Chaos Theory and Microeconomics: An Application to Model Specification and Hedonic Estimation
Author(s): Steven G. Craig, Janet E. Kohlhase, David H. Papell
Source: *The Review of Economics and Statistics*, Vol. 73, No. 2 (May, 1991), pp. 208-215
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/2109510
Accessed: 02/07/2010 15:50

CHAOS THEORY AND MICROECONOMICS: AN APPLICATION TO MODEL SPECIFICATION AND HEDONIC ESTIMATION

Steven G. Craig, Janet E. Kohlhase, and David H. Papell*

Abstract—This paper is the first to apply the theory of deterministic chaos to a microeconomic problem. Previous applications of chaos theory to time series data, while successful in uncovering nonlinearities, have not provided guidelines for resolving uncovered misspecification problems. In contrast, we show that a modified test statistic from chaos theory is an extremely valuable tool in microeconomic model specification because it shows when excluded information is correlated with included information. This test, applied to hedonic estimation of marginal housing prices, is able to distinguish among alternative regression specifications and assists in discovering a parsimonious specification devoid of nonlinear effects.

I. Introduction

INTEREST in the theory of deterministic chaos has recently exploded. Although the theory originated in the hard sciences, applications of chaos theory to economics, as described by Grandmont (1985), Brock (1988), and Baumol and Benhabib (1989), are becoming more common. All of the empirical applications to date, however, have used macroeconomic or financial time series data. Even though researchers have been successful in detecting evidence of nonlinear structure, they have not yet used chaos theory to provide guidelines for model specification to purge nonlinearities from the data. Our application of chaos theory to microeconomic data offers a natural way to attempt to resolve the presence of nonlinear structure through an examination of model specification.

In this paper we present an innovative statistical test, based on the BDS test of Brock, Dechert, and Scheinkman (1987), to detect the presence of nonlinear structure in microeconomic data. The essence of our test is to examine the residuals of a regression for systematic patterns. In a time series context, the data follow a natural order. With cross section data, the choice by which variable to order the data becomes important. We demonstrate that our modified BDS test examines whether there is information omitted from the regression that is correlated with the variable by which the residuals are ordered. Omitted information correlated with included variables can either be truly omitted variables, or can be a nonlinear combination of the information already included in the regression. While there are no systematic methods for finding excluded variables, it is possible to explore nonlinear combinations of included variables as alternative regression specifications.

To illustrate the importance of chaos theory to microeconomic research, we apply the test for nonlinear structure to a hedonic housing price regression. This application is particularly important since model specification is a crucial issue without clear theoretical guidance in the housing literature.[1] We show that the residuals from the standard specification typically exhibit nonlinearities; that is, the residuals are not independently and identically distributed (iid). For some of the cases that we examine, we are able to purge the nonlinearities from the system by examining alternative specifications of the regression model.

There are two separate aspects of specification in applied economics that are highlighted in the hedonic housing literature. One is to find the best overall fit of a regression equation. The second is to determine whether the central coefficient of the model is correctly estimated. While the econometric literature has recently proliferated with new specification tests that address the first of these problems, no test exists that is focused

Received for publication September 1, 1989. Revision accepted for publication August 9, 1990.

* University of Houston and Hunter College and Graduate Center, CUNY; University of Houston and Hunter College and Graduate Center, CUNY; and University of Houston, respectively.

We would like to thank Dee Dechert for unravelling some of the complexities of chaos theory and the BDS test for us. We have also benefited from discussions with Chera Sayers, the comments of two excellent referees, and from seminars at the Dallas Federal Reserve Bank, the University of Houston, Hunter College, and Wayne State University. Partial research support for Craig and Kohlhase was provided by Texas Advanced Research Program Grant 003652-113.

[1] Applications in labor economics, such as the specification of earnings equations (Murphy and Welch, 1990), may also benefit from our approach.

upon a single variable in order to address the second of these problems. The modified BDS test is applied with respect to a single variable in the equation and does not rely on the overall goodness of fit of the regression model.

Complicated functional forms are often needed to obtain the highest possible goodness of fit measures. Achieving the best overall goodness of fit sacrifices, however, the parsimonious specification needed to make precise statements about single regression coefficients. The problem, as discussed in Cassel and Mendelsohn (1985), is that complicated functional forms can make it difficult to calculate marginal prices and standard errors of specific housing characteristics. We find that the modified BDS test gives valuable information when applied to the analysis of individual variables in hedonic housing regressions.

We apply the modified BDS test to the specification of a hedonic housing price equation used in Kohlhase (1991). We search for the existence of systematic nonlinearities in the error term of a regression. If systematic nonlinearities are found, we can conclude that the error term is not independently and identically distributed and that the regression specification needs to be explored further. We apply the test and show how it leads to a parsimonious specification with ease of interpretation of the coefficients of a single variable.

The test is described in section II. We show that ordering by one of the right hand side variables of the regression allows the BDS test to detect correlations between the included variable and information excluded from the regression. Section III discusses the application of the modified BDS test to the hedonic housing equation. This application shows that a parsimonious specification, which allows straightforward analysis of the effects of a single variable and relatively uncomplicated calculation of the resulting standard errors, retains some of the desirable features of more complex functional forms. Section IV offers conclusions and suggestions for further research.

II. A One Variable Specification Test

The problem we analyze is to determine the correct specification of a hedonic housing price equation such as:

$$g(P) = f(\mathbf{X}) + e \tag{1}$$

where P is the price of a residential dwelling, and $\mathbf{X}$ represents a vector of characteristics of the house, the neighborhood characteristics of the area, and its location. Goodness of fit tests, such as Box and Cox (1964), examine the overall specification of (1) and maximize an appropriately defined likelihood function. As pointed out in Cassel and Mendelsohn (1985), however, if the true specification of (1) is

$$\ln(P) = \sum_i a_i \ln(X_i) + b_j X_j + e \tag{2}$$

then goodness of fit tests will prefer a log-log[2] specification over all others. This is in spite of the fact that X_j is not specified correctly in the log-log form. Such a misspecification may be crucial if X_j is the variable of interest. The modified BDS test developed here is the first diagnostic test capable of examining the error structure with respect to a single variable.

A. *Outline of the Basic BDS Test*

The BDS test developed in Brock et al. (1987) examines the pattern of a series of data. Compare the ith observation to the $(i + 1)$th observation. There is some positive probability that this pair of observations is within a known distance (d) of each other. Now consider a third observation. Again there is a probability that it is within the same distance of the second. If these three observations are from an iid series, the probability that each pair is within d is well defined. The BDS test compares how often a series of data points are actually within d of each other to the expected value if the series were distributed iid.

Define an Indicator function I, where $I = 1$ if the absolute value of the difference between two points is less than or equal to d, and $I = 0$ otherwise. The embedding dimension m defines the number of points that are compared in a series.[3] The function C depends on the sum over the product of the indicator functions for each m

[2] Here we define log-linear to mean $\ln(P)$ is a function of linear X's, linear to mean a linear specification, and log-log to indicate logs of both the left and right hand side variables.

[3] The embedding dimension defines the "m dimensional cube" that would be filled by the data if the data are truly random. If the data are deterministic, then a function of the data would show a relationship in m dimensional space. For our purposes, the embedding dimension simply represents a parameter to be specified, where m can be specified larger the larger is the data set.

as

$$C(m,n,d) = C\left[\sum_{1\leq i\leq j\leq n}\left(\prod_{k=0}^{m-1} I_d(X_{i+k}, X_{j+k})\right)\right] \quad (3)$$

where m is the embedding dimension, n is the number of observations minus m, d is the distance between two points that causes the indicator function to equal one, and X represents a data point. The expression in parentheses will equal one if all $m - 1$ pairs of points are within d of each other, and it will equal zero if even one of the pairs of points is farther apart than d. The summation shows how many times in an entire data series there is a succession of $m - 1$ pairs that are close (within d).

The BDS test relies on a proof that $C(m,n,d)$ asymptotically approaches C^m in probability when the data series is independent and identically distributed. Brock et al. (1987) use this result to derive a statistic, called a C statistic, that depends on comparing the actual calculation of $C(m,n,d)$ to the calculation of $[C(1,n,d)]^m$. If the data are independent and identically distributed, this difference should approach zero. Asymptotically, the difference between the two functions divided by the standard error approaches a standard normal distribution. Dechert (1987) has written a computer program to calculate the BDS statistic defined as

$$BDS = n^{1/2}\{C(m,n,d) - [C(1,n,d)]^m\}/b_m \quad (4)$$

where the numerator is called the C statistic and b_m is the standard deviation of the C statistic.[4] If there are more (or less) points close together than would be expected from iid data, where $[C(1,n,d)]^m$ is the number of times a series of iid points would be expected to be close, then the BDS statistic indicates the likelihood that those points could have been generated by an iid series.

B. Application of the BDS test to Micro Data

The BDS test described in (4) has been applied to macroeconomic time series data to search for occurrences of systematic nonlinearities (for example Brock and Sayers, 1988). In a time series context, the data have a natural order. In the microeconomic context we must impose sequencing on the data by sorting on the variable of interest. What further differentiates our application of the BDS test from the time series context is that application to micro data allows consideration of why a series of residuals might or might not be "close" to one another. If the observations are ordered randomly, it would be expected that the BDS statistic would be zero, as the error term should be iid. On the other hand, assume the data are ordered by one of the right hand side variables. If there is a misspecification error, either through incorrect functional form or an omitted variable that is correlated with the included variables, then the BDS statistic will be significantly different from zero.

Brock and Dechert (1988) have proven that the BDS test is a good specification test for time series data provided the residuals of the estimated regression can be written as a function of the current and lagged values of the residuals of the true relation. In the case of omitted variables for micro data, the intuition works the same way because the estimated residuals are a function of the omitted variables. This can be seen by assuming the true model is

$$g(P) = f(X) + h(Z) + u \quad (5)$$

where u is independently and identically distributed. Assume the equation is estimated omitting the Z terms:

$$g(P) = f(X) + v \quad (6)$$

so that the error term of the estimated regression is

$$v = h(Z) + u. \quad (7)$$

If the excluded information (Z) from the true regression (5) is correlated with the included information (X) in the actual regression (6), then equation (7) can be rewritten to depend upon only the included information:

$$v = w(X) + e + u = g(P) - f(X) \quad (8)$$

where e is an iid error term independent of u and $w(X)$ describes how the excluded information is correlated with the included information.

Brock and Dechert (1988) show that if $g(P) - f(X)$ is independent of X, then v is iid. The

[4] This is the C test based on Corollary 2 of Theorem 2 of Brock et al. (1987), rather than the S test which relies on the ratio of C functions.

modification to the BDS test that we propose involves ordering the data by X, the included variable of interest. If there is a relation such as $w(X)$, then sorting by X will uncover the fact that $g(P) - f(X)$ will depend on the size of X. Thus the indicator function in the BDS test will capture the relationship in $w(X)$ by showing that the error terms do not have equal probability of being close to each other. We therefore are able to use the BDS test as a specification test to look for the importance of excluded information from the regression.

To see why the pattern of residuals reflects the information omitted from the regression, consider a set of residuals sorted by X and the probability that a given pair of residuals will be within, in absolute value terms, d of each other. If the series of residuals are iid, the probability that one pair of residuals is close to each other is unaffected by whether the previous pair of residuals is close or not. If the residuals reflect the omitted information $w(X)$, however, then the probability that they are within d of each other can be predicted based on knowledge of whether the previous residuals are close to each other. The prediction is possible only because the previous pair reflects the impact of the omitted information. That is, the BDS test will show a significantly greater probability than independence of finding residuals close to each other when the residuals have been ordered by an included variable. Therefore the probability that the BDS statistic is significantly different from zero depends on the probability that there is an omitted variable that is correlated with the included variable.

The key to the test we propose is how the excluded information impacts the included variables. When the functional form of a regression is not correct, it is quite likely that the excluded information is related to the included information. Even in the case of truly omitted variables, however, the BDS test will detect the influence of omitted information if it is correlated with the included information. This poses a potential problem for the applied researcher, in that a significant value of BDS (showing a lack of iid) may mean more than one problem exists. On the other hand, a BDS value insignificantly different from zero shows that excluded information does not have any significant effect on the estimated parameters of interest. Thus our modified BDS test has the advantage of providing a stopping rule in the search for an appropriate specification.

III. Empirical Application of the Modified BDS Test

We apply the modified BDS test to the residuals from the housing price equation used in Kohlhase (1991). Hedonic estimation to find marginal prices of various housing attributes provides a particularly useful application because there is no theoretically preferred functional form and the correct functional form is likely to be rather complicated. For many purposes, however, what is of interest is the marginal price of a specific housing attribute. We focus here on the marginal price of square feet of living area ($SQFT$) for illustrative purposes.

The data set consists of 1366 cross sectional observations on housing prices in Houston in 1985. There are three types of variables differentiated by geographic level.[5] The first type differs by observation and includes housing characteristics such as $SQFT$. The second and third types vary by geographic area, Keymap number and census tract level.[6] The second and third type of variables may have systematic measurement error due to geographic aggregation bias, so we concentrate on discussing the specification of a house-specific variable $SQFT$, and report other specifications for comparison. Variables that are truly constant over geographic area, however, should not pose a problem for the modified BDS test.

Our goal is to find an estimate of the marginal price of $SQFT$ where we are confident that we have captured any nonlinearities in the specification of this variable. We proceed in two steps.

[5] Variables included in the regression are of 4 types: (1) house-specific—square feet, lotsize, parking, bedrooms, fireplaces, baths, condition, age, central air, range, dishwasher, (2) Keymap number specific—distance to toxic waste site, (3) census tract specific—distance to employment, % owner occupied, % with high school education, income, poor, blue collar, black, hispanic, % under 19 years, and (4) time period—dummies for quarter. A data appendix is available from the authors.

[6] There are 286 keymap letter areas and 155 census tract areas in the data set. A Keymap letter area is about 0.56 square miles and the average census tract is about 3.44 square miles.

First we examine three primary simple specifications, linear, log-log, and log-linear to find a preferred basic functional form. Because the preferred log-linear specification does not pass the modified BDS test, we examine alternative interaction and higher order terms. This allows us to find a relatively parsimonious specification where the marginal price of *SQFT* and its standard error are straightforward to calculate, but which still meets the requirements of statistical reliability as identified by an iid error structure.

A. *Evidence for Systematic Nonlinearities*

Table 1 presents the *C* statistics, their standard errors, and the resulting BDS statistics for the three alternative hedonic housing price equation specifications with the data ordered by *SQFT*.[7] There are two basic results presented in table 1. First, there is considerable evidence of nonlinear structure, as the BDS statistics are all over 4.[8] Second, table 1 shows that the log-linear functional form is clearly preferred to the linear, and is weakly preferred to the log-log. This can be seen because the log-linear specification has a *C* statistic significantly below that of the linear model, and a *C* statistic lower than that of the log-log model.[9] We will therefore use the log-linear specification in the remainder of the paper.[10]

We find the BDS test to have considerably more power for detecting complex patterns in the error structure than do commonly used tests for functional form. We perform two standard functional form tests: MacKinnon, White and Davidson's (1983) PE test and Kmenta's (1986) Durbin-Watson (DW) statistic test. Neither provides any clear guidance about which functional form of the regression is preferred and the inconclusive results in the DW test do not even warn the researcher of the nonlinearity problem.

TABLE 1.—MODIFIED BDS TEST RESULTS FOR ALTERNATIVE FUNCTIONAL FORMS

Basic Model[b]	C^a (Standard Error)	BDS Statistic[c]
Linear	1.57 (0.15)	10.61
Log-Log	0.43 (0.09)	4.84
Log-Linear	0.29 (0.07)	4.06

[a] Calculated from the regression residuals sorted by *SQFT*.
[b] See text.
[c] Significance test from zero.

We also estimate an autoregressive moving average (ARMA) model on the basic residuals ordered by *SQFT* for the log-linear regression, and find that, based on the Ljung-Box *Q* Statistic, the appropriate specification for our data is an ARMA (1, 1). We then apply the modified BDS test to the residuals from the ARMA model. Purging the linear structure from the residuals, however, appears to have little effect. The *C* statistic falls insignificantly from 0.29 (0.07) for the basic model residuals to 0.27 (0.07) on the ARMA residuals. Further, the resulting BDS statistic of 3.82 indicates that significant nonlinear structure remains.

B. *Respecification of the Basic Model*

While the preference for the log-linear specification is consistent with results elsewhere, the BDS statistics in table 1 indicate that further model specification is needed. There is statistical evidence that the regression residuals vary systematically, but nonlinearly, with the size of *SQFT*. When the regression residuals are ordered randomly, the BDS statistic falls to 0.003 with a standard error of 0.06. It is therefore because of ordering by *SQFT* that we uncover a pattern in the residuals through the BDS test.

Possible causes of a significant BDS statistic include omitted unobserved variables, higher order terms, or interaction terms. Clearly, it is not

[7] We use an embedding dimension (m) of 5 and a distance (d) of the standard deviation of the data divided by the range. The results are robust for values of m between 3 and 10 and for values of d between 0.5 and 1.5 times that chosen.

[8] We use standard normal tables to calculate probabilities. Hsieh and LeBaron (1988) show that small sample properties of the BDS statistic only become important for sample sizes less than 500.

[9] The test of significance is approximate only and is based on comparing the statistic $z = (C1 - C2)/(\text{Var}(C1 - C2))^{.5}$ to a standard normal distribution, where $C1$ is the C-statistic in model 1. This calculation is approximate because the covariance between $C1$ and $C2$ has not yet been theoretically derived and is assumed here to be zero. z equals 7.73 comparing the linear model to the log-linear, and z equals 1.21 comparing the log-log to the log-linear model.

[10] Both the linear and the log-linear specification include *SQFT* squared as in the basic Kohlhase model. The BDS statistic without the squared term is significantly higher under both of these specifications, but the relative results are unchanged.

possible to include unobserved variables. We therefore first search higher order terms of the included variables. While squared terms have a theoretical basis, to capture variation in marginal effects for example, further higher order terms are not very appealing theoretically. Including terms up to *SQFT* to the sixth power, however, reduce the BDS to 3.12 which is still significantly different from zero. We conclude that simple nonlinearities are not the explanation for the observed pattern of the regression residuals.

To refine the model further, we desire a parsimonious set of interaction terms that will still allow a relatively straightforward calculation of the marginal price of *SQFT*. There is no clear guidance as to which interaction terms, or how many interaction variables, might be appropriate. Table 2 presents BDS results when each of the house-specific characteristics, and two region-specific variables, is used as an interaction variable with all of the other variables in the regression model. In all cases, we ordered the data by *SQFT* to calculate the BDS statistic. We find that an interaction with the number of bedrooms (*BED*) passes the BDS test.[11] The BDS statistic of 0.92 shows that the hypothesis of iid residuals cannot be rejected at the 5% level. This specification has the further attraction that the marginal price of *SQFT* calculation only involves *SQFT*, the squared term, and the interaction of these two variables with BED.

Choosing a specification with *BED* as the interaction variable is attractive on a priori grounds. The interaction allows the marginal price of *SQFT* to vary by the number of bedrooms. Table 3 presents the marginal price of *SQFT* and its standard error under the two alternative specifications. In the basic model, a marginal price of a square foot of living area is estimated to cost \$52.49 irrespective of other house characteristics. Under the interactive and correctly specified model, however, the marginal price varies from \$33.42 for a two bedroom house to \$85.95 for five bedrooms. The modified BDS test allows a certain confidence that the interaction specification has captured any systematic nonlinearities with respect to square feet. In addition, the advantage of the parsimonious specification is shown because the marginal prices of *SQFT* calculated under the interaction model are still estimated relatively precisely.

Because of the success of the *BED* interaction model, we examine other housing characteristics

TABLE 2.—MODIFIED BDS TEST RESULTS WITH VARIOUS INTERACTION TERMS

	C^a (Standard Error)	BDS statistic[b]
Interaction Variable		
SQFT	.18 (.07)	2.66
Lotsize	.24 (.06)	4.16
Bedrooms	.06 (.07)	0.92
Fire	.23 (.07)	3.30
Bath	.15 (.07)	2.18
Cond	.29 (.07)	4.08
Age	.30 (.07)	4.11
Air	.24 (.07)	3.60
Range	.30 (.07)	4.54
Toxic	.21 (.08)	2.77
CBD	.41 (.08)	5.16
Other Specifications		
No Sqft Squared	.44 (.07)	5.97
With Higher Order Sqft Terms	.23 (.07)	3.12

[a] Calculated from the regression residuals sorted by SQFT.
[b] Significance test from zero.

TABLE 3.—MARGINAL PRICE OF A SQUARE FOOT OF LIVING AREA[a]

Basic Model	\$52.49 (2.34)
Bedroom Interaction	
—Two Bedrooms	33.42 (4.06)
—Three Bedrooms	50.93 (2.32)
—Four Bedrooms	68.44 (3.64)
—Five Bedrooms	85.95 (6.33)

[a] Asymptotic standard errors are in parentheses.

[11] The *BED* interaction specification is not the only possible specification that passes the modified BDS test, for example, additional interaction variables also pass. Brock and Dechert (1988) conclude their paper by speculating on whether the BDS test is a unique test for specification error. Our empirical results indicates that it is not.

TABLE 4.—MODIFIED BDS TEST RESULTS WITH INTERACTION TERMS

	Basic Model		Bedroom Interaction	
	C[a] (Standard Error)	BDS statistic[b]	C[a] (Standard Error)	BDS statistic[b]
Sorted Variable				
SQFT	.29 (.07)	4.06	.06 (.07)	0.92
Lotsize	.32 (.07)	4.44	.28 (.07)	4.25
Bedrooms	.51 (.07)	7.09	.25 (.07)	3.78
Fire	.34 (.07)	4.69	.14 (.07)	2.12
Bath	.28 (.07)	3.98	.11 (.07)	1.64
Cond	.29 (.07)	4.08	.11 (.07)	1.65
Age	.53 (.07)	7.44	.32 (.07)	4.76
Air	.37 (.07)	5.18	.11 (.07)	1.66
Range	.44 (.07)	6.15	.20 (.07)	3.01
Toxic	.71 (.07)	9.87	.55 (.07)	8.26
CBD	.82 (.07)	11.42	.49 (.07)	7.40

[a] Calculated from the regression residuals sorted by indicated variable.
[b] Significance test from zero.

and their performance on the BDS test. Table 4 presents the results. The first two columns show BDS test results for the basic model when the residuals are ordered by each of the housing specific characteristics, plus the Keymap variable *TOXIC* as well as the Census tract variable *CBD*.[12] None of the variables that we use to sort the residuals show an insignificant BDS statistic. The second two columns show the residuals of the *BED* interaction model when ordered by each of the variables. In addition to *SQFT*, three other variables now show BDS statistics insignificantly different from zero at the 5% level: the number of baths, condition of the house, and the presence of air conditioning. In fact, with the exception of lotsize, all of the variables in the model show a significant improvement in the C statistics. Thus the number of bedrooms appears to be an important source of nonlinearities with respect to many of the housing characteristics.

[12] We omitted from table 4 two house-specific characteristics, dishwashers and number of parking spaces, because the coefficient estimates in the basic model were quite low and imprecise (t-statistics less than 0.1).

One of the most valuable aspects of the modified BDS test is that unlike many other tests, the BDS test provides a stopping rule in the search for an appropriate specification. When the residuals are found to be iid, information excluded from the regression is not correlated with the ordering variable. For example, the BDS test shows that the residuals from the basic model augmented by bedroom interactions are iid when ordered by *SQFT*. In contrast, a standard F-test about the inclusion of more variables, such as *CBD* or *TOXIC* interactions, would lead the researcher to add unnecessary complexity.

We have shown that the modified BDS test is an important diagnostic tool when examining the specification of a regression equation with respect to a single variable. There are, however, important limitations to using the modified BDS test. As discussed earlier, when variable measurement exhibits geographic aggregation bias, the BDS test may show significant nonlinearities for all specifications. Further, we found several specifications where the BDS statistic is larger using three interactive variables than one.[13] This may be because some interaction variables induce nonlinearities into the data rather than control for them. Even more important, perhaps, is that in addition to incorrect functional form, another source of significant BDS statistics can be omitted variables. The finding that the BDS statistic when the regression residuals are ordered by lotsize do not fall with the *BED* interaction is not very reassuring. The BDS statistic using the basic log-linear model is 4.44. With the *BED* interaction, it fell slightly to 4.25. With three interaction variables, it is still 4.46. In fact, we failed to discover a specification for lotsize that provides a significant improvement in, much less passes, the modified BDS test. As with all diagnostic tests, the test is able to indicate where more work is needed without being very specific about how to correct the problem.

IV. Summary and Conclusions

This paper has shown that an application of chaos theory provides an excellent tool for investigating microeconomic model specification. The modified BDS test proposed here provides an important diagnostic test for the presence of nonlinear structure in microeconomic data, and provides an important indication about whether a particular variable in a regression is correctly specified. The modified BDS test examines whether the residuals of a regression equation exhibit a systematic pattern when ordered by the right hand side variable of interest. The test therefore determines whether excluded information from the specified regression is correlated with the variable of interest; a determination of no correlation indicates that the resulting parameter estimates are unbiased. The modified BDS test applied to micro data is an important specification test with respect to a single variable because possible sources of excluded information consist of alternative functional forms employing included variables.

Specification tests for a single variable are especially important in applications where a precise coefficient on a specific variable is desired. We have applied the modified BDS test to a hedonic housing price regression, and are able to find a relatively parsimonious specification where the residuals are iid when ordered by square feet of living area. Thus we offer the first attempt to mediate the trade-off between a complex specification of the overall equation and ease of manipulation for a particular regression coefficient.

The test for nonlinear structure, when used on macroeconomic data, can give evidence about whether the residuals are iid, but does not indicate how the model can be transformed to purge the systematic component from the residuals. In the microeconomic context, the modified BDS test is an excellent signpost about potential errors in modelling because ordering by the variable of interest provides a method for uncovering systematic nonlinearities in the residuals.

[13] For example, interactions with *BED*, *TOXIC*, and *CBD* produce a BDS statistic of 0.93, despite the fact that *BED* only as an interaction variable produces a BDS of 0.92.

REFERENCES

Baumol, William J. and Jess Benhabib, "Chaos: Significance, Mechanism, and Economic Applications," *The Journal of Economic Perspectives* 3 (Winter 1989), 77–106.

Box, G. E. P. and D. R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society Series B*, 26 (1964), 211–252.

Brock, William A., "Introduction to Chaos and Other Aspects of Nonlinearity," in W. A. Brock and A. G. Malliaris, *Differential Equations, Stability and Chaos in Dynamic Economics* (New York: North-Holland, 1988).

Brock, William A., and W. Davis Dechert, "A General Class of Specification Tests: The Scalar Case," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association (1988), 70–79.

Brock, William A., W. Davis Dechert, and Jose A. Scheinkman, "A Test for Independence Based on the Correlation Dimension," SSRI Working Paper #8702, University of Wisconsin-Madison, 1987.

Brock, William A., and Chera L. Sayers, "Is the Business Cycle Characterized by Deterministic Chaos?" *Journal of Monetary Economics* 22 (1988), 71–90.

Cassel, Eric, and Robert Mendelsohn, "The Choice of Functional Forms for Hedonic Price Equations: Comment," *Journal of Urban Economics* 18 (Sept. 1985), 135–42.

Dechert, W. Davis, "A Program to Calculate BDS Statistics for the IBM PC," University of Houston, 1987.

Grandmont, Jean-Michel, "On Endogenous Competitive Business Cycles," *Econometrica* 53 (Sept. 1985), 995–1045.

Hsieh, David A. and Blake LeBaron, "Finite Sample Properties of the BDS Statistic I: Distribution Under the Null Hypothesis," University of Chicago mimeo, May 1988.

Kmenta, Jan, *Econometrics* (New York: Macmillan, 1986).

Kohlhase, Janet E., "The Impact of Toxic Waste Sites on Housing Values," *Journal of Urban Economics* 30 July 1991), 1–26.

MacKinnon, J. G., H. White, and R. Davidson, "Tests for Model Specification in the Presence of Alternative Hypotheses: Some Further Results," *Journal of Econometrics* 21 (Jan. 1983), 53–70.

Murphy, Kevin M. and Finis Welch, "Empirical Age-Earnings Profiles," *Journal of Labor Economics* 8 (Apr. 1990), 202–229.