STUDIES IN

Nonlinear Dynamics and Econometrics

Volumes 1–3

Published by The MIT Press

The contents of this print collection were first published electronically in *Studies in Nonlinear Dynamics and Econometrics* (ISSN 1081-1826), a publication of the MIT Press. Except as otherwise noted, copyright for each article is owned by the Massachusetts Institute of Technology.

Subscriptions

Studies in Nonlinear Dynamics and Econometrics is an electronic journal, published quarterly on the Internet. Abstracts are freely accessible to all. Paid subscribers receive access to the full text of articles, e-mail notification of new material, an annual print edition and long-term archival access. For ordering information, contact:

MIT Press Journals
Five Cambridge Center
Cambridge MA 02142-1407 USA
Tel: 617-253-2889
Fax: 617-577-1545

journals-orders@mit.edu

or visit the journal's web site at:

http://mitpress.mit.edu/SNDE/

Studies in Nonlinear Dynamics and Econometrics (SNDE) is an electronic journal in the field of nonlinear analysis. The journal is a peer-reviewed, fully copyrighted, fee-based periodical hosted entirely on the Internet. SNDE was formed in recognition that advances in statistics and dynamical systems theory may increase our understanding of economic and financial markets. The journal seeks both theoretical and applied papers that characterize and motivate nonlinear phenomena. Researchers are encouraged to assist replication of empirical results by providing copies of data and programs online. Algorithms and rapid communications are also published.

Guide for Subscribers

The electronic versions of the published papers *Studies in Nonlinear Dynamics and Econometrics* (SNDE) will be made available in 2 formats:

- PostScript
- PDF

To view and/or download the SNDE papers for archiving and printing, you will need a standard Web browser.
In order to view PDF files through your browser, and also download them to your printer, you will need a copy of the Adobe Acrobat Reader available separately from the Adobe Web site:

http://www.adobe.com/prodindex/acrobat/readstep.html.

ISBN 0-262-75904-7

Studies in Nonlinear Dynamics and Econometrics

Volumes 1–3

On Cycles and Chaos in Economics

Jess Benhabib
Department of Economics
New York University
benhabib@fasecon.econ.nyu.edu

The possibility of cycles and chaos arising from nonlinear dynamics in economics emerged in the literature in the 1980s, and it came as a surprise.[1] The possibility of deterministic cycles in economic models had been noted before, for example in the well-known multiplier-accelerator models, but not in equilibrium models with complete markets, no frictions, and full intertemporal arbitrage.[2] The reason for the surprise was understandable: deterministic fluctuations in equilibrium models involve predictable changes in relative prices which should be ruled out by intertemporal arbitrage. In models of overlapping generations, however, finite lives can restrict complete arbitrage over time. As a result, some people thought, and still think, that cycles that are shorter than the agents' postulated life-spans would not be possible in equilibrium models, and therefore are irrelevant for business-cycle analysis. This view is clearly wrong, and of course ignores the extensive literature on cycles and chaos in optimal growth models with infinitely lived agents. In such models deterministic cycles in relative prices occur easily, but the amplitudes of the cycles remain within bounds of the discount rate.[3] It is not difficult to show in the context of multisector growth models, say with Cobb-Douglas production functions, that for any positive discount rate there is a large class of technologies for which cycles occur. (See Benhabib and Rustichini [1990].) Getting chaos, however, is harder. Recent works by Sorger (1992), by Mitra (1995), and by Nishimura and Yano (1995) give lower bounds for the discount rate, below which chaos is ruled out for one-sector models of optimal growth. Yet even in that context, going to a multisector framework may considerably lower the bounds on the discount rate thus far established.

A second reason for the attention that chaotic dynamics received in the economics literature regards prediction. The common wisdom has been that economic fluctuations are driven by exogenous shocks. Chaotic dynamics not only supplied an alternative explanation for at least some part of economic fluctuations, but also provided an excuse for economists' difficulties with forecasting. Sometimes, however, an important feature of chaotic dynamics that makes forecasting difficult, namely, sensitive dependence on initial conditions, is used in a cavalier way to explain short-run dynamics, forgetting that the effect of sensitive dependence becomes significant only after some periods, but not in the very short run.

When it became obvious that very-standard equilibrium models could easily generate cycles and chaos, the attention in the literature naturally turned to the empirical plausibility of such dynamics. The most interesting approach, inspired by developments in natural sciences and mathematics, was also atheoretical, and reminiscent of VAR methods of time series.[4] The idea was to try to infer whether a particular economic time series was generated by a deterministic, low (at most four- or five-) dimensional system that was chaotic, or whether it came from a simple (linear) stochastic system. It is not difficult to see that such inferences are hard to make when the time-series data is short, as is the case with most economic series, with the exception of financial data. It is not surprising, then, that many applications of this approach are in the area of finance, but even there, where we have very high-frequency data, it is hard to pick up fluctuations that may occur at lower

---

[1] Among the many authors contributing to this early literature were, Benhabib, Boldrin, Day, Deneckere, Grandmont, and Montrucchio. See the survey in this volume by Nishimura and Sorger (1996).

[2] Some earlier examples include the well-known cyclical equilibria in overlapping generations model by David Gale (1973), and some not very robust cyclical counterexamples to the turnpike theorems by Sutherland (1970).

[3] For example if the discount rate is 4%, relative prices can change by up to 12% over a three-year upswing; how quantities respond would depend on production possibilities and labor-supply elasticities.

[4] Among the many authors who contributed to this literature are Brock, Dechert, LeBaron, Scheinkman, and others. A good starting point is the book by Brock, Hsieh, and LeBaron (1991).

 Studies in Nonlinear Dynamics and Econometrics, April 1996, 1(1): 1–2

business-cycle frequencies. Another serious limitation of the atheoretical approach is that its conclusions must be limited to determining whether or not the time series come from a low-dimensional (aggregative) chaotic system. When resources in the economy are close to fully utilized, much of the fluctuations and disruptions may occur at the micro level and between sectors. In this context the underlying system may not be "low" dimensional, and it may also be contaminated by noise. Not surprisingly, therefore, many of the results that emerge from this empirical approach are either inconclusive or simply reject low-dimensional chaos.

Another (and maybe more fruitful) approach would be to bring back economic theory for purposes of empirically evaluating the plausibility of chaos and cycles. Maybe less-aggregative models of economic dynamics could be estimated or appropriately calibrated using microeconomic data. We could then directly check whether the estimated or calibrated parameters generate significant nonlinearities and fall within the theoretical range that would generate cycles and chaos. Of course, the theoretical characterization of high-dimensional systems generating chaotic dynamics is still not fully developed, and is difficult to apply. In the meantime, we may complement theoretical methods by simulating sensibly parametrized models to obtain some insight into their dynamics.

References

Benhabib, Jess and Aldo Rustichini (1990). "Equilibrium Cycling with Small Discounting," *Journal of Economic Theory* 5, 423–432.

Brock, William A., David Hsieh, and Blake LeBaron (1991), *Nonlinear Dynamics, Chaos and Instability*. Cambridge, Massachusetts: MIT Press.

Gale, David (1973), "Pure Exchange Equilibrium of Dynamic Economics Models," *Journal of Economic Theory* 6, 12–36.

Mitra, T. (1996), "An Exact Discount Factor Restriction for Period Three Cycles in Dynamic Optimization Models," *Journal of Economic Theory* 69, 281–305.

Nishimura, Kazuo and Gerhard Sorger (1996), "Optimal Cycles and Chaos: A Survey," *Studies in Nonlinear Dynamics and Econometrics* 1, 11–28.

Nishimura, Kazuo and M. Yano (1996), "On the Least Upper Bound of Discount Factors that are Compatible with Optimal Period-Three Cycles, *Journal of Economic Theory* 69, 306–333.

Sorger, Gerhard (1992), "On the Minimum Rate of Impatience for Complicated Optimal Growth Paths," *Journal of Economic Theory* 56, 160–179.

Sutherland, William A. (1970), "On Optimal Development in a Multi-Sector Economy: The Discounted Case," *Review of Economic Studies* 37, 585–589.

Power Properties of Linearity Tests for Time Series

Timo Teräsvirta
Department of Economic Statistics
Stockholm School of Economics
sttimo@hhs.se

Abstract. *This paper examines the power properties of several linearity tests applied in time-series analysis. The tests are the ones Lee et al. (1993) used in their Monte Carlo study. The main tool used for power comparisons in this paper is the Pitman asymptotic relative efficiency. The results generally strengthen the outcome of the simulations and complement some results in Lee et al. (1993). They also suggest guidelines for designing Monte Carlo experiments for linearity tests.*

Keywords. Bilinear model, local asymptotic power, nonlinear time series, Pitman asymptotic relative efficiency, threshold autoregressive model.

Acknowledgments. Financial support from the Academy of Finland and the Yrjö Jahnsson Foundation is gratefully acknowledged. A part of the research for this paper was carried out while the author was visiting the University of California, San Diego, and the work continued at the Research Department of the Central Bank of Norway, Oslo. I am grateful to the econometricians in both places for providing such a stimulating research atmosphere, and to Clive Granger in particular for helpful discussions on the present topic. I am also indebted to a referee for a series of helpful comments. Any remaining errors and shortcomings are my own responsibility.

1 Introduction

With increasing interest in nonlinear time-series models, testing linearity against nonlinearity has become an important issue in time-series analysis; see Granger and Teräsvirta (1993, chapter 6) for an overview. Recently, Lee, White, and Granger (1993), henceforth LWG, conducted a wide array of simulation experiments to study the power properties of a few linearity tests against different types of nonlinearity, including bivariate nonlinear models. Another study with a similar aim is Luukkonen, Saikkonen, and Teräsvirta (1988a). The main purpose of the present paper is to demonstrate that most of the simulation results of LWG can be explained or further illuminated using linearization and statistical theory, and, particularly, the concept of Pitman asymptotic relative efficiency (ARE). This theory can also be used when designing new simulation experiments. The plan of the paper is as follows. Section 2 briefly discusses most of the tests LWG considered by means of auxiliary regressions. Section 3 discusses the concept of ARE. Section 4 considers power properties of tests applied to Block1 models of LWG, and section 5 contains final remarks.

2 Linearity Tests

I shall concentrate on the Block1 models in LWG. For this purpose, consider the following artificial model:

$$y_t = a_0 + \sum_{j=1}^{k} a_j z_t^j + u_t, \qquad u_t \sim \text{nid}(0, \sigma^2). \tag{2.1}$$

If either $z_t = y_{t-1}$ (univariate models) or $z_t = x_t$ (bivariate models), $a_0 = 0$, $k = 2$, then the F test of $a_2 = 0$ is TSAY1 (Tsay, 1986) as in LWG. If $k = 3$ and the above holds except that I assume $a_0 \neq 0$, the F test of $a_2 = a_3 = 0$ is a Lagrange-multiplier-type test of linearity against the hypothesis that the true model is a single hidden-layer artificial neural network model, as discussed in Teräsvirta, Lin, and Granger (1993). This test often has better power than the test LWG preferred, which is based on the same neural network model. However, when the true model is not a single hidden-layer neural network model, there are situations in which the test in Teräsvirta et al. (1993) is less powerful than the test in LWG and is not even consistent. A simple example is the model $y_t = a_4 y_{t-1}^4 + u_t$, $u_t \sim \text{nid}(0, \sigma^2)$, because the test in Teräsvirta et al. (1993) is only based on the first three moments and cross-moments of lags of y_t. In this paper, I nevertheless take that test to represent the neural network (NN) test. The reason is that many nonlinear models simulated in LWG may at least locally in a neighborhood of the null of linearity be approximated by a model resembling equation (2.1). As will be seen, comparing (2.1) to such a representation makes it possible to find an explanation for the fact that the NN test sometimes seems to have better power than some other linearity tests.

Furthermore, if $z_t = f_t$ where f_t is the OLS fit of the linear model $y_t = ay_{t-1} + u_t$, the F test of $a_j = 0$, $j \geq 2$, in equation (2.1), assuming $a_0 = 0$ is RESET (Ramsey, 1969). For Block1 models, $f_t = \hat{a}y_{t-1}$ (or $f_t = \hat{a}x_{t-1}$), so that f_t can be replaced by y_{t-1} (or x_t). If $k = 3$, RESET thus coincides with the auxiliary regression version of the NN test, except that $a_0 = 0$. The Keenan test (Keenan, 1985) is a special case of RESET such that $k = 2$, and for Block1 models the test is thus identical to TSAY1. LWG, however, carry out RESET differently. In their version, $z_t = y_{t-1}$, $z_t^j = \beta_j' w_t$, $j = 2, \ldots, k$, in equation (2.1), where: $w_t = (y_{t-1}^2, \ldots, y_{t-1}^k)'$ and $\beta_j' w_t$ is the t-th element of the j-th principal component of $W = (w_1', \ldots, w_T')'$. For bivariate models, x_t takes the place of y_{t-1}.

Next, consider the auxiliary regression

$$y_t = ay_{t-1} + c_1 u_{t-1} + c_{11} u_{t-1} y_{t-1} + c_{12} u_{t-1} y_{t-2} + c_{112} u_{t-1} y_{t-1} y_{t-2} + u_t,\ u_t \sim \text{nid}(0, \sigma^2). \tag{2.2}$$

The χ^2 (Lagrange multiplier, or LM) test of $c_1 = c_{11} = c_{12} = c_{112} = 0$ in equation (2.2) is WHITE3 of LWG. For the origins of this test, see White (1987). The lagged errors are replaced by their estimates from the OLS regression of y_t on y_{t-1}. Note that the test is not a pure linearity test, as the alternative also contains a first-order moving average term. This explains the results for White tests in Table 5 of LWG for Model1 of Block2. Of the remaining tests, TSAY2 (another version of the test of Tsay, 1986) requires an auxiliary regression with 16 regressors, 15 of which have zero coefficients under the null of linearity. These auxiliary regressions form a starting point for considering and comparing the power of different tests.

3 Asymptotic Relative Efficiency

To introduce the concept of Pitman asymptotic relative efficiency (ARE), I consider two nonlinear models and assume that one of them has generated the data. This model is characterized by the log-likelihood function

$$q_T(a, \phi) = \sum_{t=1}^{T} \bar{q}_t(a, \phi), \tag{3.1}$$

where a is a $p \times 1$ and ϕ an $s \times 1$ parameter vector, and T is the sample size. The other (misspecified or inappropriate) model is characterized by the log-likelihood function $p_T(a, \psi) = \sum_{t=1}^{T} \bar{p}_t(a, \psi)$, where ψ is an $r \times 1$ parameter vector. Generally, $p_T(a, \phi) \neq q_T(a, \psi)$, but $p_T(a, 0) = q_T(a, 0)$. I want to test $\text{H}_{0\psi} : \psi = 0$ against $\text{H}_{1\psi} : \psi \neq 0$, whereas the relevant null hypothesis is $\text{H}_{0\phi} : \phi = 0$ against $\text{H}_{1\phi} : \phi \neq 0$, because equation (3.1) characterizes the true model. Consider a sequence of local alternatives $\phi = \delta / T^{\frac{1}{2}}$, $\delta \neq 0$, so that the data are generated by a model with the log-likelihood function $q_T(a, \delta / T^{\frac{1}{2}})$. Define

$$\bar{k}_t = \left[\{\partial \bar{p}_t(a, 0) / \partial a\}' \ \{\partial \bar{p}_t(a, 0) / \partial \psi\}' \ \{\partial \bar{q}_t(a, 0) / \partial \phi\}' \right]$$

with covariance matrix $\Sigma = E\bar{k}_t\bar{k}_t' = [\Sigma_{ij}]$, $i, j \in \{a, \psi, \phi\}$, where the partition conforms to that of $\bar{k}_t$. Then the asymptotic distribution of the LM test of $H_{0\psi}$ $(LM_{0\psi})$ follows an asymptotic χ_r^2 distribution with noncentrality parameter $\lambda_\psi(a, \delta) = \delta'\Sigma_{\phi\psi\cdot a}\Sigma_{\psi\psi\cdot a}^{-1}\Sigma_{\psi\phi\cdot a}\delta$, where $\Sigma_{ij\cdot k} = \Sigma_{ij} - \Sigma_{ik}\Sigma_{kk}^{-1}\Sigma_{kj}$ (see Saikkonen [1989] or Luukkonen et al. [1988a]). On the other hand, the asymptotic distribution of the LM test of $H_{0\phi}$ is a noncentral χ_s^2 distribution with noncentrality parameter $\lambda_\phi(a, \delta) = \delta'\Sigma_{\phi\phi\cdot a}\delta$. The asymptotic relative efficiency of $LM_{0\psi}$ is the following ratio (Saikkonen, 1989):

$$\text{ARE}_\psi(a, \delta, \alpha, \beta) = \frac{\lambda_\psi(a, \delta)d(s, \alpha, \beta)}{\lambda_\phi(a, \delta)d(r, \alpha, \beta)}\,. \tag{3.2}$$

In equation (3.2), $d(h, \alpha, \beta)$ is the noncentrality parameter of a noncentral χ_h^2 distribution, such that the 1–β fractile of that distribution and the 1–α fractile of the (central) χ_h^2 distribution coincide. Values of $d(h, \alpha, \beta)$ are tabulated, for example, in Pearson and Hartley (1972, Table 25). If $r = s$, equation (3.2) does not depend on d. If $r = s = 1$, ARE is also independent of δ. If $\lambda_\psi(a, \delta) = 0$, the asymptotic power of LM_ψ against the local alternative equals the size of the test. It is worth noting already that the McLeod-Li test that LWG considered has this property for all models included in Block1, as is clear from Luukkonen et al. (1988a).

4 Interpreting Simulation Results

I shall now consider the simulation results of Block1 models in LWG using the auxiliary regression interpretation of TSAY1, TSAY2, the NN test, RESET, and WHITE3. The nonlinear models in Block2 consist of another bilinear model and two nonlinear moving-average models. The latter type are rarely applied in practice. For this reason, the focus will be on Block1 models.

The Threshold Autoregressive (TAR) model

Consider the following nonlinear model:

$$y_t = a_0 y_{t-1} + a_1 F_1(y_{t-1}) + a_2 F_2(y_{t-1}) + u_t, \tag{4.1}$$

where

$$F_j(y_{t-1}) = (1 + \exp\{-\gamma(y_{t-1} - c_j)\})^{-1} - \frac{1}{2}, \qquad j = 1, 2, \quad \gamma > 0, \quad c_1 < c_2. \tag{4.2}$$

LWG simulated equation (4.1) with $a_0 = 0.3$ and $a_1 = -a_2 = 1.2$ when $c_1 = -c_2 = -1$ and $\gamma \to \infty$ in equation (4.2). To study the performance of the tests, I linearize equation (4.1). An appropriate way of doing this is to replace F_j by a third-order Taylor expansion about $\gamma = 0$, because equation (4.1) is linear when $\gamma = 0$.

The third-order Taylor expansion to F_j is:

$$\begin{aligned} T_j(y_{t-1}) &= b_1(y_{t-1} - c_j) + b_3(y_{t-1} - c_j)^3 = b_3 y_{t-1}^3 - 3b_3 c_j y_{t-1}^2 \\ &\quad + (3b_3 c_j^2 + b_1)y_{t-1} - c_j(b_3 c_j^2 + b_1), \qquad j = 1, 2, \end{aligned} \tag{4.3}$$

where $b_1 = \gamma/4$ and $b_3 = \gamma^3/16$. Assuming $a_1 = -a_2$ and $c_1 = -c_2$ yields

$$\begin{aligned} \{a_1 T_1(y_{t-1}) + a_2 T_2(y_{t-1})\}y_{t-1} &= a_1\{T_1(y_{t-1}) - T_2(y_{t-1})\}y_{t-1} \\ &= -6a_1 b_3 c_1 y_{t-1}^3 - a_1(6b_3 c_1^3 - 2b_1 c_1)y_{t-1}. \end{aligned} \tag{4.4}$$

Thus the corresponding approximation of equation (4.1) is of the form:

$$y_t = a_1^* y_{t-1} + a_3^* y_{t-1}^3 + u_t', \tag{4.5}$$

and the linearity hypothesis is $H_0' : a_3^* = 0$. The fourth- and second-order terms theoretically present in equation (4.4) vanish because both $a_1 = -a_2$ and $c_1 = -c_2$.

From equation (4.5) it is seen that the NN test has power against equation (4.1), as its auxiliary regression contains the crucial third-order term y_{t-1}^3. On the other hand, y_{t-1}^2 in TSAY1 is a very poor substitute for y_{t-1}^3.

Assume equation (4.5) is the correct model, and equation (2.1) with $a_0 = 0$ and $k = 2$ is the inappropriate alternative. Then

$$\Sigma = E\bar{k}_t\bar{k}_t' = \begin{bmatrix} Ey_{t-1}^2 & Ey_{t-1}^3 & Ey_{t-1}^4 \\ & Ey_{t-1}^4 & Ey_{t-1}^5 \\ & & Ey_{t-1}^6 \end{bmatrix}$$

so that $\Sigma_{\phi\psi} = Ey_{t-1}^5 = 0$, which implies $\Sigma_{\phi\psi\cdot a} = 0$. Thus for any $|a_1^*| < 1$, $\lambda_\psi(a_1^*, \delta) = 0$, so that the local asymptotic power of TSAY1 against equation (4.5) is not higher than the size of the test. This explains the low empirical power of the test.

Next I apply ARE to consider the performance of WHITE3. Assume again that equation (4.5) is the true model. If I choose $\alpha = 0.05$ and $\beta = 0.5$, say, then the ARE of WHITE3 with respect to equation (4.5) is:

$$\mathrm{ARE}_W(a_1, \delta, 0.05, 0.5) = 1.896a_1^2(1 - a_1^2), \tag{4.6}$$

so that

$$\max_{|a_1|<1} \mathrm{ARE}_W(a_1, \delta, 0.05, 0.5) = 0.474 \text{ at } a_1 = \pm 1/\sqrt{2},$$

and $\mathrm{ARE}_W(0.3, \delta, 0.05, 0.5) = 0.155$. The computation of equation (4.6) (see the Appendix) shows that the only term in WHITE3 based on equation (2.2) that contributes to the (local) power is the third-order term $y_{t-1}y_{t-2}\hat{u}_{t-1}$. The test thus may be expected to detect nonlinearity in equation (4.1) for $a_1 = 0.3$ because equation (4.5) is a local approximation to (4.1) in the neighborhood of $\gamma = 0$. However, for the present parametrization, a much larger sample size than T = 200 is needed for that power to show. Note that the RESET based on principal components is also without power, although the linear combinations serving as regressors in the test do contain y_{t-1}^3.

The Sign (SGN) model

The sign model of LWG is:

$$y_t = \mathrm{sgn}(y_{t-1}) + u_t,\ u_t \sim \mathrm{nid}(0, 1), \tag{4.7}$$

where $\mathrm{sgn}(x) = 1,\ x > 0$; $\mathrm{sgn}(x) = 0,\ x = 0$; $\mathrm{sgn}(x) = -1,\ x < 0$.

Consider the generalization:

$$y_t = aF_H(y_{t-1}) + u_t, \tag{4.8}$$

where $F_H(y_{t-1}) = (1 + \exp\{-\gamma y_{t-1}\})^{-1} - \frac{1}{2}$, $\gamma > 0$. Choosing $a = 2$ and letting $\gamma \to \infty$ yields equation (4.7). Note that equation (4.8) is a special case of the single hidden-layer feedforward artificial neural network model (LWG; White, 1989) with a single hidden unit. The NN test of $H_0 : a = 0$ is therefore powerful by definition. Replacing $F_H(y_{t-1})$ in equation (4.8) by its third-order Taylor approximation about $\gamma = 0$ (linearity) yields $y_t = b_1 y_{t-1} + b_3 y_{t-1}^3 + u_t'$.

The arguments used above apply again. TSAY1 has little power because it lacks y_{t-1}^3. In fact, its empirical power in the experiments of LWG does not increase with the sample size. The ARE of the White test equals zero for $\alpha = 0.05$ and $\beta = 0.5$, which explains its lack of power. For comparison, RESET with principal components is clearly now more powerful than in the previous design.

Bivariate models

The above considerations also help explain the simulation results for the two bivariate models. For the SQ model $y_t = x_t^2 + u_t$, TSAY1 is the best test. The regression version of the neural network test indicates that the NN test of LWG is not as powerful, but may still have considerable power. Taking $y_t = ax_t + bx_t^2 + u_t$ to be the true model, it is easy to see that the asymptotic local power of WHITE3 is not higher than the size of the test because $\Sigma_{\phi\psi} = 0$. This is because x_t and u_{t-j} are independent at any lag j. Nevertheless, the empirical global power of WHITE3 is fairly high. This helps to put the performance of the BDS test into perspective. Its empirical power is high, but no higher than that of a test whose ARE equals zero. However, a problem with this test is that its size is not completely under control in small samples. This is because its asymptotic null distribution depends on two nuisance parameters: the embedding dimension and the nearness parameter; see for example Brock and Potter (1993). In fact, as to the SQ model, Table 8 in LWG is more informative than Table 4. The differences in power one may expect appear more clearly when the error standard deviation $\sigma = 20$. TSAY1 is then the most powerful test, followed by the NN test. The power of WHITE3 is low.

The power of the linearity tests against the model EXP($y_t = \exp(x_t) + u_t$) is best evaluated by approximating the exponent by a Taylor expansion. This shows that TSAY and the NN test should be the most powerful tests, and that both x_t^2 and x_t^3 are useful terms in the auxiliary regression. From Table 9 in LWG it is seen that TSAY1 and the NN test have about the same empirical power. Assuming that the true model is $y_t = ax_t + bx_t^2 + cx_t^3 + u_t$ and selecting equation (2.2) with $z_t = x_t$ to be the inappropriate one again gives $\Sigma_{\phi\psi} = 0$. This leads one to expect WHITE3 to be clearly less powerful than the other two tests, which also turns out to be the case.

Bilinear model

LWG simulated the (Block1) bilinear model

$$y_t = ay_{t-1} + c_{ij}u_{t-i}y_{t-j} + u_t, \qquad u_t \sim \text{nid}(0, \sigma^2), \tag{4.9}$$

where $i = 2$, $j = 1$, $a = 0$ and $c_{21} = 0.7$. WHITE3 had the highest power. As seen from equation (2.2), its auxiliary regression does not contain $u_{t-2}y_{t-1}$. To find out where the power comes from, I consider two misspecified models based on WHITE3, both of type (4.9). The idea is to separately consider contributions of different components to the power of the test in the WHITE3 equation (2.2). The first model has $i = j = 1$ and $c_{11} \neq 0$, whereas the second one has $i = 1$, $j = 2$ with $c_{12} \neq 0$. The ARE of the linearity test of $c_{11} = 0$ in the first misspecified model when the data were generated by equation (4.9) is:

$$\text{ARE}_1(a) = 4a^2(1 - a^2)^2/\{(1 + 2a^2 - 2a^4)(3 - 2a^2)\},$$

whereas that of the second inappropriate model equals

$$\text{ARE}_2(a) = (1 - a^2)^2/(1 + 2a^2 - 2a^4)$$

(see the Appendix). Now, $\text{ARE}_2(0) = 1$, indicating that the component $u_{t-1}y_{t-2}$ in the White tests is important, contributing power as $a = 0$ in equation (4.9). The McLeod-Li test also has high power, although its ARE compared to the test based on the true model (4.9) equals zero. The low power of TSAY1 therefore requires an explanation. It is sufficient to look at its ARE with respect to the test based on equation (4.9), which is:

$$\text{ARE}_{T1}(a) = 3a^2(1 - a^2)/(1 + 2a^2 - 2a^4) = 0,$$

for $a = 0$ (see the Appendix). On the other hand, TSAY2 is more powerful than TSAY1 although it contains more regressors. This suggests that some of the additional lags introducing cross-terms in the auxiliary regression of the Tsay test increase the power of the test. To investigate this, consider the following simple artificial model:

$$y_t = ay_{t-1} + c_{12}y_{t-1}y_{t-2} + u_t, \qquad u_t \sim \text{nid}(0, \sigma^2). \tag{4.10}$$

The ARE of the test of $c_{12} = 0$ in equation (4.10) compared to that of $c_{21} = 0$ in equation (4.9) equals

$$\text{ARE}_{T'}(a) = (1 - a^2)(1 + 2a^2)/(1 + 2a^2 - 2a^4),$$

so that $\text{ARE}_{T'}(0) = 1$. Thus, adding the second lag is likely to increase the power of the Tsay test substantially. TSAY2 uses five lags, which reduces the power again because many of the terms in the auxiliary regression of TSAY2 are redundant if one tests $c_{21} = 0$ in equation (4.3). A larger sample size and/or more noise in equation (4.9) would be needed to render TSAY2 more powerful than the McLeod-Li test. Note that the ARE of TSAY1 is maximized at $a = \pm 1/\sqrt{2}$, and the maximum equals $\frac{1}{2}$. Thus for $|a| = 1/\sqrt{2}$, TSAY1 would be clearly more powerful than the McLeod-Li test. The power of TSAY2 would decrease by increasing $|a|$ sufficiently as $\text{ARE}_{T'}(1) = 0$.

The bilinear model constitutes an interesting example, because the McLeod and Li test with zero ARE compared to the test based on equation (4.9) is quite powerful. The ARE is based on local considerations, and a test with zero ARE with respect to another test may still have plenty of power against a global alternative. The ARE thus does not necessarily say much about the empirical power of a test in a given experiment. What the ARE comparisons do is establish a ranking of tests in terms of power: for a given number of observations, a test with a higher ARE with respect to another test is expected to have more power than one with a lower ARE with respect to the same test.

This has implications for the design of simulation experiments. Assume that, in a pilot study, a test with zero ARE has high power (say, 0.9). Then one knows that the experiment will not be very informative about tests with higher ARE, because their power will only vary between 0.9 and unity. Suppose the experiment includes a test whose ARE is difficult or impossible to determine, and whose power turns out to lie between 0.9 and unity as well. Then it is difficult to say much about the relative performance of the test, because a test with zero ARE already is very powerful. In the present example, the BDS test is a case in point. It would be useful to redesign the experiment in such a way that the power of the tests with zero ARE would not be too high. This would allow more spread between powers of tests with zero ARE and those with high ARE, and give a better indication of the power properties of tests with unknown ARE. The designs in Block2 (Model5 and Model6) are not ideal in this respect either.

The argument also works the other way around. Suppose that in a simulation experiment the tests with high ARE have low empirical power. Then the tests with low ARE are even less powerful, and the results of the experiment probably turn out not to be particularly informative if power comparisons between tests are the main object of interest. To avoid that, and to ensure a sufficient amount of variation in the simulation results, tests with high ARE should be designed to have fairly high power. Alternatively, to obtain interesting information about power differences, the tests should be carried out at a sufficiently large number of sample sizes.

Nonlinear Autoregressive (NLAR) model

A linearization of the model indicates that TSAY1, TSAY2, the NN test, and RESET may have power against NLAR because their auxiliary regressions contain y_{t-1}^2. That of RESET also contains y_{t-1}^4, albeit in linear combinations with other powers of y_{t-1}. That the regressors give power is seen from Table 4 in LWG. However, TSAY2 has low power because it contains a large number of redundant regressors.

5 Final Remarks

I have shown that the simulation results in LWG can to a large extent be explained by linearizing some of the models and applying the concept of ARE. The following conclusions emerge. First, the mediocre performance of the TSAY1 against TAR and SGN models is due to the particular parametrization of these models. The designs in LWG did not include a two-regime TAR model, against which the Tsay test usually has power; see Luukkonen, Saikkonen, and Teräsvirta (1988b) and Petruccelli (1990) for examples. Second, the results of LWG speak against the principal component RESET, not RESET as such. For Block1 models, the original RESET with f_{t-1}^2 and f_{t-1}^3 as regressors is almost the same as the NN test in Teräsvirta et al. (1993), and should therefore have excellent power against the TAR and SGN models. Third, the design of the bilinear alternatives is not informative enough for evaluating the relative performance of the BDS test. In general, whenever the BDS test has high power, then at least one test with zero ARE with respect to the test based on the correct alternative also is quite powerful. The power of several tests against bilinearity crucially depends on the coefficient of y_{t-1}, which equals zero in all simulations of LWG. Luukkonen et al. (1988a) have made a similar point.

Finally, an LM or LM-type test against the appropriate alternative would be a useful addition to any Monte Carlo design. This is because it gives an upper bound to the empirical power for that particular design, and thus helps to better assess the relative performance of the other tests. In some cases LWG have already included it: an example is TSAY1 in the SQ model.

References

Brock, W.A. and S.M. Potter (1993), "Nonlinear Time Series and Macroeconometrics." In: G.S. Maddala, C.R. Rao, and H.D. Vinod, eds., *Handbook of Statistics*, vol. 11. Amsterdam: North-Holland, pp. 195–229.

Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

Keenan, D.M. (1985), "A Tukey Nonadditive Type Test for Time Series Nonlinearity," *Biometrika* 72, 39–44.

Lee, T.-H., H. White, and C.W.J. Granger (1993), "Testing for Neglected Non-Linearity in Time Series Models. A Comparison of Neural Network Methods and Alternative Tests," *Journal of Econometrics* 56, 269–290.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988a), "Testing Linearity in Univariate Time Series Models," *Scandinavian Journal of Statistics* 15, 161–175.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988b), "Testing Linearity against Smooth Transition Autoregressive Models," *Biometrika* 75, 491–499.

Pearson, E.S. and H.O. Hartley, eds. (1972), *Biometrika Tables for Statisticians*, vol. 2. Cambridge: Syndics of the Cambridge University Press.

Petruccelli, J. (1990), "A Comparison of Tests for SETAR-Type Non-Linearity in Time Series," *Journal of Forecasting* 9, 25–36.

Ramsey, J.B. (1969), "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis," *Journal of the Royal Statistical Society, B*, 31, 350–371.

Saikkonen, P. (1989), "Asymptotic Relative Efficiency of the Classical Test Statistics under Misspecification," *Journal of Econometrics* 42, 351–369.

Teräsvirta, T., C.-F. Lin, and C.W.J. Granger (1993), "Power of the Neural Network Linearity Test," *Journal of Time Series Analysis* 14, 209–220.

Tsay, H. (1986), "Non-Linearity Tests for Time Series," *Biometrika* 73, 461–466.

White, H. (1987), "Specification Testing in Dynamic Models." In: T.F. Bewley, ed., *Advances in Econometrics, Fifth World Congress of the Econometric Society*, vol. 1. Cambridge: Cambridge University Press, pp. 1–58.

White, H. (1989), "Some Asymptotic Results for Learning in Single Hidden-Layer Feedforward Network Models," *Journal of the American Statistical Association* 84, 1003–1013.

Appendix: Computing Asymptotic Relative Efficiency

A1. True model $y_t = ay_{t-1} + c_{21}u_{t-2}y_{t-1}$; the inappropriate model $y_t = ay_{t-1} + u_t + \psi u_{t-1}y_{t-1}$

In all these cases, σ^2 may be regarded as known because the information matrix of σ^2 and the remaining parameters is block diagonal such that one block contains σ^2 and the other the rest of the variables. Then the components of $\bar{k}_t$ are:

$$\partial \bar{p}_t(a, 0)/\partial a = -\sigma^{-2}\tilde{u}_t y_{t-1} \tag{A.1}$$

$$\partial \bar{p}_t(a, 0)/\partial \psi = -\sigma^{-2}\tilde{u}_t \tilde{u}_{t-1} y_{t-1} \tag{A.2}$$

$$\partial \bar{p}_t(a, 0)/\partial \phi = -\sigma^{-2}\tilde{u}_t \tilde{u}_{t-2} y_{t-1} \tag{A.3}$$

where $\tilde{u}_t = y_t - ay_{t-1}$. Then $r = s = 1$, $\Sigma_{ij} = \sigma_{ij}$ is a scalar, and

$$\Sigma = \begin{pmatrix} \sigma_{aa} & \sigma_{a\psi} & \sigma_{a\phi} \\ & \sigma_{\psi\psi} & \sigma_{\psi\phi} \\ & & \sigma_{\phi\phi} \end{pmatrix} = \sigma^{-2} \begin{pmatrix} \sigma_y^2 & 0 & 0 \\ & \sigma^2\sigma_y^2(3 - 2a^2) & 2a\sigma^4 \\ & & \sigma^2\sigma_y^2(1 + 2a^2 - 2a^4) \end{pmatrix}$$

where $\sigma_y^2 = Ey_t^2 = \sigma^2/(1 - a^2)$. Thus $\sigma_{\psi\psi \cdot a} = \sigma_{\psi\psi}$ and $\sigma_{\phi\phi \cdot a} = \sigma_{\psi\psi}$, so that $\text{ARE}_\psi(a) = \sigma_{\psi\phi}^2/\sigma_{\psi\psi}\sigma_{\phi\phi} = 4a^2(1 - a^2)^2/\{(3 - 2a^2)(1 + 2a^2 - 2a^4)\}$.

A2. True model $y_t = ay_{t-1} + c_{21}u_{t-2}y_{t-1}$; the inappropriate model the auxiliary regression of TSAY1

The auxiliary regression of TSAY1 is:

$$y_t = ay_{t-1} + a_2 y_{t-1}^2 + u_t.$$

Then equations (A.1) and (A.3) remain unchanged, but equation (A.2) is replaced by $\partial \bar{p}_t(a, 0)/\partial \psi = -\sigma^{-2}\hat{u}_t y_{t-1}^2$. It follows that

$$\Sigma = 4\sigma^{-2} \begin{pmatrix} \sigma_y^2 & 0 & 0 \\ & 3\sigma_y^4 & 3a\sigma^2\sigma_y^2 \\ & & \sigma^2\sigma_y^2(1 + 2a^2 - 2a^4) \end{pmatrix}$$

which leads to $\text{ARE}_\psi(a_1) = 3a^2(1 - a^2)/(1 + 2a^2 - 2a^4)$.

A3. True model $y_t = a_1 y_{t-1} + a_3 y_{t-1}^3 + u_t$; the inappropriate model equation (A.4)

The inappropriate model is based on WHITE3, and can be written as:

$$y_t = a_1 y_{t-1} + u_t + \psi_1 u_{t-1} + \psi_2 y_{t-1} u_{t-1} + \psi_3 y_{t-2} u_{t-1} + \psi_4 y_{t-1} y_{t-2} u_{t-1}, \tag{A.4}$$

where $u_t \sim \text{nid}(0, \sigma^2)$. This yields $\sigma^2 \Sigma_{aa} = \sigma_y^2$, $\sigma^2 \Sigma_{a\phi} = 3\sigma_y^4$, $\sigma^2 \Sigma_{\phi\phi} = 15\sigma_y^6$, and

$$\sigma^2 \Sigma_{\phi\phi \cdot a} = 6\sigma_y^6. \tag{A.5}$$

Furthermore,

$$\sigma^2 \Sigma_{\psi\psi} = \begin{pmatrix} \sigma^2 & 0 & 0 & a_1\sigma^2\sigma_y^2 \\ & 3\sigma^4 + a_1^2\sigma^2\sigma_y^2 & a_1\sigma^2\sigma_y^2 & 0 \\ & & \sigma^2\sigma_y^2 & 0 \\ & & & 3\sigma^2\sigma_y^4 \end{pmatrix}$$

$$\sigma^2 \Sigma_{a\psi} = (\sigma^2, 0, 0, 2a_1\sigma^2\sigma_y^2), \quad \sigma^2 \Sigma_{\phi\psi} = (3\sigma^2, \sigma_y^2, 0, 0, 12a_1\sigma^2\sigma_y^4)$$

$$\sigma^2 \Sigma_{\phi\psi \cdot a} = \sigma^2(\Sigma_{\phi\psi} - \Sigma_{\phi a}\Sigma_{aa}^{-1}\Sigma_{a\psi}) = (0, 0, 0, 6a_1\sigma^2\sigma_y^4).$$

Thus the only term that contributes to the asymptotic local power of WHITE3 in this case is $y_{t-1}y_{t-2}\hat{u}_{t-1}$. Straightforward computation shows that the southeast corner element of $\sigma^{-2}\Sigma_{\psi\psi \cdot a}$ equals $(2\sigma^2\sigma_y^4)^{-1}$, so that

$$\lambda_\psi(a_1, \delta) = \delta' \Sigma_{\phi\psi \cdot a} \Sigma_{\psi\psi \cdot a}^{-1} \Sigma_{\psi\phi \cdot a}^{-1} \delta = 18 a_1^2 \sigma_y^4 \sigma^2.$$

Using equation (A.5), one obtains $\lambda_\psi(a_1, \delta) = 6\sigma^{-2}\sigma_y^6\delta^2$, so that $\lambda_\psi(a_1, \delta)/\lambda_\phi(a_1, \delta) = 3a_1^2(1 - a_1^2)$ and

$$\text{ARE}_\psi(\delta, a_1, \alpha, \beta) = 3a_1^2(1 - a_1^2)\{d(1, \alpha, \beta)/d(4, \alpha, \beta)\}.$$

Optimal Cycles and Chaos: A Survey

Kazuo Nishimura
Institute of Economic Research
Kyoto University
Japan
nishimura@kier.kyoto-u.ac.jp

Gerhard Sorger
Department of Economics
University of Vienna
Austria
a4411dal@helios.eduz.univie.ac.at

Abstract. *This paper surveys the literature on cyclical and chaotic equilibrium paths in deterministic optimal growth models with infinitely lived agents. We focus on discrete time models but also briefly mention results for continuous time models. We start by reviewing those results that have been proved for optimal growth models in reduced form. Then we discuss results for two-sector optimal growth models in primitive form. Finally, we summarize a few results that have been obtained for other variants of the model, including models with recursive preferences and models with heterogeneous agents.*

Acknowledgment. We thank William A. Brock, Cars Hommes, Luigi Montrucchio, and Makoto Yano for useful discussions and comments.

1 Introduction

Both the interest and the research activity in economic growth theory has increased dramatically during the last decade. One can explain this development in at least two different ways, namely, (i) that the models studied in the recent endogenous growth literature are capable of explaining long-run growth and technological advancement in a more convincing way than the models from 20 or 30 years ago, and (ii) that new mathematical concepts and methods have been developed that allow us to study growth models with dynamically complicated optimal paths or equilibrium paths. In this article we try to survey a number of new results for optimal growth models with infinitely lived households, which can be attributed to fact (ii) from above.[1] Particular emphasis will be put on results that are concerned with cyclical and chaotic optimal growth paths.

The importance of the theory of nonlinear dynamical systems for the study of economic processes has already been noted by many authors. Nonlinear dynamics are ubiquitous in economics, and the reader is referred to the following survey articles or books to get an impression of how much work has been done in

[1] A recent and comprehensive survey of developments triggered by fact (i) can be found in Barro and Sala-i-Martin (1995).

this area over the last couple of years: Benhabib and Baumol (1989), Boldrin and Woodford (1990), Brock (1993), Brock and Dechert (1991), Brock and Malliaris (1989), Chiarella (1990), Day (1994), Frank and Stengos (1988), Lorenz (1989), Medio (1993), and Scheinkman (1990). Although models from almost any field in economics have been analyzed by methods from nonlinear dynamical systems theory, optimal growth theory is probably the one field that has received the most attention. This is not surprising, since one of the main objectives of growth theory is the explanation of the fluctuations of economic variables around their general upward trends. For the purpose of this explanation, endogenous cycles and chaos seem to be at least as appropriate as the exogenous shocks commonly used in real business-cycle theory.[2]

The two most commonly considered paradigms in the theory of economic growth are the overlapping generations model going back to Samuelson (1958) and Diamond (1965), and the model with infinitely lived households originating from Ramsey (1928). In this survey we exclusively consider models of the second variety.[3] Among the models using infinitely lived agents we can again distinguish between those that use a continuous-time framework and those that use discrete time periods. This distinction is not innocuous for the subject of the survey, because it is well known that complicated dynamics can arise more easily in low-dimensional systems of difference equations than in corresponding systems of differential equations. We shall focus our attention on discrete time models and mention possible generalizations or extensions to the continuous-time framework only briefly in Section 4.3.

Most of the results on cycles and chaos in economic growth theory deal with models in which the dynamics can be reduced to a one-dimensional difference equation. This emphasis on one-dimensional dynamics is also reflected in the present survey. It seems, therefore, worthwhile to point out those results that deal with higher dimensional dynamical systems. The general possibility theorem (Theorem 2.2) from Boldrin and Montrucchio (1986) and the result on topological entropy (Theorem 2.3) from Montrucchio and Sorger (1996) hold for state spaces of arbitrary (finite) dimension, and some of the heterogeneous agent models discussed in Section 4.1 lead to two-dimensional difference equations. As will be discussed in Section 4.3, one also needs higher dimensional state spaces to generate cycles and chaos in continuous-time models.

The general outline of the paper is as follows. In Section 2 we discuss results for models in reduced form. These models do not contain any control variables such as consumption rates or investment rates, but are stated in state variables (capital stocks) only. The advantage of the reduced form is that it has a simpler mathematical structure than nonreduced models, and is therefore more amenable to a clear-cut characterization of those properties that generate cycles and chaos. We start with definitions and basic properties of reduced utility function models in Section 2.1, discuss optimal cycles in Section 2.2, and optimal chaos in Section 2.3. Section 2.4 studies the influence of the rate of depreciation on the structure of optimal growth paths, whereas Section 2.5 analyzes the relationship between the size of the discount factor and the occurrence of complicated dynamics. Section 3 deals with nonreduced models. Since cycles and chaos cannot occur in the standard one-sector model (see Dechert and Nishimura [1983] and Dechert [1984]), a nonreduced optimal growth model that can generate complicated dynamics must have at least two production sectors. The existing literature on cyclical or chaotic optimal growth has focused almost entirely on the simplest such case, which is the case of a two-sector model, and we follow this tradition. We introduce the basic framework in Section 3.1, discuss optimal cycles in Section 3.2, and chaos in Section 3.3. Section 4 summarizes results for models that did not fit easily into Sections 2 or 3. More specifically, these are models with heterogeneous agents (Section 4.1), models with nonadditive utility functions (Section 4.2), and continuous-time models (Section 4.3). Finally, Section 5 presents some concluding remarks.

2 Reduced-Form Utility Function Models

2.1 Definitions and assumptions

Reduced-form utility function models have been used by many authors because of their simple mathematical structure and their wide applicability in economics. For a comprehensive survey of methods for and applications of such models, we refer to McKenzie (1986) and Stokey and Lucas (1989). In this subsection we briefly summarize the definitions and results that are needed in the remainder of the paper.

[2]See Stadler (1994) for a recent survey of real business-cycle theory.

[3]The occurrence of chaos and cycles in models with overlapping generations of finitely lived agents has been discussed, for example, in Grandmont (1985).