlinear dynamics has focused on several concepts, including the use of the correlation integral in nonparametric tests of misspecification for time series models.[1] By detecting serial dependence in a time series they allow researchers to check for dependence without specifying an alternative model. In this paper we use the correlation integral to describe a test for the independence of a variable and a vector that can be used to detect model misspecification in both independently and identically distributed (iid) and serially dependent data.

To clarify how our test fits into the family of nonparametric specification tests, we can distinguish members of the family by their null hypotheses. For example, Wooldridge (1992) and Bradley and McClelland (1994) propose linearity tests in iid data that are consistent against almost all alternatives, the former using a sieve estimator and the latter a kernel estimator. In addition, Lee, White and Granger (1993) propose a linearity test for use on serially dependent data that is based upon the test of Bierens (1990).

Although requiring iid data, the Bierens test has a more general null hypothesis than linearity: it is consistent against almost all misspecifications of the first conditional moment of a nonlinear least squares model.[2] Similarly, Bradley and McClelland (1993) and Lewbel (1993) describe conditional moment tests for use with iid data without requiring linearity and de Jong (1992) extends the Bierens approach to time series data.

While the above tests work on individual moments of a distribution, other tests have even broader null hypotheses by checking assumptions about the whole distribution of the residual or the dependent variable. This is useful for

---

[1] For example, see Brock, Dechert and Schienkman [1987], Baek and Brock [1988] and [1992], Brock, Hsieh and LeBaron [1991].

[2] For a more detailed discussion of the Bierens test and some potential problems, see Bradley and McClelland (1993).

procedures that require strong distributional assumptions, such as the adaptive estimation method of Manksi (1984) or some quasi-maximum likelihood methods, or more generally when the researcher wishes to test distributional assumptions or a priori knowledge other than simple conditional moment restrictions. Examples of this set of tests include the many types of Pearson χ^2 independence tests, such as the tests for iid data in Andrews (1988). Tests for serial independence of a variable or vector, such as Robinson (1991) and Skaug and Tjøstheim (1993) also fall into this class.

Because these tests are nonparametric, assumptions about the distributional form of the variable are not necessary. Both the Robinson test and the test of Skaug and Tjøstheim are consistent against all alternatives when a single lag is chosen. However, the consistency properties against dependence in the more useful case of higher lags is unknown. It is also unclear if either of these tests is appropriate for use upon residuals from an estimated model. Finally, it should be noted that the Skaug and Tjøstheim test does not have a standard normal or χ^2 distribution, which can make comparisons with other tests more difficult.[3]

Another example of a time series independence test is the BDS test (see Brock, Hsieh and LeBaron (1991)). This test checks for serial dependence of a random variable by estimating the correlation integral of the series, which measures spatial correlation in phase space with the frequency with which pairs of observations are close. The BDS test is asymptotically normally distributed under the null hypothesis of independence and De Lima (1996) shows that correlation integral-based tests make relatively weak assumptions about the

---

[3] The Skaug and Tjøstheim test has the distribution of the test of Blum, et.al. (1961), which is defined only by its characteristic function. Although the Robinson test does have a standard distribution, Robinson notes that a parameter affecting the power of the test may also distort its size.

necessary moment conditions. By using a nuisance parameter theorem, Brock, Hsieh and LeBaron also show that the test can be used as a specification test by checking for the dependence of the residuals from a regression. However, it has been pointed out that consistent estimation of ordinary least squares coefficients does not require that residuals are iid, only that they are a martingale difference sequence (MDS).

In this paper, we use the correlation integral to test for the dependence of the residual upon the set of explanatory variables. By checking the dependence of the residuals upon the regressors our test resembles the test of Andrews, using the correlation integral instead of the standard χ^2 test. Unlike the tests of Andrews (1988), Bierens (1990), Wooldridge (1992), and Bradley and McClelland (1993), our test can be used with serially dependent data. As in the BDS test, our test is normally distributed and does not require bootstrapping of the distribution. It also does not use a pseudo-random number generator, as in the tests of Bierens (1990), de Jong (1992), Lee, White and Granger (1993), and Bradley and McClelland (1994).

Because we use correlation integrals, our test shares the moment condition properties of the BDS test. Unlike the BDS test however, our test explicitly allows for serial dependence in the residuals, so that our test is insensitive to dependence among residuals that do not affect the regressors. We more directly test for correct specification since we check for dependence of the estimated residual upon the regressors, rather than the dependence of the residuals upon the regressors through the past residuals. This means that in univariate time series the GD test may have greater finite sample power against functional form misspecification. The BDS test, however, should have better power against dependence in the residuals upon their own histories (e.g., autoregressive conditional heteroskedasticity).

In the case where regressors do not include lagged values of the dependent variable, the BDS and GD tests are essentially complementary. While the BDS test checks for time dependence in the estimated residuals (which can come from time dependence in the true residuals), the GD test checks for contemporaneous dependence between the regressors and the estimated residuals. A rejection of the null hypothesis by the BDS test but not by the GD test indicates time dependence in the actual residuals but a correct functional form, while acceptance of the null by the BDS test but rejection by the GD test indicates functional form misspecification but no dependence in the residuals.

To compare the size and power properties of these two tests we use Monte Carlo simulations under the null and several alternative hypotheses. We show that for some univariate time series models the GD test performs nearly as well or better than the BDS test. We also use Monte Carlo simulations to illustrate the complementarity of the BDS and GD tests in a distributed lag model.

As an example we use both tests to examine various models of U.S. unemployment data. While we find evidence of nonlinearities and asymmetries, none of the existing models appears to adequately fit the data. However, using the null of a four-regime threshold autoregressive model, the GD test rejects the null while the BDS fails to reject. This illustrates the complementarity of the tests and indicates that dependence in the residuals is likely to be less of a concern than the issue of proper functional form specification in each regime.

The remainder of the paper is organized as follows. The next section describes our test and compares and contrasts it with the BDS test. Section two presents the results of Monte Carlo simulations. Section three applies our test to detect misspecifications in models of unemployment rates and the final section concludes.

I The GD Test for Independence

Our test can be applied to general nonlinear models of the form

$$y_t = G(x_t, \beta) + e_t, \tag{1.}$$

where y_t is an observed random variable produced by the data generating process $G(\cdot, \cdot)$ with a K-dimensional vector of observed explanatory variables x_t and parameter vector β, and an unobserved disturbance e_t distributed independently of x_t. We assume that x_t and e_t satisfy the following two conditions: (1) $\{(x_t, e_t)\}$ is an absolutely regular stationary sequence with mixing coefficients[4] $\alpha(t)$ such that, for some $\lambda < \frac{1}{2}$ and $\delta > 0$, $\alpha(t)^{\delta/(2+\delta)} = O(t^{(-2+\lambda)})$ and (2) $\{(x_t, e_t)\}$ has a smooth joint distribution with a bounded density. Note that assumption (1) excludes unit-root processes but allows Martingdale difference sequences in $\{x_t\}$ and $\{e_t\}$. Under the assumption of independence, the regressors x_t hold no information about the disturbance e_t, so that the conditional distribution of e_t given x_t is equal to the marginal distribution of e_t.

Equation (1.) can be estimated with

$$y_t = G(x_t, \beta_T) + \hat{e}_t, \quad t = 1,\ldots,T \tag{2.}$$

where $\hat{e}_t$ is the estimated residual and β_T is a consistent estimator of β. Using the results of Brock, Hsieh and LeBaron (1991) and de Lima (1996), we can test the independence of x_t and e_t by examining the estimated residuals $\hat{e}_t$. Combining equations (1.) and (2.), the estimated residuals can be described as follows:

$$\hat{e}_t = G(x_t, \beta) - G(x_t, \beta_T) + e_t. \tag{3.}$$

---

[4] Denker and Keller (1983) define as $\alpha(t) = \sup_j E[\sup\{ | P(A \mid \mathscr{F}_1^j) - P(A) | : A \in \mathscr{F}_{j+t}^\infty \}]$ such that $\mathscr{F}_i^j$ is the σ-algebra generated by $\{(x_t, u_t) : 1 \le i \le t \le j \le \infty\}$. A sequence is absolutely regular if the mixing coefficients converge to zero.

If β_T is a consistent estimator of β and we have correctly specified the data generating process, then in the limit $\hat{e}_t$ is equal to e_t and hence independent of x_t. A test of the independence of $\hat{e}_t$ and x_t is then a test of the correct specification.

In general, e_t is independent of x_t if the joint probability of some event is equal to the product of the marginal probabilities of that event. For our test, we note that if x_t and e_t are independent then

$$\Pr[\|(x_t, e_t) - (x_s, e_s)\| < \varepsilon] = \Pr[\|x_t - x_s\| < \varepsilon] \cdot \Pr[|e_t - e_s| < \varepsilon] \quad \text{for all } \varepsilon, \qquad (4.)$$

where $\|\cdot\|$ is the sup norm.[5] This means that the probability that the pairs, (x_t, e_t) and (x_s, e_s), are close is equal to the product of the probability that x_t and x_s are close and the probability that e_t and e_s are close.

These probabilities can be given in terms of correlation integrals, i.e., $C((x,e), \varepsilon) \equiv \Pr[\|(x_t, e_t) - (x_s, e_s)\| < \varepsilon] = E I((x_t, e_t) - (x_s, e_s), \varepsilon)$, where $I(\cdot)$ is an indicator function given by:

$$I((x_n, e_n) - (x_m, e_m), \varepsilon) = \begin{cases} 1 & \text{if } \|(x_n, e_n) - (x_m, e_m)\| < \varepsilon \\ 0 & \text{otherwise} \end{cases}. \qquad (5.)$$

Hence, independence implies that $C((x,e), \varepsilon) = C(x, \varepsilon) \cdot C(e, \varepsilon)$.

To estimate these correlation integrals we use the theory of U-statistics. An estimator of $C((x,e), \varepsilon)$ is $C_T((x,e), \varepsilon)$, which is a U-statistic of the following form:

$$C_T((x,e), \varepsilon) = \frac{2}{T(T-1)} \sum_{1 \leq t < s \leq T} I(e_t - e_s, \varepsilon) \cdot \left\{ \prod_{k=1}^{K} I(x_t^k - x_s^k, \varepsilon) \right\} \qquad (6.)$$

The correlation integral estimator for the regressors, $C_T(x, \varepsilon)$, is given by:

$$C_T(x, \varepsilon) = \frac{2}{T(T-1)} \sum_{1 \leq t < s \leq T} \left\{ \prod_{k=1}^{K} I(x_t^k - x_s^k, \varepsilon) \right\} \qquad (7.)$$

The correlation integral estimator for the disturbances, $C_T(e, \varepsilon)$, is defined similarly.

---

[5] For the sup norm, $\|(x_t, u_t) - (x_s, u_s)\| < \varepsilon$ if and only if $|u_t - u_s| < \varepsilon$ and $|x_t^k - x_s^k| < \varepsilon$ for k: $1 \leq k \leq K$.

Using the results of Brock, Hsieh and LeBaron (1991) and Denker and Keller (1986), it can be shown that if our assumptions described above hold then for a fixed ε

$$\sqrt{T}[C_T((x,e),\varepsilon) - C_T(x,\varepsilon)C_T(e,\varepsilon)] \Rightarrow \mathcal{N}(0,\sigma_C^2), \tag{8.}$$

where

$$\sigma_C^2 = 4\left[K_0(x,\varepsilon)K_0(e,\varepsilon) - K_0(x,\varepsilon)C(e,\varepsilon)^2 - K_0(e,\varepsilon)C(x,\varepsilon)^2 + C(e,\varepsilon)^2C(x,\varepsilon)^2 + \right.$$

$$\left. 2 \sum_{j \geq 1} \left(K_j(x,\varepsilon)K_j(e,\varepsilon) - K_j(x,\varepsilon)C(e,\varepsilon)^2 - K_j(e,\varepsilon)C(x,\varepsilon)^2 + C(e,\varepsilon)^2C(x,\varepsilon)^2\right)\right], \tag{9.}$$

$$K_0(x,\varepsilon) = K_0(x_i,\varepsilon) = E\left[E[I(x_t - x_s,\varepsilon) \mid x_t = x_i]^2\right], \tag{10.}$$

$$K_j(x,\varepsilon) = K_j(x_i,\varepsilon) = E\left[E[I(x_t - x_s,\varepsilon) \mid x_t = x_i]E[I(x_t - x_s,\varepsilon) \mid x_s = x_{j+i}]\right], \tag{11.}$$

and

$$C(x,\varepsilon) = E[I(x_t - x_s,\varepsilon)] \quad \text{(similarly for } K(e,\varepsilon) \text{ and } C(e,\varepsilon)).[6]$$

Hence, the GD test statistic is given by:

$$\sqrt{T}\,\frac{W_T(x,e,\varepsilon)}{s_C}, \tag{19.}$$

where $W_T(x,e,\varepsilon) = C_T(x,e,\varepsilon) - C_T(x,\varepsilon) \cdot C_T(e,\varepsilon)$ and s_C^2 is calculated by substituting the sample analogs, $C_T(\cdot,\varepsilon)$, $K_{0_T}(\cdot,\varepsilon)$ and $K_{j_T}(\cdot,\varepsilon)$, for $C(\cdot,\varepsilon)$, $K_0(\cdot,\varepsilon)$ and $K_j(\cdot,\varepsilon)$ in equation (9.).

Equation (8.) states that the distribution of our statistic converges to a normal distribution when using the disturbances, e_t. To apply our tests to residuals $\hat{e}_t$, we must show that under the appropriate conditions the GD test evaluated at the residuals converge to the statistics evaluated at the disturbances, i.e., that

$$\sqrt{T}\left[\frac{W_T(x,\hat{e},\varepsilon)}{s(\hat{e})} - \frac{W_T(x,e,\varepsilon)}{s(e)}\right] \xrightarrow{p} 0 \tag{20.}$$

---

[6] This follows from proposition 1 in Johnson and McClelland [1996], where the variance σ_C^2 is calculated using the results of Denker and Keller [1986] to allow for the vector-pair, $\{(x_t,u_t)\}$, to follow the mixing condition. A formal proof is available from the authors upon request.

This type of convergence is shown in nuisance parameter theorems such as those in Baek and Brock (1992) and de Lima (1996). De Lima (1996) shows a nuisance parameter theorem for a family of BDS tests applied to univariate time series models. A nuisance parameter theorem for the GD test would be a simple application of de Lima's results if our regressors were the K lags of a univariate time series, that is, $(x_t^1,...,x_t^K) = (y_{t-1},...,y_{t-K})$ and the disturbances were iid. In this case, we can apply de Lima's results by requiring the series $\{\,y_t\,\}$ to satisfy the conditions in his theorems.

Nuisance Parameter Theorem: Suppose that the data generating process is given by $y_t = G(x_t,\beta) + e_t$, and the following conditions are satisfied:

i)   The regressors and disturbances are contemporaneously independent and the vector $\{x_t,e_t\}$ is a strong mixing processes (absolutely continuous processes are strong mixing) with mixing coefficients that satisfy the summability condition $\sum\alpha(t)^{1/2} < \infty$. (i.e., the processes are strong mixing processes of order two). $G(\cdot)$ is a measurable function of x_t.

ii)   For all d there is some constant C_1 such that

$$\mathop{\mathrm{E}}\ \sup_{\{\,\beta_T:\,\|\beta_T-\beta\|\,<\,d\}}\ |\ I(\hat{e}_t - \hat{e}_s,\varepsilon) - I(e_t - e_s,\varepsilon)\ |\ < C_1 d.$$

iii) $\sqrt{T}(\beta_T-\beta) = O(1)$.

iv)   The series $\{(x_t,e_t)\}$ has a joint distribution, $F(\cdot,\cdot)$ that is continuously differentiable with a bounded density.

Then the following convergences hold:

$$\sqrt{T}(W_T(x,\hat{e},\varepsilon) - W_T(x,e,\varepsilon)) \xrightarrow{P} 0 \ \text{ and } \ (s_C^2(\hat{e}) - s_C^2(e)) \xrightarrow{P} 0.$$

The definition of $W(\cdot,\cdot,\cdot)$ and the independence of x_t, and e_t imply that both of these convergences are satisfied if $\sqrt{T}(C_T(\hat{e},\varepsilon) - C_T(e,\varepsilon)) \xrightarrow{P} 0$. We can now apply

Theorem 2.1 in de Lima (1996) to show this convergence.[7]   Here, (i) takes the

place of assumption (A) in de Lima (1996).   For a univariate time series, our

assumption (i) is weaker than assumption (A).   De Lima assumes that the series

$\{y_t\}$ is strong mixing of order two, which implies that $\{x_t\}$, where x_t is the k lags of

y_t (i.e., $(x_t^1,...,x_t^K) = (y_{t-1},...,y_{t-K}))$, is strong mixing of order two. De Lima's additional

assumption that the disturbances are iid along with the fact that $\{x_t\}$ is a mixing

process implies that $\{x_t,e_t\}$ is also strong mixing of order two. De Lima uses

assumption (A) to show that the residuals and the kernal are strong mixing

processes of order two.

In addition, assumption (ii) is assumption (B) in de Lima (which given our

indicator kernel also implies de Lima's assumption (C)), (iii) is assumption (D) and

(iv) is assumption (E).

While there is a nuisance parameter theorem for the BDS test, it differs

from the GD test because it examines the independence among the residuals.

Using the 2-histories of the residuals, $e^2 \equiv \{(e_t, e_{t+1})\}$, if the series is iid then the

correlation integral for 2-histories is equal to the product of two one-history

correlation integrals, i.e.,

$$C(e^2,\varepsilon) = C(e,\varepsilon)\cdot C(e,\varepsilon) \qquad\qquad (21.)$$

so that the BDS statistic is given by:

$$\sqrt{T}\ \frac{C_T(e^2,\varepsilon) - C_T(e^1,\varepsilon)^2}{s_B}, \qquad\qquad (22.)$$

where s_B^2 is the sample variance given in Brock, Hsieh and LeBaron (1991) and

the correlation integral for the 2-histories is:

$$C_T(\hat{e}^2,\varepsilon) = \frac{2}{(T\text{-}1)(T\text{-}2)} \sum_{1\leq i<j\leq T-1} I((\hat{e}_i,\hat{e}_{i+1})\text{-}(\hat{e}_j,\hat{e}_{j+1}),\varepsilon). \qquad\qquad (23.)$$

---

[7] A formal proof is available from the authors upon request.

We can compare the BDS and GD test in a univariate time series model in which the lagged value is the only regressor, i.e.,

$$y_t = G(y_{t-1}, b) + e_t. \tag{24.}$$

Using (3), misspecification implies that

$$\hat{e}_t = G(y_{t-1}, b) - G(G(y_{t-2}, \beta_T) + \hat{e}_{t-1}, \beta_T) + e_t. \tag{25.}$$

The BDS test will detect dependence of $\hat{e}_t$ on $\hat{e}_{t-1}$ because $\hat{e}_{t-1}$ affects $\hat{e}_t$ through $G(\cdot, \cdot)$. Alternatively, the GD test examines the dependence between $\hat{e}_t$ and y_{t-1} directly.

II Monte Carlo Simulations

In this section, we describe the relative properties of the GD test and the BDS test in univariate and distributed lag time series models. All tests use 5,000 iterations, errors with a standard normal distribution and a window width, ε, set to one standard deviation. We apply the BDS test to the two-histories of the residuals. When calculating the GD test we follow Hiemstra and Jones (1994) and include a weight, w(j), on the crossproduct terms in (9.) such that

$$w(j) = 1 - j/(trunc(T^{.25})+1). \tag{26.}$$

This weight determines the rate at which the terms in the summation in (9.) increase with the sample size.[8]

In figure one we show the results of simulations when the null hypothesis is true.[9] In this case we correctly specified a AR model in which the lags were selected using the Bayesian information criteria. The horizontal axis depicts critical points for a normal distribution, while the vertical axis indexes the actual

---

[8] The second component of (9) becomes:

$$2 \sum_{j=1}^{trunc(T^{.25})} w(j)\left(K_j(x,\varepsilon)K_j(u,\varepsilon) - K_j(x,\varepsilon)C(u,\varepsilon)^2 - K_j(u,\varepsilon)C(x,\varepsilon)^2 + C(u,\varepsilon)^2C(x,\varepsilon)^2\right)\Big].$$

[9] The authors would like to thank the editor for suggesting this graphical method of presenting test statistic simulation results. For more information, see Davidson and Mackinnon (1994).

percent of simulations that exceed that critical point. The curve of a perfectly-sized statistic will follow a line formed of points where the actual percent of simulations that exceeds the critical point equals the critical points , while the curve of a statistic that rejects too often will be above that line..

As figure one shows, the BDS test and the GD test for regressors formed with one and four lags from the series reject too frequently with 50 observation data sets, with the rates for the BDS test everywhere above the GD test lines.[10] Because of the use of the sup norm in the correlation integral, the size of the GD test changes with the number of lags. However, this does not occur if we aggregate the regressors, such as by using the conditional expected value of the dependent variable in place of the regressors. Figure two shows results of simulations with 250 observations in each data set. As expected, both the GD test and the BDS test are converging to the proper size.

Although it is impossible to draw any global conclusions, it is also useful to examine the power of the GD test and the BDS test against the models in table two. Models one, two, and three are inspired by those in Lee, White and Granger (1993). Model four is an AR(1) model with ARCH disturbances. This model allows us to compare the BDS and GD tests when the disturbances are an MDS process. Finally, model five is from Rothman (1992b). In each case, the model is misspecified by an AR model with the number of lags chosen by the Bayesian information criteria. The horizontal axis in the figures now lists the rejection rates generated by the simulations described in figures one and two. More powerful tests will be reflected in curves that most closely approximate an inverted "L".

Figures three through seven show the results from these simulations. In figures three and four it appears that the GD test and the BDS test are comparable against nonlinear AR and nonlinear MA alternatives. In figure five

---

[10] The lines for two and three lags lie everywhere between the lines for the one and four lags.

the GD test seems to outperform the BDS test, to the point where the GD test rejects on data sets with 50 observations at approximately the same rate as the BDS test with 250 observations. Figure six shows that, for data sets with 250 observations, the BDS test detects the dependence in the residuals, while the GD test shows almost no power. Finally, figure seven illustrates that the GD test appears to strongly outperform the BDS test against sign AR models with 250 observations and is approximately equal in power (which is to say no power) with 50 observations.[11]

To demonstrate how the BDS test and the GD test complement each other, table two presents a distributed lag model with two variables that are independent of each other but time dependent. The true model is given by these two variables minus their product and a disturbance term that is either iid (model six) or autocorrelated (model seven). In the estimation of model six the product term is omitted, while in model seven the autocorrelation in the error term is not modeled. As figures eight and nine show, when the cross-product term is omitted from estimation only the GD test detects the misspecification, while failure to model the error term properly is detected only by the BDS test.

III An Application to Unemployment Rates

In this section we use the BDS test and the GD test to examine various models of the U.S. unemployment rate. This series provides a good application because of the current interest in applying models of nonlinearities and asymmetries to macroeconomic data. For example, Brock and Sayers (1988), Frank and Stengos (1988), Ham and Sayers (1990) and Rothman (1992a)

---

[11] The results for models 1, 2 and 3 can be compared to the results of the neural net test given in Lee, Granger and White (1993). At the 5% critical value they report better power against model 3, while the power of the GD exceeds that of the neural net test in model 1 and the tests have the same power against model 2.

demonstrate this asymmetry by examining nonlinearities in unemployment rates, while Potter (1991), Sichel (1991), Beaudry and Koop (1993) and Pesaran and Potter (1994) use GNP to demonstrate these business cycle asymmetries.

To examine these models, we use aggregate seasonally adjusted monthly unemployment rates from January 1948 to December 1993. Since the published seasonally adjusted unemployment rate series is only reported at one decimal point, we calculate the unemployment rate from the seasonally adjusted levels of the unemployed and labor force. This allows us to retain another significant digit.

Figure 10 shows the unemployment rate data in which the shaded areas are the periods of recession (as calculated by the National Bureau of Economic Research and published in USDC (1994)). This figure illustrates the observation of Keynes (1936) that expansions (where the rate is falling) are longer and slower than contractions (the shaded areas). The figure also illustrates the insight of Sichel (1991) of a third recovery phase in which the rate falls faster than in the expansion phase. Kydland and Prescott (1990) report that Mitchell suggested in the 1920's that there are four phases: prosperity (expansion), crisis (recession), depression (contraction), revival (recovery).

While many suggest the presence of asymmetries in the unemployment data, a simple AR model fits the data fairly well (see also Brock and Sayers (1988)). However, the coefficient on the first lag is close to one, suggesting the presence of a unit root. Applying the test of Dickey and Fuller (1981), we fail to reject the null hypothesis of a unit root. Hence, for the remainder of the analysis the data is transformed by taking the log first differences.

A reasonable next step is to estimate a linear model on the transformed data, as in Ham and Sayers (1990) and Rothman (1992a).   The Akaike Information Criteria (AIC) suggests an AR(10) fits the data fairly well, yet the BDS

and GD tests both reject the null of independence. Similar to Ham and Sayers and Rothman, this suggests that there may be nonlinearities unaccounted for by the AR model.

To model these nonlinearities, many researchers use threshold autoregressive models since this class of models can be viewed as piecewise linear approximations to general nonlinear models. In general, a self-exciting autoregressive (SETAR) model with K regimes, SETAR($K,p_1,...,p_K$), is given by:

$$y_t = \alpha_0^{(k)} + \sum_{i=1}^{p(k)} \alpha_i^{(k)} y_{t-i} + \sigma^{(k)} u_t \quad , \text{ for } \tau_{k-1} < y_{t-d} \le \tau_k, \quad k = 1,...,K. \qquad (27.)$$

In (25.), d represents the delay parameter, τ_k represent the thresholds that define the regimes $(-\infty = \tau_0 < \tau_1 < ... < \tau_K < \tau_{K+1} = +\infty)$ and u_t is an iid disturbance term. This specification allows the order of the $AR(p_k)$ process and the variance of the disturbance term to vary across regimes.

To estimate the parameters we use the procedures of Tong (1983) and conduct a grid search to choose the threshold value, delay parameter and orders of the regimes by minimizing the normalized (the sum of the AIC for each regime) AIC.[12] We allow the delay parameter to vary from one to three and the switching parameter to vary from the 15th to the 85th percentile of the distribution of the series. Finally, we standardize the residuals from each regime before applying our tests.

First, we follow Ham and Sayers (1990) and fit a SETAR model with two regimes to the data. As expected, the regimes are determined by whether the unemployment rate is increasing or decreasing. The first regime consists of the expansionary months in which the rate was decreasing in the previous month $(x_{t-1} \le -.0145)$, while the second regime represents the contractionary periods. In

---

[12] Selecting the parameters by minimizing the mean squared error does not significantly alter the results.

fact, the expansionary regime is longer and slower with an order of 11 and smaller coefficients on the first two lagged terms. Although there is an improvement in the AIC, both the BDS and GD tests still reject the null hypothesis, even though the residuals have been standardized.

These rejections of both an AR and SETAR model contrast with those of Hansen (1994), who provides a test which has a null of an AR(p) and an alternative of a SETAR($2,p_1,p_2$). Using a model similar to Potter's model of U.S. GNP growth rates, Hansen shows that the hypothesis of a single regime (AR(p) model) cannot be rejected. Pesaran and Potter (1994), however, find that when the alternative is a more general SETAR model, the null hypothesis of linearity can be rejected.

To attempt to capture this remaining dependence, we estimate three models that are generalizations of the two-regime SETAR model: an exponential autoregressive (EAR) model (as in Rothman (1992a)), a threshold model similar to that in Beaudry and Koop (1993) and a four-regime SETAR model.

The advantage of the EAR model is that it allows for smooth transitions between regimes. Basically, the EAR(p) model is an AR(p) model with additional terms that depend on the delay parameter and is defined by:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + u_t , \qquad\qquad\qquad (28.)$$

where

$$\phi_i = a_i + \pi_i \exp(-\gamma y^2_{t-1}). \qquad\qquad\qquad (29.)$$

Table 3 shows that both the BDS and GD tests indicate misspecification.

Next we estimate a threshold model similar to that in Beaudry and Koop (1993). Their model is an AR(p) model with an additional depth-of-recession variable. While they apply their model to GNP, we apply their model to unemployment data by measuring the depth of a recession as the difference

between the current unemployment rate and the previous twelve month low. Our application of their model is given by:[13]

$$y_t = \alpha_0 + \sum_{i=1}^{p}\alpha_i y_{t-i} + CDR_{t-1} + u_t, \tag{30.}$$

where

$$CDR_{t-1} = y_{t-1} - \min\{y_{t-12-1},...,y_{t-1}\}. \tag{31.}$$

Again, table 3 shows that the BDS and GD tests suggest misspecification.

Finally, because Pesaran and Potter (1994) find that periods of high expansions and contraction are significantly different than mild expansions and contractions, we estimate a four regime SETAR model. The AIC suggests that this model fits better than the previous models and the switching parameters of -.016, 0, and .038 capture the regimes suggested by Mitchell. The first regime consists of fairly substantial expansions (recoveries), the second consists of mild expansions, the third consists of mild contractions and the fourth consists of recessions. Unlike the previous models, the BDS test fails to reject the null of independence. The GD test, however, rejects the null. This suggests that the problem is not dependence among the residuals, but dependence between the residuals and the regressors. We can show that the acceptance of the four regime model by the BDS is obtained in part because the residuals in each regime are standardized before the test is applied. When the BDS test is applied to the non-standardized residuals, it rejects the null of independence with a value of 3.32.

We can also use the GD test to examine each regime separately. As table 3 show, the GD test for the first three regimes is about the same as for the whole series and is about 1.93 for the last regime. These values coupled with the BDS

---

[13] Pesaran and Potter (1994) show how this model can be viewed as a generalized SETAR model.

test's failure to reject suggests that for this model the specification of the functional form within each regime is more of a concern than dependence in the residual.

IV Conclusion

In this paper we have presented a nonparametric test for the independence of a variable from a vector and described how this test can be used as a specification test in both cross-sectional and time series models. Monte Carlo simulations suggest that the test has power in several models in iid and serially dependent data. Finally, we have used the GD test in conjunction with the BDS test to detect misspecification in models of U.S. unemployment rates.

Figures

Figure 1: Size of Tests with 50 Observations

Figure 2: Size of Tests with 250 Observations

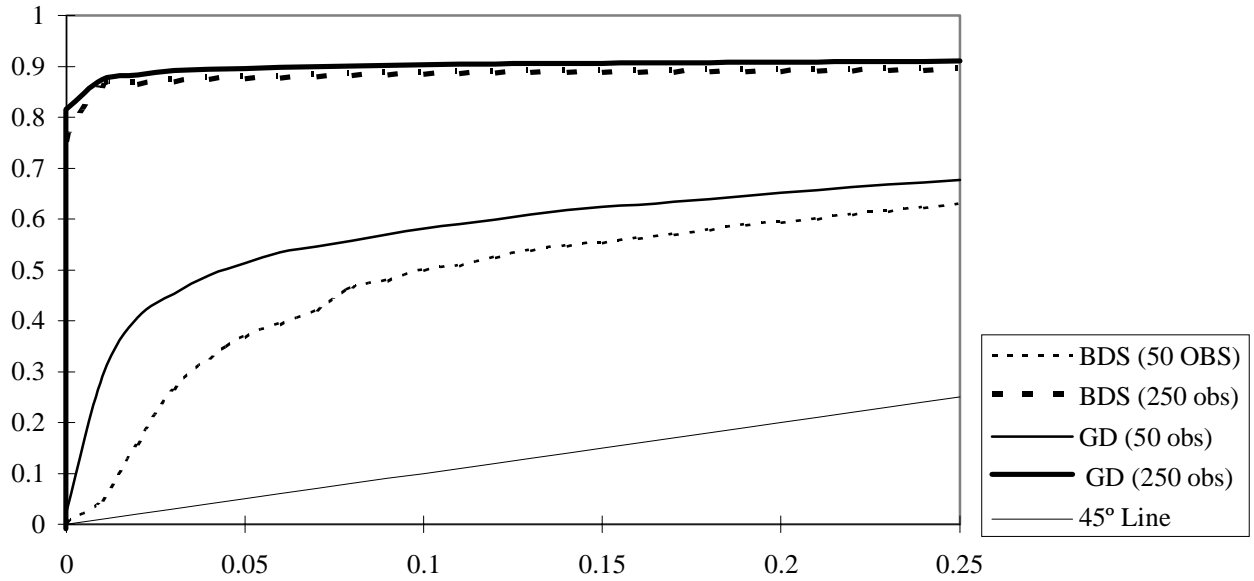Figure 3: Power of Tests Against Model 1 (Nonlinear AR)

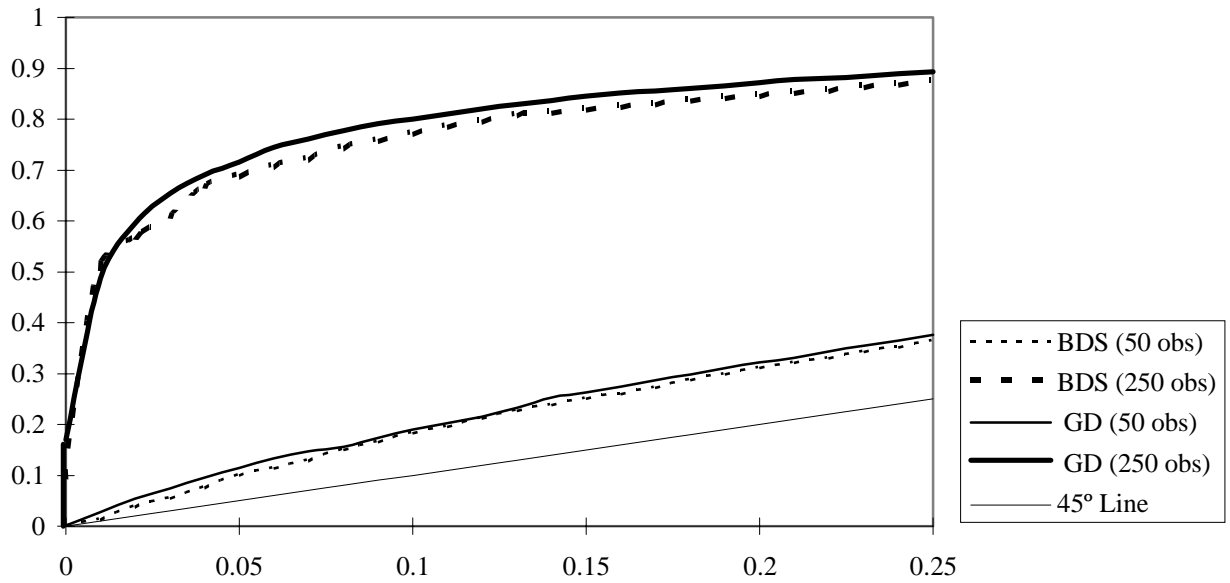Figure 4: Power of Tests Against Model 2 (Nonlinear MA)

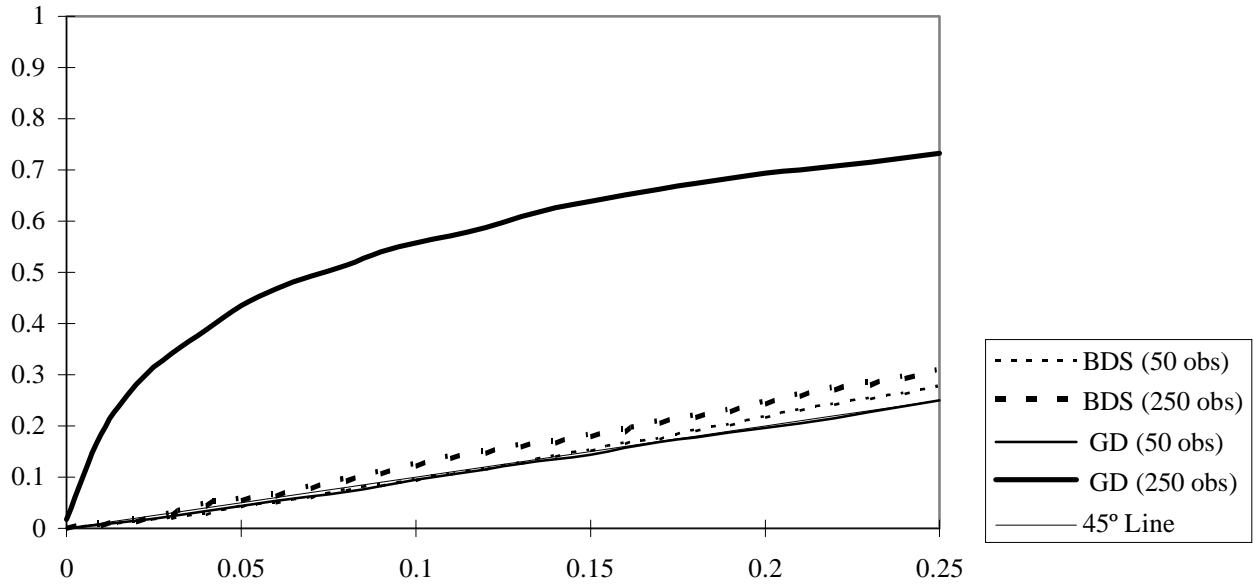## Figure 5: Power of Tests Against Model 3 (Sign AR)

Legend:
- BDS (50 obs)
- BDS (250 obs)
- GD (50 obs)
- GD (250 obs)
- 45º Line

Figure 6: Power of Tests Against Model 4 (AR/ARCH)

Legend:
- BDS (50)
- BDS (250)
- GD (50)
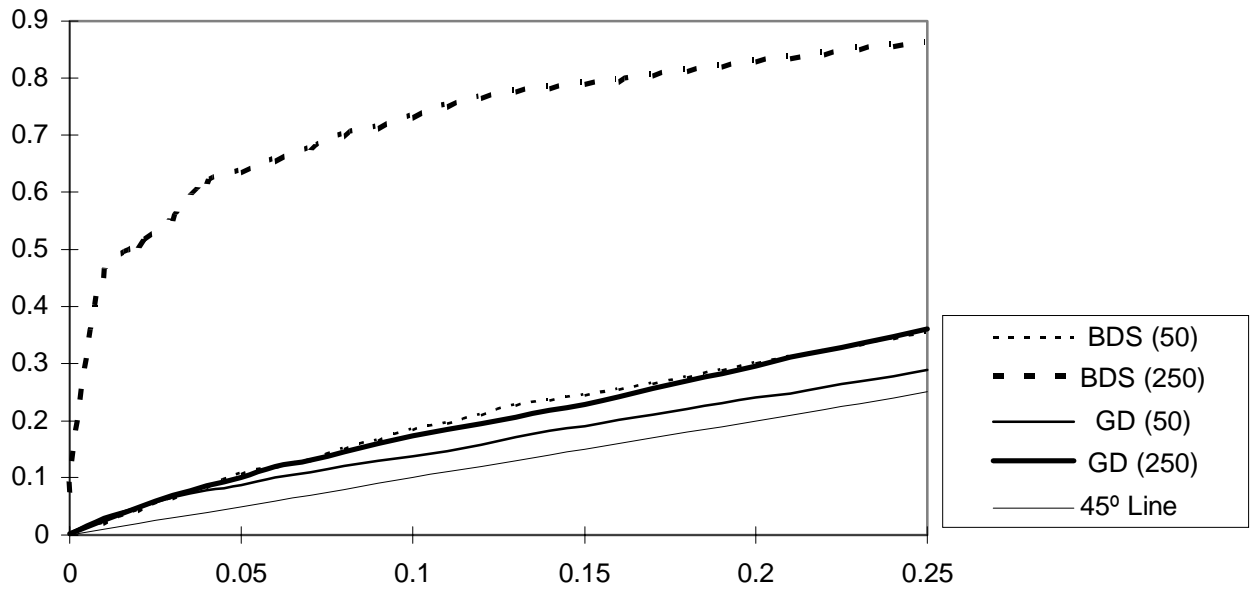- GD (250)
- 45º Line

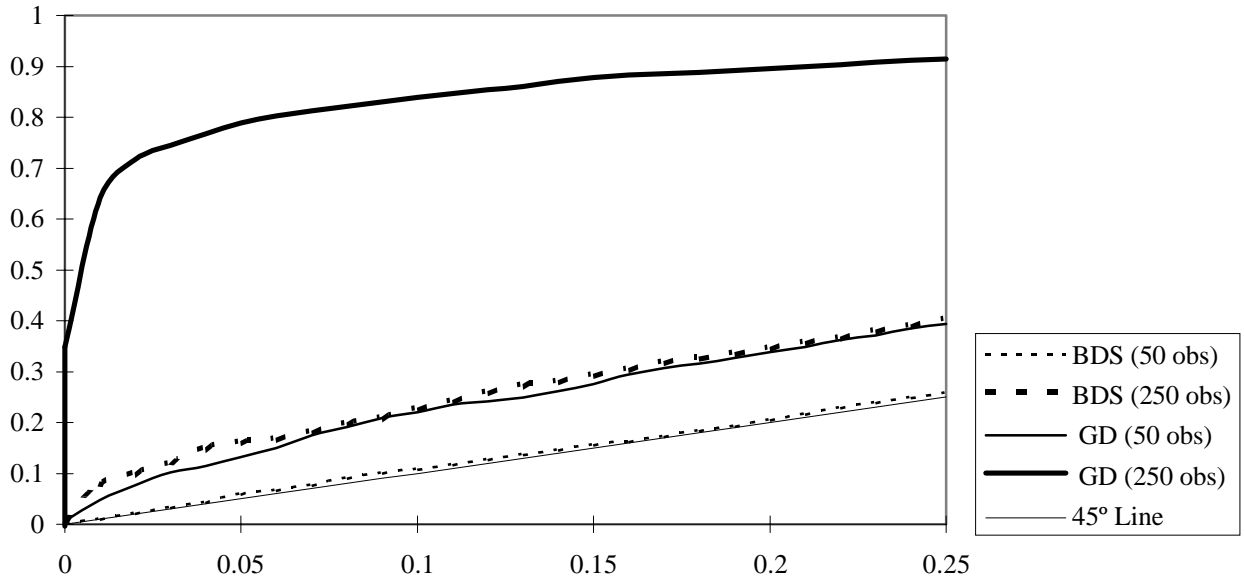Figure 7: Power of Tests Against Model 5 (SETAR)

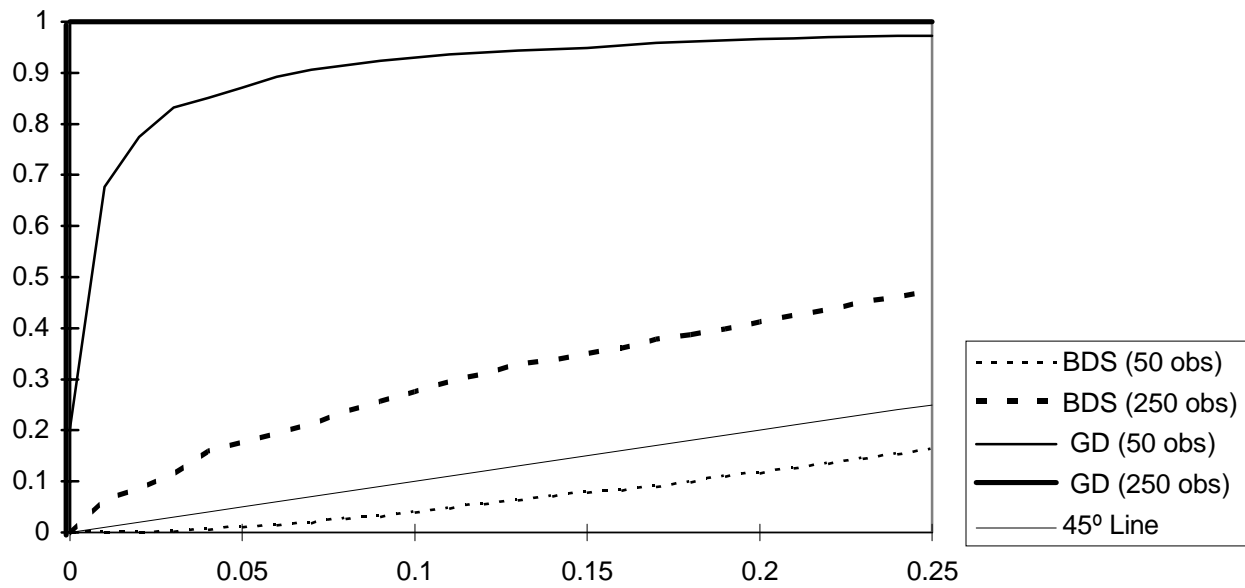Figure 8: Power of Tests Against Distributed Lag with Misspecified Functional

Form

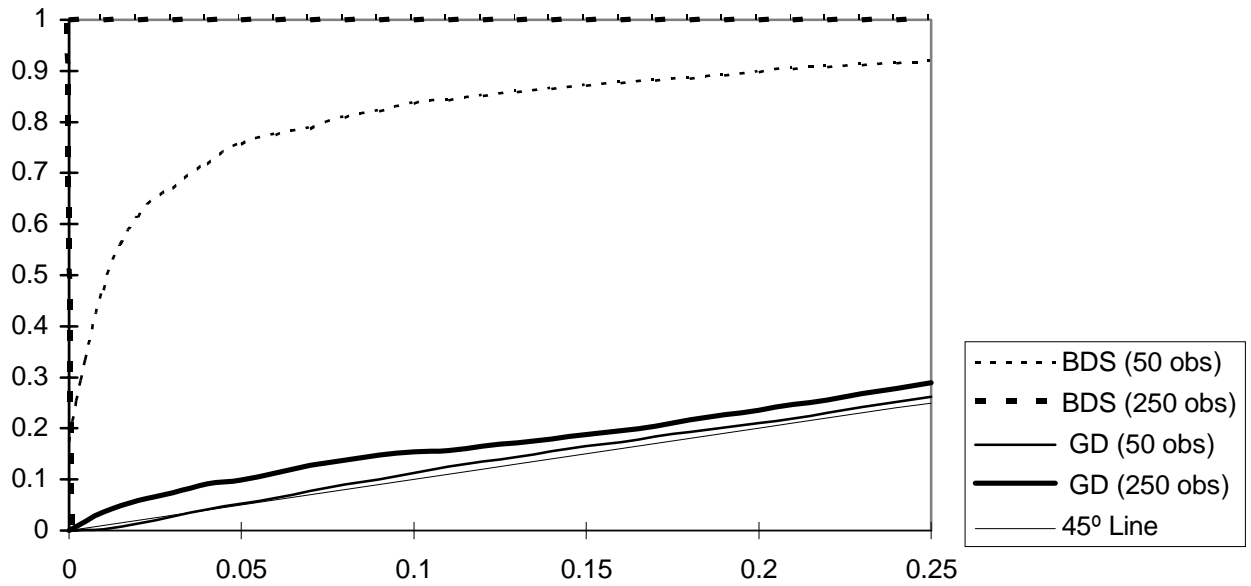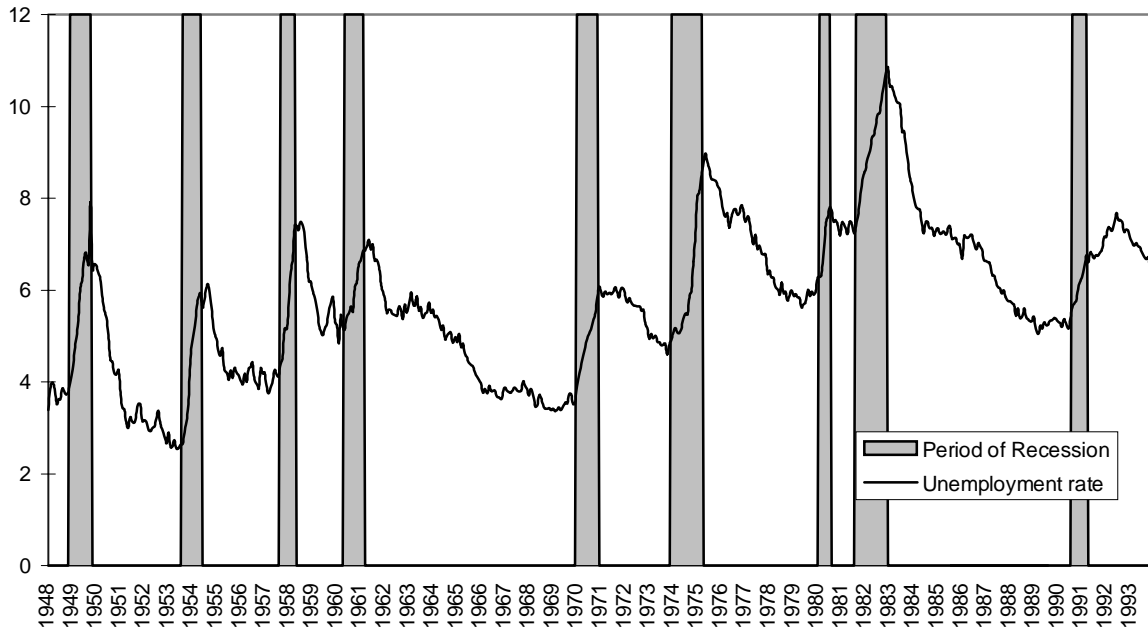**Figure 9: Power of Tests Against Distributed Lag with Misspecified Residual**

Figure 10: Monthly Unemployment Rate 1948-1993 (seasonally adjusted)

Tables

Table 1: Models Used in Time Series Misspecification Testing

Number	True Model	Name of Model
Model 1	$y_t = .7y_{t-1}/(y_{t-2}+2) + u_t$	Nonlinear AR
Model 2	$y_t = .8u_{t-2}u_{t-1} + u_t$	Nonlinear MA
Model 3	$y_t = \begin{cases} -1 + u_t \text{ if } y_{t-1}< 0 \\ 1 + u_t \text{ otherwise} \end{cases}$	Sign AR
Model 4	$y_t = .9y_{t-1}+ u_t, \ u_t \sim N(0, (1+ .25u_{t-1})^2)$	AR/ARCH
Model 5	$y_t = \begin{cases} 0.62 +1.25y_{t-1} -0.43y_{t-2} +u_{1t} \text{ if } y_{t-2}\le 3.25 \\ 2.25 +1.52y_{t-1} -1.24y_{t-2} +u_{2t} \text{ otherwise} \end{cases}$ $u_{1t} \sim N(0,.0381); \ u_{2t} \sim N(0,.0626)$	SETAR

In each case, u is drawn from a standard Normal distribution

Table 2: Models used in Misspecification of Distributed Lag Models

Number	True Model	Model Misspecification
Model 6	$y_t = 2x_t + 1.5z_t - x_tz_t + e_t$ $x_t = .5x_{t-1} - .2x_{t-2} + u_{1,t}$ $z_t = .5z_{t-1} + u_{2,t}$ $e_t = u_{3,t}$	x_tz_t omitted
Model 7	$y_t = 2x_t + 1.5z_t - x_tz_t + e_t$ $x_t = .5x_{t-1} - .2x_{t-2} + u_{1,t}$ $z_t = .5z_{t-1} + u_{2,t}$ $e_t = .7e_{t-1} + u_{4,t}$ e_t normalized to have zero mean and a standard deviation of 1	e_{t-1} omitted

u is drawn from a standard Normal distribution

Table 3: Statistics on Alternative Models of Monthly Unemployment Rates

Model	AIC	σ^2	BDS Test	GD Test
AR(10)	-6.486	.00147	4.692	5.403
SETAR(2;11,7)	-6.566	.00133	5.079	5.357
Regime I: $x_{t-1} \leq -.0145$	-6.775	.00101		2.048
Regime II: $x_{t-1} > -.0145$	-6.459	.00150		5.817
EAR	-6.568	.00134	4.748	4.254
Beaudry/Koop	-6.495	.00144	4.193	5.389
SETAR(4;10,9,2,3)	-6.624	.00129	1.069	2.489
Regime I: $x_{t-1} \leq -.0156$	-6.723	.00107		2.664
Regime II: $-.0156 < x_{t-1} \leq 0$	-7.029	.00076		2.365
Regime III: $0 < x_{t-1} \leq .0384$	-6.556	.00138		2.595
Regime IV: $.0384 \leq x_{t-1}$	-5.904	.00249		1.929

References

Andrews, D. (1988), 'Chi-Square Diagnostic Tests for Econometric Models,' Journal of Econometrics, 37, 135-156.

Baek, E. and W. Brock (1988), 'An Alternative Approach to Specification Error Test,' manuscript, University of Wisconsin.

Baek, E. and W. Brock (1992), 'A Nonparametric Test for Independence of a Multivariate Time Series,' Statistica Sinica, 2, 137-156.

Beaudry, P. and G. Koop (1993), 'Do Recessions Permanently Affect Output?' Journal of Monetary Economics, 31, 149-163.

Bierens, H. (1990) 'A Consistent Conditional Moment Test of Functional Form,' Econometrica, 58, 1443-1458

Blum, J.R., J. Kiefer & M. Rosenblatt (1961), "Distribution Free Tests of Independence Based on the Sample Distribution Function" Annals of Mathematical Statistics, 32, 485- 498

Bradley, R. and R. McClelland (1993), 'An Improved Nonparametric Test for Misspecification of Functional Form,' manuscript.

————— (1994), 'A Kernel Test for Neglected Nonlinearity,' manuscript.

Brock, W., Dechert, W.D., and J. Scheinkman (1987), 'A Test for Independence Based on the Correlation Dimension', University of Wisconsin Working Paper No. 8702.

Brock, W., Hsieh, D. and B. LeBaron (1991), Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence, MIT Press, Cambridge.

Brock, W. and C. Sayers (1988), 'Is the Business Cycle Characterized by Deterministic Chaos?', Journal of Monetary Economics, 22, 71-90.

Craig, S. G., Kohlhase, J.E. and D.H. Papell (1991), 'Chaos Theory and Microeconomics: An Application to Model Specification and Hedonic Estimation', Review of Economics and Statistics, 73, 208-215.

Davidson, R. and J. MacKinnon (1994), 'Graphical Methods for Investigating the Size and Power of Hypothesis Tests,' Queen's University Discussion Paper #903.

de Jong, R. M. (1991), 'The Bierens Test Under Data Dependence,' manuscript

de Lima, P. (1996), 'Nuisance Parameter Free Porperties of Correlation Integral Based Statistics,' forthcoming in Econometric Reviews.

Denker, M. and G. Keller (1983), 'On U-Statistics and v. Mises' Statistics for Weakly Dependent Processes', Zeitscrift fur Wahrscheinlichkeitstheories und Verwandte Gebiete 64, 617-656.

Dickey, D.A. and W.A. Fuller (1981), 'The Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root', Econometrica, 49, 1057-1072.

Frank, M. and T. Stengos (1988), 'Some Evidence Concerning Macroeconomic Chaos', Journal of Monetary Economics, 22, 423-428.

Ham, M. and C. Sayers (1992), 'Testing for Nonlinearities in United States Unemployment by Sector', manuscript, University of Virginia and University of Houston.

Hansen, B. (1994), 'Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis, manuscript.

Hiemstra, C. and J. Jones (1994), 'Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation,' Journal of Finance, 49

Johnson, D. and R. McClelland (1992), 'Nonparametric Test for the Independence of Regressor and Disturbances', Working Paper #231, U.S. Bureau of Labor Statistics

Keynes, J.M. (1936), The General Theory of Employment, Interest and Money, Macmillan Press, London.

Kydland, F. and E. Prescott (1990), 'Business Cycles: Real Facts and a Monetary Myth,' Quarterly Review, Spring 1990, 3-18.

Lee, T., White, H. and C. Granger (1993), 'Testing for Neglected Nonlinearity in Time Series Models', Journal of Econometrics, 56, 269-290.

Lewbel, A. (1993), 'Consistent Tests with Nonparametric Components' manuscript.

Manski, C. (1984), 'Adaptive Estimation of Nonlinear Regression Models,' Econometric Reviews, 3, 145-194.

Pesaran, M.H. and S. Potter (1994), 'A Floor and Ceiling Model of US Output,' manuscript.

Potter, S. (1991), 'A Nonlinear Approach to U.S. GNP', manuscript, Department of Economics, University of Wisconsin.

Robinson, P.M. (1991), 'Consistent Nonparametric Entropy-Based Testing,' Review of Economic Studies, 58, 437-453.

Rothman, P. (1992a), 'Forecasting Asymmetric Unemployment Rates', Department of Economics, University of Delaware Working paper #92-03.

Rothman, P. (1992b), 'The Comparative Power of the TR Test Against Simple Threshold Models', Journal of Applied Econometrics, 7, S187-S195.

Sichel, D. (1991) 'The Three Phases of the Business Cycle: What goes down must come up,' manuscript.

Skaug, J. & D. Tjøstheim (1993) 'A nonparametric test of serial independence based on the empirical distribution function,' Biometrika, 80, 591-602.

Tong, H. (1983) Threshold Models in Nonlinear Time Series Analysis, Springer-Verlag, New York.

U.S. Department of Commerce (1994), Survey of Current Business, 74:4.

Wooldridge, J.M. (1992) 'A Test for Functional Form Against Nonparametric Alternatives', Econometric Theory, 8, 452-475.