University of
California, Davis

IMPROVED TESTING AND SPECIFICATION OF SMOOTH TRANSITION REGRESSION MODELS

Alvaro Escribano & Oscar Jordá

Working Paper #97-26

Department
of Economics

Improved Testing and Specification of Smooth Transition Regression Models

Alvaro Escribano
Universidad Carlos III de Madrid

Oscar Jordá
University of California, Davis

Working Paper Series No. 97-26
October, 1997

The first author acknowledges support from Spanish DGCIYT, PB95-0298. The second author acknowledges support from "La Caixa" and the Bank of Spain. We wish to thank R. F. Engle, J. Gonzalo, C. W. J. Granger, J. D. Hamilton, M. H. Pesaran and T. Terävirta for useful comments and suggestions.

Abstract

This paper extends previous work in Escribano and Jordá (1997) and introduces new LM specification procedures to choose between Logistic and Exponential Smooth Transition Regression (STR) Models. These procedures are simpler, consistent and more powerful than those previously available in the literature. An analysis of the properties of Taylor approximations around the transition function of STR models permits one to understand why these procedures work better and it suggests ways to improve tests of the null hypothesis of linearity versus the alternative of STR-type nonlinearity. Monte-Carlo experiments illustrate the performance of the different tests introduced. The new procedures are then implemented on a study of the dynamics of the U.S. unemployment rate.

1 Introduction

This paper provides tools for the empirical practitioner interested in modelling Smooth Transition Regression (STR) models. Nonlinear time series models are being used more frequently in empirical applications, leaving the researcher with a virtual infinity of models and specifications from which to choose. However, STR models are a general class of state-dependent, reduced form, non-linear time series models in which the transition between states is generally endogenously generated. They encompass as particular cases the Smooth Transition Autoregressive (STAR), the Exponential Autoregressive (EAR), the Threshold Autoregressive (TAR) and the SETAR models.[1] Together with Hamilton's Regime Switching model[2] (where the transition between regimes is exogenously generated by a Markov chain), state-dependent models have proven particularly useful in modelling the asymmetric behavior of economic fluctuations.[3]

Even after restricting attention to a certain class, the rich parametrization and flexibility of these models makes the task of specifying the model complicated. Model building usually starts by performing a nonlinearity test. If there is not enough evidence of nonlinearity, there is no reason to pursue a model that is much more difficult to specify, estimate and evaluate. Chan and Tong (1986) discuss the possibility of using likelihood ratio test statistics for testing linearity against SETAR models. The drawback of this approach is that the distribution of the statistic has to be determined by simulation for each application. Based on work by

[1] See Teräsvirta (1994), Haggan and Ozaki (1981), Tsay (1989) and Chan and Tong (1986).

[2] See Hamilton (1989).

[3] A variety of recent studies justify this assertion. See Neftçi (1984), Sichel (1989) and Rothman (1991) for example.

Tsay (1986). Luukkonen et al. (1988a) introduce a set of Lagrange Multiplier (LM) type tests that have asymptotic χ^2 distributions. Saikkonen and Luukkonen (1988) considered LM tests against bilinear and EAR alternatives. Teräsvirta (1994) uses these procedures in several stages of the specification of STAR models.

This paper introduces two new main results, which rely on Taylor series approximations to the transition function (between states) around the scale parameter.[4] The results, derived in the general context of STR models (which encompass STAR-based tests as a particular case), are an extension of previous work in Escribano and Jordá (1997). First, we introduce a new specification strategy extended to include the choice between logistic and exponential STR models. Our selection procedure has higher correct selection frequencies of the right model, avoids the pitfalls of the rules proposed in Teräsvirta (1994) and is much simpler to apply. Second, we suggest that in some cases, tests of the null hypothesis of linearity against STR-type nonlinearity[5] should be augmented to include up to fourth order terms. Including these terms is necessary to gain power against alternatives where an exponential STR might be involved. We complement this analysis with an exhaustive study of the dynamics of STR models and the shapes of transition functions in practice through simulations. The paper also provides ample Monte-Carlo evidence in support of these claims (providing several additional cases to those available in Escribano and Jordá (1997)) and shows how these procedures work in practice with a different empirical example - a brief study of the dynamics of U.S. Unemployment.

[4] Luukkonen et al. (1988a) and Teräsvirta (1994), based on Davies (1977)) introduced this solution for Smooth Transition Autoregressive models.

[5] Note that this alternative is an assumption imposed by the analysist. The test has power against alternatives other than STR.

The paper is organized as follows: Section 2 briefly defines and introduces the basic properties of STR models; Section 3 presents the tests of nonlinearity and their properties; Section 4 discusses the selection procedure proposed by Teräsvirta (1994) and then introduces the new alternative; Section 5 presents Monte Carlo simulations of the new tests; Section 6 applies the new techniques to the U.S. unemployment rate; and Section 7 concludes.

2 Smooth Transition Regression (STR) Models

Consider the following STR model:

$$y_t = \pi' \underline{x}_t + \theta' \underline{x}_t F(z_{t-d}, \gamma, c) + u_t \tag{1}$$

where y_t is a scalar, $\underline{x}_t = (1, y_{t-1}, ..., y_{t-r}; \underline{w}_{t-1}, ..., \underline{w}_{t-q})' = (1, \tilde{\underline{x}}_t)'$ where $\underline{w}_t$ is a vector of exogenous variables; z_{t-d} will be a scalar, although in general, it could be a vector. Usually $z_{t-d} = y_{t-d}$ where d is the delay parameter which satisfies $1 \leq d \leq p$ for $p = \max(r, q)$ and is assumed known.[6] Note that when $z_{t-d} = y_{t-d}$ and $\underline{w}_t = \underline{0}$ we have a conventional STAR model. $\pi' = (\pi_0, \pi_1, ..., \pi_p) = (\pi_0, \tilde{\pi}')$; $\theta' = (\theta_0, \theta_1, ..., \theta_p) = (\theta_0, \tilde{\theta}')$. u_t is a martingale difference sequence with constant variance[7] and y_t is assumed stationary and ergodic. The function $F(z_{t-d}, \gamma, c)$ is at least fourth order, continuously differentiable with respect to the scale parameter, γ.

The transition function, $F(.)$, is traditionally chosen to be either a logistic or an exponential distribution function.[8] The exponential transition function is:

[6] See Teräsvirta (1994) when d is unknown.

[7] This assumption is sufficient to derive the LM tests below. See White (1984).

[8] In general, it could be any continuously differentiable function such that $0 \leq F(z_{t-d}, \gamma, c) \leq 1$ for all z_{t-d}, $1 \leq d \leq p$, $\gamma \neq 0$ and c. However, in practice the logistic and the exponential permit the neccesary

$$F(z_{t-d},\gamma,c)=\left[1-\exp\left\{-\gamma(z_{t-d}-c)^2\right\}\right] \tag{2}$$

When $\theta_0 = c; \underline{w}_t = \underline{0}\ \forall t$ and $z_{t-d} = y_{t-d}$, the corresponding ESTR model is reduced to the exponential autoregressive model (EAR). The logistic transition function[9] is:

$$F(z_{t-d},\gamma,c)=\left[\{1+\exp\left(-\gamma(z_{t-d}-c)\right\}^{-1}-\frac{1}{2}\right] \tag{3}$$

In order to illustrate the properties and shapes of the logistic and exponential transition functions, we simulate two specifications from Teräsvirta (1994). This exercise will provide a better understanding of the details involved in the specification and nonlinearity testing of STR models. Consider the following simulation:

$$y_t = 1.8y_{t-1} - 1.06y_{t-2} + (\theta_0 - 0.9y_{t-1} + 0.795y_{t-2})\,F\left(y_{t-1},\gamma,c\right) + u_t \tag{4}$$

where:

$$F(z_{t-d},\gamma,c)=\left[1+\exp\left(-1000(y_{t-1}-c)^2\right)\right] \qquad \text{(ESTAR)}$$

and

$$F(z_{t-d},\gamma,c)=\left[1+\exp\left(-100(y_{t-1}-c)\right)\right]^{-1} \qquad \text{(LSTAR)}$$

with u_t i.i.d. $N(0, 0.02^2)$ and $T = 300$. Figure 2.1 depicts 3 graphs of the plots of the data, y_{t-1}, and $F(y_{t-1},\gamma,c)$ for the following cases: (1) $\theta_0 = c = 0$ and ESTAR specification:

degree of generality.

[9] The term $\frac{1}{2}$ is added here for convenience but does not affect the results.

(2) $\theta_0 = 0.04$; $c = 0.02$ and ESTAR specification; and (3) $\theta_0 = 0.02$; $c = 0$ and LSTAR specification. The plots on the first column depict the time series of the data and the values of the transition function. The plots in the second column depict the scatter plots of the data versus the corresponding values of the transition function, thus illustrating its shape.

Consider case (1). The data is symmetrically distributed at each side of the threshold c. When $|y_{t-1}| \gg c$, then $F(y_{t-1}, \gamma, c) \simeq 1$, and the dynamic behavior of the data is characterized by the "upper" linear regime. When $|y_{t-1}|$ is "close" to c then $F(y_{t-1}, \gamma, c) \simeq 0$ and the dynamics of the "lower" linear regime dominate. The speed with which the transition from one regime to the other takes place is set by the value of the scale parameter, γ. The shape of the plot of the values of the transition function and the data that we obtain in case (1) therefore represent that of the usual ESTAR model.

Now consider case (3), the graph of the typical transition function for a LSTAR model. When $y_{t-1} \ll c$, then $F(y_{t-1}, \gamma, c) \simeq 0$ -- what we denominated, the "lower" regime. Conversely, when $y_{t-1} \gg c$, then $F(y_{t-1}, \gamma, c) \simeq 1$ (the "upper" regime). γ again determines how smooth the transition between regimes is. In the extreme case where $\gamma \longrightarrow \infty$, then we get a particular case of LSTAR -- the Threshold Autoregressive (TAR) model.

Finally, the most interesting scenario is that of case (2). The model is a ESTAR but notice that the data, y_{t-1} tends to cluster around the "upper" regime, to the right of the threshold c. In fact, the graphs for cases (2) and (3) are rather similar although the dynamic behavior of each model is quite different. This asymmetric behavior in cases when $\theta_0 \neq 0$ and/or $c \neq 0$ will influence both the ability to specify and test for a STAR model as we shall see in the next sections.

3 Testing Linearity Against STR Models

Testing linearity against STR-type nonlinearity implies testing the null hypothesis $H_0 : \theta' = \underline{0}$ in equation 1. However, under the null, the parameters γ and c are not identified, that is, they can take any value. Alternatively, we could choose as our null hypothesis $H_0 : \gamma = 0$ in which case neither c nor θ' would be identified. Davies (1977) showed that conventional maximum likelihood theory is not directly applicable to this problem. A solution proposed in Luukkonen et al. (1988a) and adopted in Teräsvirta (1994), is to replace $F(z_{t-d}, \gamma, c)$ in equation 1 with a suitable Taylor series approximation. Under the null of linearity, the LM test is shown to possess asymptotically the usual χ^2 distribution.[10] In practice, the test is performed on the following auxiliary regression:

$$y_t = \pi' \underline{x}_t + \left[\theta' \underline{x}_t \left\{\gamma F_\gamma(z_{t-d}, \gamma = 0, c)\right\}\right] + v_{1t} \tag{5}$$

where $F_\gamma(.)$ indicates the first derivative of F with respect to γ — $\gamma F(.)$ is obviously the first term of the Taylor approximation of F around $\gamma = 0$. In particular, $F_\gamma(.)$ for the logistic transition function is:

$$F_\gamma(z_{t-d}, \gamma = 0, c) = \frac{1}{4}\left(z_{t-d} - c\right) \tag{6}$$

which substituted into 5 yields:

$$y_t = \pi' \underline{x}_t + \frac{1}{4}\gamma\theta' \underline{x}_t \left(z_{t-d} - c\right) + v_{1t} \tag{7}$$

[10] The delay parameter, d, is usually unknown. Based on Tsay (1989), Teräsvirta (1994) proposes choosing d such as to minimize the p-value of the nonlinearity test.

The parameters under this specification cannot be identified. After combining terms, we get the final version of the auxiliary regression:

$$y_t = \delta_0 + \widetilde{\delta}'\widetilde{\underline{x}}_t + \beta_1'\widetilde{\underline{x}}_t z_{t-d} + v_{1t} \tag{8}$$

where $\delta_0 = \left(\pi_0 - \frac{1}{4}\gamma c\theta_0\right)$; $\widetilde{\delta}' = \left(\widetilde{\pi}' - \frac{1}{4}\gamma c\widetilde{\theta}'\right)$ where the d^{th} term is $\delta_d = \left(\pi_d + \frac{1}{4}\gamma c(\theta_0 - \theta_d)\right)$; $\beta_1' = \frac{1}{4}\gamma\widetilde{\theta}'$. The null hypothesis of linearity therefore becomes $H_0' : \beta_1' = \underline{0}$. Note that equation 8 is explosive and generally not a meaningful time series model (see Granger and Andersen (1978)[11]). Luukkonen et al. (1988a) realized that this test would have low power against alternatives where $\widetilde{\theta}'$ is "small" and θ_0 is "large" in absolute value since β_1' does not include the θ_0 coefficient — "detecting a shift in the mean" problem.

To overcome this difficulty, they proposed to approximate the transition function with a third order Taylor series expansion, that is:

$$y_t = \pi'\underline{x}_t + \frac{1}{4}\gamma\theta'\underline{x}_t\left(z_{t-d} - c\right) + \frac{1}{48}\gamma^3\theta'\underline{x}_t\left(z_{t-d} - c\right)^3 + v_{3t} \tag{9}$$

where it is important to note that the square powers of the approximation are identically zero. Recombining in terms of identified parameters, we obtain:

$$y_t = \delta_0 + \widetilde{\delta}'\widetilde{\underline{x}}_t + \beta_1'\widetilde{\underline{x}}_t z_{t-d} + \beta_2'\widetilde{\underline{x}}_t z_{t-d}^2 + \beta_3'\widetilde{\underline{x}}_t z_{t-d}^3 + v_{3t} \tag{10}$$

where now $\beta_{2d} = \frac{1}{48}\gamma^3\theta_0$ and thus we avoid the "detecting a shift in the mean" problem.[12]

The null hypothesis now becomes $H_0''' : \beta_1' = \beta_2' = \beta_3' = \underline{0}$. Following Saikkonen and

[11] Also note that the alternative now includes models other than STR.

[12] The details of the correspondence between the parameters in 9 and 10 is derived in the Appendix.

Luukkonen (1988), Teräsvirta et al. (1994) and Teräsvirta (1994), a convenient procedure for computing the LM statistic by OLS consists on estimating the auxiliary regression 10 under the null hypothesis and compute the sum of squared residuals, SSR_0. Next, estimate 10 under the alternative hypothesis and compute the sum of squared residuals, SSR_1. Finally, the statistic, $T(SSR_0 - SSR_1)/SSR_0$ is shown to have asymptotically a χ^2 distribution under the null. It is usually recommended to use the approximation given by the F distribution because of the good size and power properties of the test in small samples. We will call the nonlinearity test based on equation 10, NL3 for short.

Note that when the model is an ESTR, we have that $F_\gamma(z_{t-d}, \gamma = 0, c) = (z_{t-d} - c)^2$ which, after substituting into 5 becomes:

$$y_t = \gamma c^2 \theta_0 + \pi' \underline{x}_t - 2\gamma c \widetilde{\theta}' \underline{\widetilde{x}}_t z_{t-d} + \gamma \widetilde{\theta}' \underline{\widetilde{x}}_t z_{t-d}^2 + v_{2t} \tag{11}$$

which in terms of identified parameters becomes:

$$y_t = \delta_0 + \widetilde{\delta}' \underline{\widetilde{x}}_t + \beta_1' \underline{\widetilde{x}}_t z_{t-d} + \beta_2' \underline{\widetilde{x}}_t z_{t-d}^2 + v_{2t} \tag{12}$$

where in particular $\beta_{1d} = \gamma \theta_0$ and therefore we do not need to pursue further terms of the expansion since we are able to identify shifts in the mean via the term β_{1d}. As a result, equation 10 is a valid nonlinearity test for either LSTR or ESTR alternatives.

3.1 Properties of the Taylor Approximation

The transition function of a STR model exhibits two important features. First, the logistic function (see equation 3, figure 2.1 case (3)) has a single inflection point while the exponential function (see equation 2, figure 2.1, cases (1) and (2)) has two inflection points. Second, the even powers of the Taylor expansion of a logistic function are all zero. Meanwhile, all the odd powers of the Taylor expansion of an exponential function are zero. The first feature suggests ways to improve the nonlinearity test NL3 of the previous section. In Section 4 we use the second feature to introduce a new specification procedure to choose between LSTR and ESTR models.

The immediate consequence resulting from the difference in shape between the logistic and the exponential function is that we need a second order Taylor series expansion in order to be able to capture its two inflection points. This is particularly the case when γ is "small" in absolute value (smooth transition) and/or the variance of the residuals, u_t is "large" such that a reasonable number of observations are driven to the "upper" regime. A solution to this problem is to expand the auxiliary regression 10 to include the terms resulting from the second order expansion of the exponential. In particular:

$$y_t = \pi' \underline{x}_t + \theta' \underline{x}_t \left[(z_{t-d} - c)^2 \left(\gamma + \gamma^2 \right) - \frac{1}{2} (z_{t-d} - c)^4 \gamma^2 \right] + v_{4t} \tag{13}$$

which rewritten in terms of identified parameters[13] simply means augmenting the auxiliary regression in 10 as follows:

$$y_t = \delta_0 + \tilde{\delta}' \tilde{\underline{x}}_t + \beta_1' \tilde{\underline{x}}_t z_{t-d} + \beta_2' \tilde{\underline{x}}_t z_{t-d}^2 + \beta_3' \tilde{\underline{x}}_t z_{t-d}^3 + \beta_4' \tilde{\underline{x}}_t z_{t-d}^4 + v_{4t} \tag{14}$$

Now the null hypothesis becomes $H_0^{IV}: \beta_1' = \beta_2' = \beta_3' = \beta_4' = \underline{0}$. We call this test NL4 for short.[14] The computation of the test is parallel to that of NL3, where either the χ^2 or the

[13] The details of how the parameters in 13 are transformes into 14 are available in the Appendix.

[14] When $z_{t-d} = y_{t-d}$, this test is similar to a higher order RESET test.

F version of the test can be used.

An alternative to the above LM tests (namely NL3 and NL4) is to use the Wald test of Hansen (1996). This procedure approximates unknown limiting distributions by generating p-values based on simulation methods. Pesaran and Potter (1997) adapt this strategy in an interesting application to their floor and ceiling model. However, this alternative involves additional elements of complexity which are easily avoided with the type of LM testing discussed here.

Testing in practice involves several important steps such as choice of lag length of the linear ARX model,[15] choice of the delay parameter,[16] d, and others which are all well documented in Teräsvirta (1994). However, an important empirical question is to realize that NL4 requires p extra regressors with respect to NL3, and that NL3 requires $2p$ extra regressors with respect to NL1. Lack of parsimony is particularly important with small sample sizes and/or when the order of the ARX polynomial, p, is high since it reduces the power of the test. Luukkonen et al. (1988a) recognized this lack of parsimony and suggested an augmented fist order procedure based on equation 8. This consisted on using the following auxiliary regression instead of 10:

$$y_t = \delta_0 + \tilde{\delta}' \underline{\tilde{x}}_t + \beta_1' \underline{\tilde{x}}_t z_{t-d} + \beta_{2d} z_{t-d}^3 + \beta_{3d} z_{t-d}^4 + e_{1at} \qquad (15)$$

which we will call NL3A. In the same spirit, we can augment the first order procedure further to take into account the results that led us to equation 14. This means including the term

[15] Too few terms can cause false rejections of linearity, excesive terms can undermine the power of the nonlinearity test.

[16] We mentioned this in the previous section.

$\beta_{4d} z_{t-d}^5$ in the auxiliary regression 15 to obtain the equivalent augmented first order version of the test which we will call NL4A. The null hypothesis simply becomes $H_0^{IVA} : \underline{\beta}_1' = \underline{0}$; $\beta_{2d} = \beta_{3d} = \beta_{4d} = 0$ which only requires $p + 3$ terms. The intuition behind this test is that the parameters β_{jd} for $j = 2, 3, 4$ collect the effect of a shift in the mean in the nonlinear regime.

4 Choosing between LSTR or ESTR

4.1 Teräsvirta's (1994) Selection Procedure

Upon rejecting the null hypothesis of linearity (with any of the tests suggested in the previous section), one might consider using a STR model as a useful nonlinear alternative — recall that the tests for nonlinearity have power against forms of nonlinearity other than the STR-type. Teräsvirta (1994) introduced the following model selection procedure (which we will denominate TP for short), based on equation 10, reproduced here for convenience:

$$y_t = \delta_0 + \widetilde{\underline{\delta}}' \widetilde{\underline{x}}_t + \underline{\beta}_1' \widetilde{\underline{x}}_t z_{t-d} + \underline{\beta}_2' \widetilde{\underline{x}}_t z_{t-d}^2 + \underline{\beta}_3' \widetilde{\underline{x}}_t z_{t-d}^3 + v_{3t}$$

1. Test the null hypothesis: $H_{03} : \underline{\beta}_3' = \underline{0}$ versus the alternative, $H_{13} : \underline{\beta}_3' \neq \underline{0}$ with an F-test (F_3). According to Teräsvirta, rejection of this null would imply rejection of the ESTR specification since cubic powers of z_{t-d} in a first order approximation of $F(z_{t-d}, \gamma, c)$ are 0.

2. Test the null hypothesis: $H_{02} : \underline{\beta}_2' = \underline{0} | \underline{\beta}_3' = \underline{0}$ with an F-test (F_2). Teräsvirta's reasoning is that the terms z_{t-d}^2 of a first order approximation to a logistic function are zero when $c = \theta_0 = 0$ (see equation 6). However, these terms will be non-zero in the ESTR case (except in the unlikely case that $\widetilde{\underline{\theta}}' = \underline{0}$). Failure to reject this null is taken as evidence

in favor of a LSTR model. Nevertheless, rejection of H_{02} is not very informative one way or the other.

3. Test the null hypothesis: $H_{01}: \beta_1' = \underline{0} \mid \beta_2' = \beta_3' = \underline{0}$ with an F-test (F_1). Following Teräsvirta, failing to reject H_{01} after rejecting H_{02} points to a ESTR model. On the other hand, rejecting H_{01} after failing to reject H_{02} supports the choice of LSTR.

4. Note which hypotheses are rejected and compare the relative strengths of the rejections. If the model is LSTR, typically H_{01} and H_{03} are rejected more strongly than H_{02}. Therefore, if the p-value of F_2 is the smallest of F_1, F_2, F_3 select the ESTR specification, otherwise, select LSTR.

To discuss the problems and pitfalls with TP we reproduce equation 14 for a LSTR and an ESTR in terms of the Taylor expansion to the transition function, namely:

$$y_t = \pi_0 + \tilde{\pi}'\tilde{x}_t + v_1' x_t (z_{t-d} - c) + v_3' x_t (z_{t-d} - c)^3 + v_{3t} \tag{16}$$

for the logistic STR third order expansion, and:

$$y_t = \pi_0 + \tilde{\pi}'\tilde{x}_t + v_2' x_t (z_{t-d} - c)^2 + v_4' x_t (z_{t-d} - c)^4 + v_{4t} \tag{17}$$

for the second order expansion of the exponential STR. The parameters v_j' are not directly identifiable but make the discussion easier to follow.

Consider the following problems with TP. First: Whenever $c \neq 0$, expansion of the term $v_4' x_t (z_{t-d} - c)^4$ will yield non-zero $\tilde{x}_t z_{t-d}^3$ terms in the auxiliary regression 10 when the model is ESTR. This is particularly problematic if π_0 and θ_0 are also non-zero. Additionally, when the variance of the error term is "large" (causing the data to be distributed asymmetrically around c as we showed in figure 2.1, case (2)), false detection of a LSTR model will become more frequent.

The second source of problems is intrinsic to the design of the rule: The three F-tests suggested in TP are nested. This feature becomes troublesome again when $c \neq 0$. For example, if the true model is LSTR, expansion of the term $\psi_3' x_t (z_{t-d} - c)^3$ yields non-zero $\tilde{x}_t z_{t-d}^2$ terms in the auxiliary regression 10. In addition, by conditioning on $\beta_2' = \underline{0}$ in H_{02}, the terms $\tilde{x}_t z_{t-d}^2$, irrespective of whether $c \neq 0$, or not, are left to approximate the transition function — an approximation that the cubic terms presumably were successfully capturing. It is therefore unclear whether F_3 or F_2 will have the smallest p-value.

4.2 A New Selection Procedure

Consider the following example. Suppose that $c = 0$. It is clear that (based on equation 16) if the model is LSTR, the terms $\tilde{x}_t z_{t-d}^j$ for $j = 2, 4, 6, ...$ are exactly zero (i.e. $\beta_2' = \beta_4' = \underline{0}$ in equation 14). Conversely, if the model is ESTR, based on equation 17, the terms $\tilde{x}_t z_{t-d}^j$ for $j = 1, 3, 5, ...$ are exactly zero (i.e. $\beta_1' = \beta_3' = \underline{0}$ in equation 14). This suggests the following selection procedure (which we will call EJP for short) based on NL4, conditional on prior rejection of linearity:

1. Test the null $H_{0E} : \beta_2' = \beta_4' = \underline{0}$ with an F-test (F_E).

2. Test the null $H_{0L} : \beta_1' = \beta_3' = \underline{0}$ with an F-test (F_L).

3. Compare the relative strength of the rejection of each hypothesis. If the minimum p-value corresponds to F_L select LSTR, otherwise, if it corresponds to F_E, select ESTR.

Note that when $c \neq 0$, the test is still effective since we rely on testing the joint significance of linear and cubic terms relative to the joint significance of quadratic and fourth order terms, without conditioning.. In addition, EJP provides information regarding non-zero thresholds, c. Linear and cubic terms are exactly zero when $c = 0$ and the model is ESTR. Quadratic and fourth order terms are exactly zero when $c = 0$ and the model is LSTR. Therefore, rejecting H_{0L} and failing to reject H_{0E} suggests a LSTR model with $c = 0$. Conversely, rejecting H_{0E} and failing to reject H_{0L} suggests a ESTR with $c = 0$. This feature provides useful starting values in the estimation stage of the model.

5 Monte Carlo Evidence

This section examines how accurate is EJP with respect to TP (and a generalization of TP for completeness) in choosing the correct STR specification (logistic or exponential). In addition, we analyze the power properties of NL3 versus NL4. The models that we simulate here are not proposed by us but rather based on previous experiments in the literature, in particular: Luukkonen et al. (1988a,b); Saikkonen and Luukkonen (1988); and Teräsvirta (1994). The specifics and technical details of the experiments are spared to the Appendix. Here we give a general overview and concentrate on the results.

The first 100 observations of each of the series simulated are disregarded to avoid initialization problems. Each experiment is replicated 1,000 times. Some specifications allow for values of the variance of the residuals that complement those originally proposed in the literature. Whenever the specification proposed in the literature was restricted to one type of model (say an ESTR for example), we constructed its counterpart (in the example, a

LSTR) with the same choice of parameters.[17] The next section examines the accuracy of the selection procedures while Section 5.2 examines the power properties of the nonlinearity tests.

5.1 Accuracy of the EJ Selection Procedure

This section will compare the performance of EJP versus TP. In addition and for the sake of completeness, we generalize TP to account for the fact that EJP is based on equation 14 and therefore uses information not available to TP. The generalized version of TP will be denominated GTP for short and is described as follows. Conditional on rejecting linearity, use the following sequence of nested tests:

1. Test the null hypothesis $H_{04}^G : \beta_4' = \underline{0}$ with an F-test $\left(F_4^G\right)$.
2. Test the null hypothesis $H_{03}^G : \beta_3' = \underline{0} \mid \beta_4' = \underline{0}$ with an F-test $\left(F_3^G\right)$.
3. Test the null hypothesis $H_{02}^G : \beta_2' = \underline{0} \mid \beta_3' = \beta_4' = \underline{0}$ with an F-test $\left(F_2^G\right)$.
4. Test the null hypothesis $H_{01}^G : \beta_1' = \underline{0} \mid \beta_2' = \beta_3' = \beta_4' = \underline{0}$ with an F-test $\left(F_1^G\right)$.
5. If the minimum p-value from these sequence of tests corresponds to either F_4^G or F_2^G select ESTR. Otherwise, if it corresponds to either F_3^G or F_1^G select LSTR.

The simulations are done by using the general guidelines introduced in the previous section and available in the Appendix. We used NL4 as our test of nonlinearity (we also used NL3 but obtained similar results). Conditional on rejecting linearity, we then applied each of EJP, TP and GTP. Furthermore, conditional on having found the correct model, we then

[17] However, some of the values of γ had to be rescaled to obtain sensible models.

looked at the accuracy of being able to determine whether the threshold is zero. The results of these exercises are reported in Tables 5.1.1 - 5.1.5. In all, we tested three LSTR models and three ESTR models for different specifications.

The results of the experiments indicate that EJP is much more accurate that either TP or GTP. EJP's success rate always increases as the sample size increases (a highly desirable "consistency" feature — the result of the particular design of the procedure). In contrast, both TP and GTP lack this feature in some cases. For example, consider $\mu = 1$ in Table 5.1.4 . TP's correct selection frequency drops from 12.9% to 9.5% and to 3.9% as the sample size *increases* from 50 to 100 and to 200 observations respectively. Compare these numbers to EJP's selection frequency for the same case: 62.4%, 70.4% and 78.5%.

Teräsvirta (1994) recognized that TP works well when the LSTR and ESTR specifications are not close substitutes. However, TP is less effective when the two models are close substitutes and the model is ESTR. It is remarkable that the most impressive gains of using the alternative EJP occur precisely in this situation. The results are fairly conclusive: EJP outperforms TP (and GTP): it is simpler to construct (it requires two simple F-tests and a straight forward choice); and is "consistent" in the sense mentioned above.

5.2 Power Properties of the NL4 Test

The key question we examine here is whether the gains in power from adding the terms $\tilde{x}_t z_{t-d}^4$ in NL4 outweigh the losses from including p extra regressors in the auxiliary regression 14. It is clear that if the DGP is a LSTR model, we will loose power because we are including redundant regressors. If the model is ESTR we should expect to perform well whenever $c = 0$ and the data is symmetrically distributed between the upper and lower regimes (recall

Figure 2.1, case (1)). If $c \neq 0$, the benefits of including extra regressors will depend on each particular case. Tables 5.2.1 - 5.2.3 report this Monte Carlo exercise.

The simulations indicate that with large sample sizes ($T \geq 300$) there is little to no loss or gain in using NL4 rather than NL3. Both tests detect nonlinearity appropriately with the power approaching 1 in most cases. However, for smaller sample sizes, NL4 performs better, particularly when the variance is "large" and/or c and θ_0 are non-zero (asymmetric cases to those mentioned in Section 2, Figure 2.1). On the other hand, there are no significant losses of power when the true model is LSTR. One should view these results with caution. While NL4 is able to capture some nonlinearities that are hard to capture with NL3, it involves p additional regressors compared to NL3. When the lag length of the model, p, is long and the sample size small, parsimony becomes an issue regarding the power of these tests. A parsimonious alternative in these cases is the test NL4A that we introduced in Section 3.1.

6 Unemployment Dynamics in the U.S.

This section will apply the techniques developed above on an empirical study of the dynamics of U.S. Unemployment — the aim is to illustrate how the previous tests work in practice rather than producing a detailed analysis of U.S. Unemployment. The existence of asymmetries in the unemployment rate has been a debated topic in recent years. Neftçi (1984) used the theory of Markov chains and applied it to the series of the signs of the first differences of the unemployment series. Symmetry is thus defined as a situation in which next period's probability of moving from a positive to a negative sign is equal to the probability of moving from a negative to a positive sign. Sichel (1989), after correcting some errors in Neftçi (1984), found no evidence of asymmetry using a second order Markov chain, while Rothman (1991), using a first order Markov chain concluded that the unemployment series was indeed asymmetric. De Long and Summers (1986) define asymmetry as non-zero skewness in a detrended series and found unemployment to be asymmetric.

The issue of asymmetry is crucial in both theoretical and empirical studies of unemployment behavior. Theoretical explanations for why unemployment might be asymmetric vary widely. Non-competitive theories can generate asymmetries as a result of nominal and real rigidities. In Layard et al. (1991), asymmetries are generated by insider-outsider mechanisms in wage-setting. Labor turnover costs (a difference between hiring and firing costs) may also induce asymmetry as in Burgess (1988); Burgess and Dolado (1989)and Pfann and Palm (1993). According to these theories, adjustment costs depend on the size of unemployment, labor market legislation and union power. Search-theoretic models can also predict asymmetries through a reduction of search effectiveness and loss of skill induced by long-term unemployment (see Pissarides (1992)).

Following the previous discussion, it is natural to test the unemployment series for nonlinearities. Before we proceed, it is useful to consider the raw data and its properties. Figure 6.1 depicts monthly, seasonally adjusted unemployment rate from January 1948 to July 1997. Perhaps the first concern that arises from observing the graph is whether the series is stationary (recall that the nonlinearity tests are based on the assumption of stationarity and ergodicity). Both the Dickey-Fuller test and the Phillips-Perron test clearly rejected the hypothesis of a unit root. This result coincides with previous studies in the literature. However, the majority of these studies incorporate a time trend — an approach that is hard to justify on theoretical grounds.

Regarding the asymmetric behavior of unemployment, it is perhaps best to begin with a simple inspection of the graph of the series and comment its most notable characteristics. Two features are most striking: First, recessions are characterized by sharp increases of unemployment within a few months. After reaching its peak, unemployment drops quickly at first but then its descent to the original level becomes protracted and slow. Second, related to the previous feature, the duration between trough to peak is much shorter than the duration from peak to trough (note that we are referring to unemployment, not the business cycle itself). It is therefore natural to consider a two-state model such as the STAR to analyze this series.

6.1 Nonlinearity and Model Selection

The first step is to construct the linear model from which to build the nonlinearity tests and from which to compare the performance of the nonlinear alternative. The usual information criteria (Akaike's and Schwartz's) select an AR(6). However, upon inspection of the autocorrelogram of the residuals and the Ljung-Box statistic, there is evidence of left-over autocorrelation corresponding to seasonal lags 12, 13, 24 and 25 (despite the fact that we used seasonally adjusted data). Recall that omitting terms in the specification of the linear model could lead to false rejections of linearity. Therefore, the linear model is expanded to include these lags as well.

Based on this model, we then perform the tests for nonlinearity that have been presented in the previous sections, namely: NL3, NL4 and their augmented versions, NL3A and NL4A. Table 6.1.1 reports the results of these tests as well as the choice of delay parameter, d, and model selection with EJP and TP. All four nonlinearity tests (NL3, NL4, NL3A and NL4A) clearly detected nonlinearity for $d = 6$. With respect to model choice, EJP clearly selected ESTAR with a non-zero threshold while TP unequivocally selected the LSTAR specification.

6.2 Estimation and Evaluation

Following the results of the previous section, we proceeded to build and estimate ESTAR and LSTAR specifications to compare their performance. Following Teräsvirta (1994), estimation is performed by nonlinear least squares. Under certain conditions, including stationarity and ergodicity of the series, the estimators of the parameters are consistent and asymptotically normal.[18] The LSTAR alternative produced models that were either not superior to the linear alternative or did not converge. As a result we do not report estimation results. The ESTAR alternative did produce models that had better fit than the linear model. The results of the preferred model are reported in Table 6.2.1.

An important result is worth noting. The ratio of the residual variance of the nonlinear model to that of the linear model was 0.97, clearly above the 0.90 level proposed by Granger and Andersen (1978) as a guideline to avoid spurious results. This indicates that while the ESTAR model has a better fit than its linear counterpart, the improvement is not very significant and therefore caution should be exercised in not putting too much weight on this specification. Given this caveat, there is perhaps still some interesting information that can be extracted from the model.

Figure 6.2.2 plots the transition function versus the data as we did in Section 2. This exponential function shows that the transition between regimes is very smooth — perhaps indicating why NL4 achieved the lowest p-value of all nonlinearity tests and why TP might

[18] See Tong (1990) Chapter 5.

have selected the LSTAR. Figure 6.2.3 plots the impulse response functions for each of the linear regimes in the ESTAR model. For both regimes, the impulse response functions reveal very strong persistence. The upper regime exhibits a sinusoidal decay, characteristic of polynomials with complex roots, with a period of approximately 40 months. The lower regime is far more persistent (a 1% increase takes over 20 years to die out). Complementing these plots, Figure 6.2.4 shows the graph of the unemployment rate and the values of the transition function at each point. An important aspect shown by this graph is the following: Prior to approximately 1975, recessions or in other words, periods of high unemployment, are associated with periods in which the lower regime dominates while periods of low unemployment are associated to the upper regime. Starting around 1975 these dynamics reverse themselves. In fact a Chow test for a break point in 1975 performed on the residuals of the ESTAR model rejects the null hypothesis of no-break.[19]

How successful is the ESTAR model in picking up asymmetries? Aside from the difference in dynamic behavior explained above, there seems to be ample scope to improve the modelling stage. The residuals exhibit ARCH, the Jarque-Bera test rejects the null of normality and there is left-over skewness in the data (asymmetry defined in the sense of De Long and Summers (1986)). This suggests that the ESTAR model, while it provided some useful insights, was not a good model in itself. For our purposes however, we did learn a few things. The ESTAR model was clearly more helpful than its LSTAR alternative to describe some of the features of the data. The nonlinearity tests that we suggested had lower p-values than those available in the literature and our selection rule did suggest using the most useful

[19] This might constitute another reason for why the LSTAR model was a poor alternative to the ESTAR.

specification within the STR family — the exponential STR, unlike the rule proposed by Teräsvirta (1994).

7 Conclusion

This paper provides a variety of useful suggestions and testing strategies for the empirical analysis with STR models. Our analysis is based on the properties of Taylor approximations to the transition function and the construction of different LM tests based on auxiliary regressions that use these approximations. The most significant result is the introduction of a new specification strategy to choose between exponential and logistic STR models: EJP. Monte Carlo evidence showed that this procedure offers much higher correct selection frequencies, it is consistent unlike its predecessor, TP, and it is straight forward to apply. Another important result was the realization that nonlinearity testing can be improved. By augmenting the existing tests with fourth order terms, we are able to approximate salient features of the exponential function that increase the power of the test. While this result is not general — there is not an equivalent gain when the alternative is a LSTR – Monte Carlo evidence suggests that the gains in power are significant. Moreover, it constitutes the natural test in view of our selection procedure.

Understanding the nature of the Taylor approximations and the testing strategies is useful to adapt the testing procedures to the empirical practice of each particular case. For example, we showed that in situations where parsimony is at stake, one can construct simplified versions of the test – namely, NL4A. Additionally, if one had transition functions other than the exponential and the logistic in mind, specific procedures can be readily constructed from the derivations in the paper. Finally, the paper expands specification and testing to the STR family. Extensions to multivariate analysis are straight forward generalizations to the equations proposed except for a few obvious details.

8 Appendix

8.1 Derivations of NL3 and NL4

Recall equation 9:

$$y_t = \pi' \underline{x}_t + \frac{1}{4}\gamma\theta' \underline{x}_t \left(z_{t-d} - c\right) + \frac{1}{48}\gamma^3\theta' \underline{x}_t \left(z_{t-d} - c\right)^3 + v_{3t} \tag{A.1}$$

from which we derived equation 10:

$$y_t = \delta_0 + \widetilde{\delta}' \underline{\widetilde{x}}_t + \beta_1' \underline{\widetilde{x}}_t z_{t-d} + \beta_2' \underline{\widetilde{x}}_t z_{t-d}^2 + \beta_3' \underline{\widetilde{x}}_t z_{t-d}^3 + v_{3t} \tag{A.2}$$

The reparametrization from equation 9 to 10 is as follows:

$$\delta_0 = \pi_0 - \frac{1}{4}\gamma\theta_0 c \left[1 + \frac{1}{12}\gamma^2 c^2\right]$$

$$\widetilde{\delta}' = \widetilde{\pi}' - \frac{1}{4}\gamma c \left[1 + \frac{1}{12}\gamma^2 c^2\right] \widetilde{\theta}'$$

$$\delta_d = \pi_d + \frac{1}{4}\gamma c \left[\theta_0 - \theta_d\right] + \frac{1}{48}\gamma^3 c^3 \left[3\theta_0 - \theta_d\right]$$

$$\beta_1' = \frac{1}{4}\gamma \left[1 + \frac{1}{4}\gamma^2 c^2\right] \widetilde{\theta}'$$

$$\beta_{1d} = \frac{1}{4}\gamma\theta_d + \frac{1}{16}\gamma^3 c \left(c\theta_d - \theta_0\right)$$

$$\beta_2' = -\frac{1}{16}\gamma^3 c \widetilde{\theta}'$$

$$\beta_{2d} = \frac{1}{48}\gamma^3 \left(\theta_0 - 3c\theta_d\right)$$

$$\beta_3' = \frac{1}{48}\gamma^3\theta_0$$

Recall equation 13:

$$y_t = \pi'\underline{x}_t + \theta'\underline{x}_t\left[(z_{t-d}-c)^2\left(\gamma+\gamma^2\right) - \frac{1}{2}(z_{t-d}-c)^4\gamma^2\right] + v_{4t} \tag{A.3}$$

from which we derived equation 14:

$$y_t = \delta_0 + \widetilde{\delta}'\underline{\widetilde{x}}_t + \beta_1'\underline{\widetilde{x}}_t z_{t-d} + \beta_2'\underline{\widetilde{x}}_t z_{t-d}^2 + \beta_3'\underline{\widetilde{x}}_t z_{t-d}^3 + \beta_4'\underline{\widetilde{x}}_t z_{t-d}^4 v_{4t} \tag{A.4}$$

This transformation is done as follows:

$$\delta_0 = \pi_0 + \theta_0\left(\gamma+\gamma^2\right)c^2 - \frac{1}{2}\theta_0\gamma c^4$$

$$\widetilde{\delta}' = \widetilde{\pi}' + c^2\left(\gamma+\gamma^2\right)\widetilde{\theta}' + \frac{1}{2}\gamma c^4$$

$$\delta_d = \pi_d - c\left(\gamma+\gamma^2\right)(c\theta_d - 2) + \frac{1}{2}\gamma c^3(c\theta_d - 4\theta_0)$$

$$\beta_1' = 2c\gamma\left(c^2 - (1+\gamma)\right)\widetilde{\theta}'$$

$$\beta_{1d} = \left(\gamma+\gamma^2\right)(\theta_0 - 2c\theta_d) + \gamma c^2(2c\theta_d - 3\theta_0)$$

$$\beta_2' = \left[\left(\gamma+\gamma^2\right) - 3\gamma c^2\right]\widetilde{\theta}'$$

$$\beta_{2d} = \left(\gamma+\gamma^2\right)\theta_d + \gamma c(2\theta_0 - 3c\theta_d)$$

$$\beta_3' = 2\gamma c\tilde{\theta}'$$

$$\beta_{3d} = \gamma\left(2c\theta_d - \frac{1}{2}\theta_0\right)$$

$$\beta_4' = \frac{1}{2}\gamma\tilde{\theta}'$$

8.2 Description of the Experiments

The series in this study were generated according to each of the models. The first 100 observations from each of the series were disregarded. Each experiment was replicated 1,000 times. The χ^2 version of the test was used for the simulations to determine the empirical power of the NL3 and the NL4 tests. Experiments are not size-corrected. Size-correction showed that the empirical values were close to the asymptotic values. The simulations of the decision rules were conditioned on prior rejection of linearity with NL4 in its F-version. Similar results were obtained by conditioning on NL3. The models used and the details of each simulation are described in Tables 5.1.1 - 5.1.5. Tables 5.2.1 - 5.2.3 are based on the same models and their specifics are therefore not repeated.

References

Burgess, S. M. (1988), "Employment Adjustment in U.K. Manufacturing," *Economic Journal*, 98, 81-103.

Burgess, S. M. and J.J. Dolado, (1989), "Intertemporal Rules with Variable Speed of Adjustment: An Application to U.K. Manufacturing Employment," *Economic Journal*, 99, 347-365.

Chan, K. S. and H. Tong, (1986), "On Estimating Thresholds in Autoregressive Models," *Journal of Time Series Analysis*, 7, 179-190.

Davies, R. B., (1977), "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247-254.

De Long, J. B. and L. H. Summers, (1986), "Are Business Cycles Asymmetrical?", in R.J. Gordon (ed.), *The American Business Cycle: Continuity and Change*. NBER Studies in Business Cycles, Vol. 25. University of Chicago Press: Chicago.

Escribano, A. and O. Jordá, (1997), "Testing Nonlinearity: Decision Rules for Selecting between Logisitc and Exponential STAR Models," working paper, Universidad Carlos III de Madrid.

Granger, C. W. J. and A. P. Andersen, (1978), *An Introduction to Bilinear Time Series Models.* Göttingen: Vandenhoeck and Ruprecht.

Haggan, V. and T. Ozaki, (1981), " Modelling Nonlinear Random Vibrations Using an Amplitude-Dependent Autoregressive Time Series Model," *Biometrika*, 68, 189-196.

Hamilton, J. D., (1989), "A New Approach to the Economic Analysis of Non Stationary Time Series and the Business Cycle," *Econometrica*, 57, 357-384.

Hansen, B., (1996), "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis," *Econometrica*, 64, 413-430.

Layard, R., S. Nickell and R. Jackman, (1991): *Unemployment: Macroeconomic Performance and the Labour Market.* Oxford University Press: Oxford.

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

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

Neftçi, S. N., (1984), "Are Economic Time Series Asymmetric over the Business Cycle?", *Journal of Political Economy*, 92, 307-328.

Pesaran, M. H. and S. M. Potter, (1997), "A Floor and Ceiling Model of U. S. Output," *Journal of Economic Dynamics and Control*, forthcoming.

Pfann, G. A. and F. C. Palm, (1993), "Asymmetric Adjustment Costs in Non-Linear Labour Demand Models for the Netherlands and U.K. Manufacturing Sectors," *Review of Economic Studies*, 60, 397-412.

Pissarides, C., (1992), " Loss of Skill during Unemployment and the Persistence of Employment Shocks," *Quarterly Journal of Economics*, 107, 1371-1392.

Rothman, P., (1991), "Further Evidence on the Asymmetric Behavior of Unemployment Rates over the Business Cycle," *Journal of Macroeconomics*, 13, 291-298.

Saikkonen, P. and R. Luukkonen, (1988), "Lagrange Multiplier Tests for Testing Nonlinearities in Time Series Models," *Scandinavian Journal of Statistics*, 15, 55-68.

Sichel, D. (1991), "Business Cycle Asymmetry: A Deeper Look," Working Paper n. 93, Board of Governors of the Federal Reserve System, Economic Activity Section, Washington, D.C.

Teräsvirta, T. and H. M. Anderson, (1992), "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models," *Journal of Applied Econometrics*, 7, S119-S136.

Teräsvirta, T., (1994), "Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, vol. 89, no. 425, 208-218.

Teräsvirta, T., D. Tjostheim and C. W. J. Granger, (1994), "Aspects of Modelling Nonlinear Time Series," in R. F. Engle and D. McFadden (eds.) *Handbook of Econometrics Vol. 4*, Amsterdam: Elsevier.

Tong, H., (1990), *Non-Linear Time-Series: A Dynamical System Approach*. New York: Springer-Verlag.

Tsay, R., (1986), "Non-Linearity Test for Time Series," *Biometrika*, 73, 461-466.

Tsay, R., (1989), "Testing and Modelling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84, 231-240.

White, H., (1984), *Asymptotic Theory for Econometricians*, London and San Diego: Academic Press.

Table 5.1.1 – Relative correct selection frequencies

Model: *LSTAR*, Fig. 2, pg. 496 , cases (a) and (b) in Luukkonen et al. (1988a).
DGP:
ε ~ N(0, 25); c = 0; d = 1; γ = 0.5

$$y_t - 0.5y_{t-1} + (\theta_1 y_{t-1})F(z_t) = \varepsilon_t$$

Sample Size	θ_1	TP	GTP	EJP	C = 0	Power NL4
		Selection Procedures				
	-0.4	0.500	0.406	0.594	0.868	0.064
	0	-	-	-	-	0.039
50	0.5	0.736	0.597	0.736	0.868	0.072
	1.0	0.853	0.776	0.871	0.838	0.170
	1.5	0.904	0.889	0.936	0.787	0.467
	-0.4	0.459	0.377	0.811	0.798	0.122
	0	-	-	-	-	0.043
100	0.5	0.811	0.724	0.724	0.793	0.127
	1.0	0.952	0.916	0.936	0.867	0.498
	1.5	0.963	0.960	0.978	0.917	0.883

Table 5.1.2 - Relative correct selection frequencies

Model: *LSTAR*, Fig. 2, pg. 496, cases (c) and (d) in Luukkonen et al. (1988a).
DGP:

$$y_t - 0.5y_{t-1} + (\theta_1 y_{t-1})F(z_t) = \varepsilon_t$$

ε ~ N(0, 25); c = 0; d = 1; γ = 0.5

Sample Size	θ_1	TP	GTP	EJP	C = 0	Power NL4
		Selection Procedures				
	-1.4	0.568	0.537	0.947	0.787	0.322
	-1.0	0.872	0.788	0.872	0.785	0.203
50	-0.5	0.802	0.630	0.630	0.667	0.081
	0	-	-	-	-	0.047
	0.5	0.841	0.735	0.690	0.782	0.113
	-1.4	0.594	0.585	0.978	0.823	0.744
	-1.0	0.941	0.916	0.935	0.858	0.491
100	-0.5	0.898	0.814	0.814	0.854	0.118
	0	-	-	-	-	0.037
	0.5	0.919	0.871	0.860	0.791	0.272

Table 5.1.3 – Relative correct selection frequencies

Model: 4.1 and 4.6 in Teräsvirta (1994).
DGP:

$d = 1$; $u_t \sim N(0, 0.04)$; $\gamma = 100$ (LSTAR), 1000 (ESTAR).

LSTAR Model

Sample Size	c	π_{20}	TP	GTP	EJP	C = 0	Power NL4
			Selection Procedure				
	0	0	0.975	0.965	0.997	0.984	0.950
100	0	0.02	0.932	0.909	0.907	0.913	0.636
	0.02	0.04	0.859	0.827	0.577	0.611	0.156
	0	0	1.000	1.000	1.000	1.000	1.000
300	0	0.02	0.998	0.996	0.985	0.979	0.993
	0.02	0.04	0.959	0.925	0.671	0.691	0.386

ESTAR Model

Sample Size	c	π_{20}	TP	GTP	EJP	C = 0	Power NL4
			Selection Procedure				
	0	0	0.885	0.922	0.971	0.963	0.729
100	0	0.02	0.673	0.698	0.670	0.863	0.838
	0.02	0.04	0.488	0.569	0.350	0.816	0.606
	0	0	0.983	0.989	1.000	1.000	0.999
300	0	0.02	0.876	0.883	0.850	0.995	1.000
	0.02	0.04	0.332	0.378	0.569	0.961	0.982

Table 5.1.4 - Relative correct selection frequencies

Model: *ESTAR*. Table 4, pg. 172 in Luukkonen et al. (1988b).
DGP:
$\varepsilon \sim N(0, 0.36)$

$$y_t - 0.3y_{t-1} - \left\{\exp\left(-y_{t-1}^2\right) - 1\right\}0.9y_{t-1} = \mu + \varepsilon_t$$

Selection Procedures

Sample Size	μ	TP	GTP	EJP	C = 0	Power NL4
	0	0.632	0.698	0.792	0.833	0.106
50	0.3	0.472	0.552	0.736	0.739	0.125
	1	0.129	0.252	0.624	0.466	0.210
	0	0.805	0.836	0.898	0.904	0.256
100	0.3	0.522	0.609	0.830	0.768	0.317
	1	0.095	0.142	0.704	0.403	0.493
	0	0.899	0.914	0.963	0.936	0.616
200	0.3	0.659	0.686	0.923	0.753	0.692
	1	0.039	0.053	0.765	0.510	0.881

Table 5.1.5 – Relative correct selection frequencies

Model: ESTAR. Table 4.3, pg. 64 in Saikkonen and Luukkonen (1988).
DGP:
$u_t \sim N(0, 0.09)$

$$y_t + \left\{a + \theta\exp(-y_{t-1}^2)\right\}y_{t-1} = u_t$$

Selection Procedure

Sample Size	a	θ	TP	GTP	EJP	C = 0	Power NL4
	0.9	0.3	0.667	0.746	0.702	0.925	0.114
	0.9	0.6	0.826	0.852	0.852	0.836	0.264
50	-0.3	-0.9	0.747	0.783	0.816	0.768	0.217
	-0.6	-0.9	0.720	0.752	0.858	0.700	0.218
	-0.6	0.9	0.379	0.485	0.591	0.667	0.066
	0.9	0.3	0.820	0.860	0.847	0.894	0.222
	0.9	0.6	0.958	0.967	0.973	0.969	0.669
100	-0.3	-0.9	0.854	0.858	0.913	0.820	0.528
	-0.6	-0.9	0.751	0.757	0.941	0.829	0.666
	-0.6	0.9	0.600	0.650	0.750	0.475	0.080
	0.9	0.3	0.924	0.940	0.946	0.954	0.485
	0.9	0.6	0.993	0.994	0.998	1.000	0.974
200	-0.3	-0.9	0.959	0.959	0.982	0.961	0.903
	-0.6	-0.9	0.805	0.806	0.980	0.931	0.966
	-0.6	0.9	0.805	0.842	0.858	0.840	0.190

Table 5.2.1 – Power simulations

Model: Table 4, pg. 172 in Luukkonen et al. (1988b).
Note: LSTAR model constructed for the same values of γ and c as original ESTAR model.

		$\sigma = 0.6$		$\sigma = 0.6$		$\sigma = 0.6$	
Sample Size	**Model**	**NL3 Power**	**NL4 Power**	**NL3 Power**	**NL4 Power**	**NL3 Power**	**NL4 Power**
$\mu = 0$							
	ESTAR	0.183	0.168	0.106	0.138	0.054	0.0712
	LSTAR	0.125	0.103	0.203	0.173	0.242	0.202
	ESTAR	0.405	0.411	0.225	0.292	0.107	0.152
	LSTAR	0.242	0.225	0.374	0.347	0.458	0.436
	ESTAR	0.725	0.746	0.459	0.623	0.157	0.263
	LSTAR	0.486	0.425	0.683	0.636	0.828	0.792
$\mu = 0.3$							
	ESTAR	0.232	0.201	0.130	0.139	0.077	0.090
	LSTAR	0.127	0.107	0.200	0.173	0.264	0.233
	ESTAR	0.439	0.433	0.291	0.339	0.102	0.142
	LSTAR	0.266	0.238	0.393	0.361	0.521	0.461
	ESTAR	0.823	0.828	0.502	0.621	0.178	0.303
	LSTAR	0.487	0.423	0.707	0.653	0.822	0.793
$\mu = 1$							
	ESTAR	0.384	0.339	0.181	0.172	0.087	0.078
	LSTAR	0.114	0.102	0.222	0.203	0.281	0.248
	ESTAR	0.702	0.661	0.321	0.319	0.155	0.169
	LSTAR	0.240	0.206	0.419	0.385	0.529	0.442
	ESTAR	0.963	0.946	0.629	0.603	0.246	0.246
	LSTAR	0.490	0.442	0.714	0.754	0.850	0.802

Table 5.2.2 – Power simulations

Model: Table 4.3, pg. 64 in Saikkonen and Luukkonen (1988).
Note: LSTAR model constructed for the same values of γ and c as original ESTAR model.

		$\sigma = 0.3$		$\sigma = 0.6$		$\sigma = 0.9$	
		NL3 Power	**NL4 Power**	**NL3 Power**	**NL4 Power**	**NL3 Power**	**NL4 Power**
Sample Size = 50							
A = 0.9	ESTAR	0.220	0.201	0.085	0.095	0.073	0.072
θ = 0.3	LSTAR	0.994	0.994	0.997	0.997	0.996	0.996
A = 0.9	ESTAR	0.462	0.478	0.144	0.158	0.074	0.077
θ = 0.6	LSTAR*	0.045	0.035	0.045	0.047	0.048	0.059
A = -0.3	ESTAR	0.387	0.352	0.308	0.306	0.173	0.167
θ = -0.9	LSTAR	0.043	0.042	0.038	0.038	0.053	0.053
A = -0.6	ESTAR	0.363	0.348	0.338	0.321	0.154	0.167
θ = -0.9	LSTAR†	0.774	0.769	0.829	0.811	0.830	0.814
A = -0.6	ESTAR	0.081	0.078	0.160	0.162	0.100	0.122
θ = 0.9	LSTAR	0.069	0.057	0.108	0.104	0.169	0.143
Sample Size = 100							
A = 0.9	ESTAR	0.354	0.368	0.117	0.115	0.070	0.067
θ = 0.3	LSTAR	1.000	1.000	1.000	1.000	1.000	1.000
A = 0.9	ESTAR	0.787	0.856	0.218	0.265	0.084	0.079
θ = 0.6	LSTAR*	0.681	0.933	0.679	0.929	0.701	0.930
A = -0.3	ESTAR	0.742	0.703	0.566	0.578	0.274	0.331
θ = -0.9	LSTAR	0.053	0.053	0.065	0.065	0.078	0.078
A = -0.6	ESTAR	0.790	0.761	0.677	0.685	0.270	0.316
θ = -0.9	LSTAR†	0.567	0.553	0.761	0.741	0.778	0.771
A = -0.6	ESTAR	0.153	0.141	0.359	0.355	0.209	0.273
θ = 0.9	LSTAR	0.084	0.081	0.205	0.174	0.400	0.351
Sample Size = 200							
A = 0.9	ESTAR	0.610	0.643	0.148	0.162	0.072	0.071
θ = 0.3	LSTAR	1.000	1.000	1.000	1.000	1.000	1.000
A = 0.9	ESTAR	0.969	0.969	0.340	0.470	0.094	0.108
θ = 0.6	LSTAR*	0.981	0.997	0.975	0.987	0.978	0.991
A = -0.3	ESTAR	0.970	0.967	0.869	0.895	0.531	0.658
θ = -0.9	LSTAR	0.085	0.085	0.073	0.071	0.091	0.090
A = -0.6	ESTAR	0.980	0.978	0.917	0.957	0.445	0.577
θ = -0.9	LSTAR†	0.576	0.557	0.724	0.703	0.786	0.774
A = -0.6	ESTAR	0.312	0.278	0.699	0.735	0.448	0.599
θ = 0.9	LSTAR	0.170	0.152	0.466	0.410	0.730	0.685

* $\gamma/10$
† $\gamma/3000$

Table 5.2.3 – Power simulations

Model: 4.1 and 4.6 in Teräsvirta (1994).

	$\sigma = 0.01$		$\sigma = 0.02$		$\sigma = 0.04$	
	NL3 Power	**NL4 Power**	**NL3 Power**	**NL4 Power**	**NL3 Power**	**NL4 Power**
Sample Size = 100						
ESTAR	0.817	0.824	0.612	0.722	0.245	0.392
LSTAR	0.875	0.828	0.962	0.951	0.981	0.975
ESTAR	1.000	1.000	0.983	0.997	0.724	0.924
LSTAR	0.061	0.058	0.691	0.656	0.946	0.931
ESTAR	0.038	0.041	0.611	0.623	0.424	0.466
LSTAR	0.028	0.035	0.157	0.139	0.883	0.870
Sample Size = 300						
ESTAR	0.927	0.911	0.825	0.835	0.358	0.414
LSTAR	1.000	1.000	1.000	1.000	1.000	1.000
ESTAR	1.000	1.000	1.000	1.000	0.876	0.937
LSTAR	0.148	0.148	0.993	0.993	1.000	1.000
ESTAR	0.113	0.114	0.984	0.993	0.903	0.947
LSTAR	0.030	0.043	0.378	0.373	1.000	1.000

Table 6.1

Non-linearity tests, choice of delay parameter and model selection tests. U.S. Unemployment Rate, 1948:01 to 1997:07, seasonally adjusted.

Test	p-value	Delay parameter
NL3A	0.02984	6
NL3	0.00012	6
NL4A	0.01926	6
NL4	0.00009	6

Selection Procedure	Model Selected
TP	LSTAR
EJP	ESTAR with $c \neq 0$

Table 6.2.2

ESTAR Estimates and Statistics. U. S. Unemployment Rate, 1948:01 – 1997:07, seasonally adjusted.

Lower Regime				Upper Regime			
Coef.	**Estimate**	**Std. Error**	**T-Statistic**	**Coef.**	**Estimate**	**Std. Error**	**T-Statistic**
π_0	0.0320	0.1031	0.3101	θ_0	0.0956	0.1262	0.7573
π_1	0.9333	0.0789	11.823	θ_1	0.2888	0.1316	2.1948
π_2	0.2044	0.0929	2.2013	θ_2	-0.3490	0.1575	2.2156
π_6	-0.1887	0.0404	4.6681	θ_6	0.1521	0.0818	1.8599
π_{12}	-0.1727	0.0429	4.0284	θ_{13}	-0.1062	0.0494	2.1476
π_{13}	0.2286	0.0431	5.3080				
π_{24}	-0.1380	0.0387	3.5697				
π_{25}	0.1262	0.0367	3.4427				
γ	1.1735	1.0422	1.1260				
c	6.8691	0.2808	24.467				

Summary Statistics

R-Squared	0.986316	ARCH Test (p-val.)	0.012627
S. S. R.	19.64620	Jarque-Bera Res.	0.000000
Log-Likelihood	151.0145	Skewness Res.	0.361244
Durbin – Watson	2.002727	Kurtosis Res.	4.286392

Figure 2.1 - STR Simulated Models

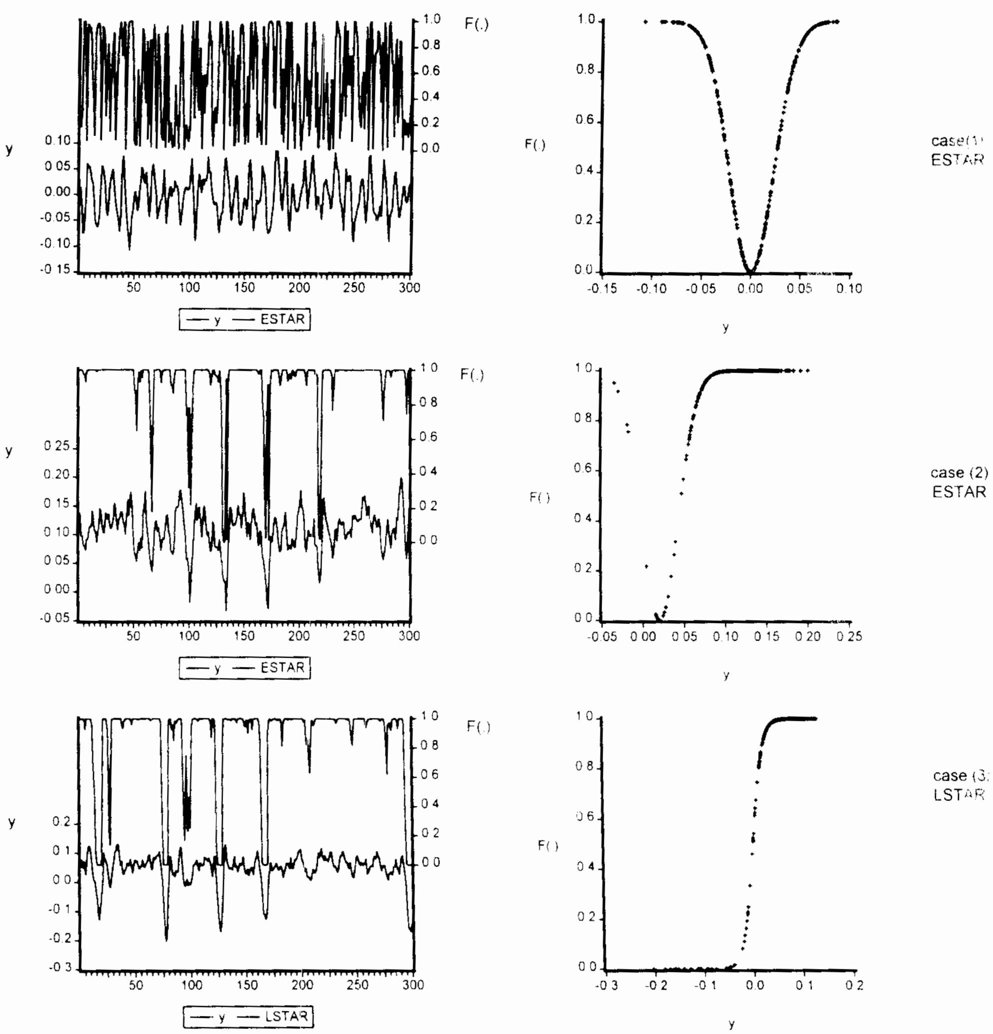

Figure 6.1 - U.S. Unemployment Rate
(1948:01 - 1997:07)

Figure 6.2.2. - Transition Function from ESTAR Model. U.S. Unemployment Rate

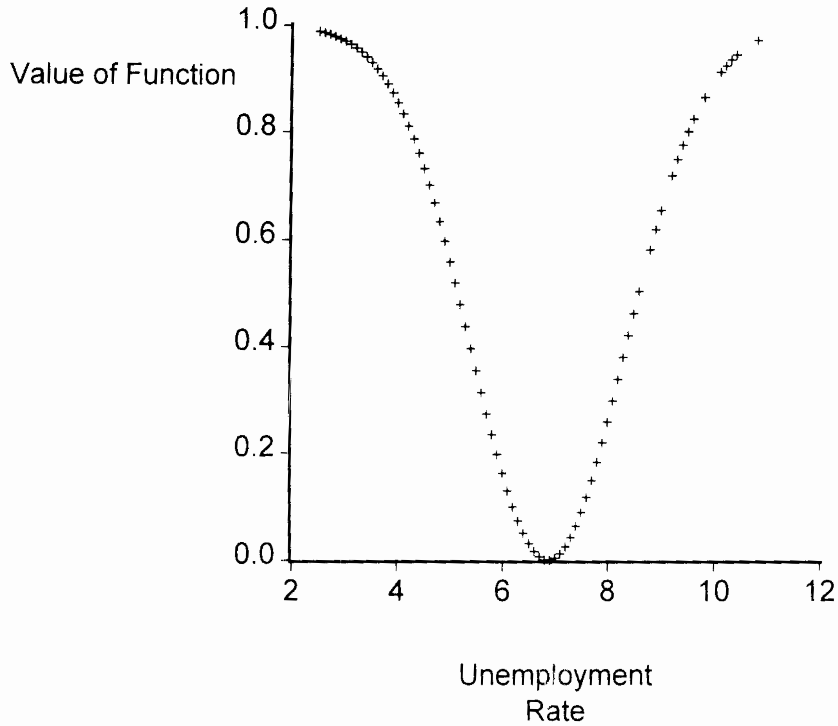

Figure 6.2.3 - Impulse Response Functions of Linear Regimes from ESTAR Model - U.S. Unemployment Rate

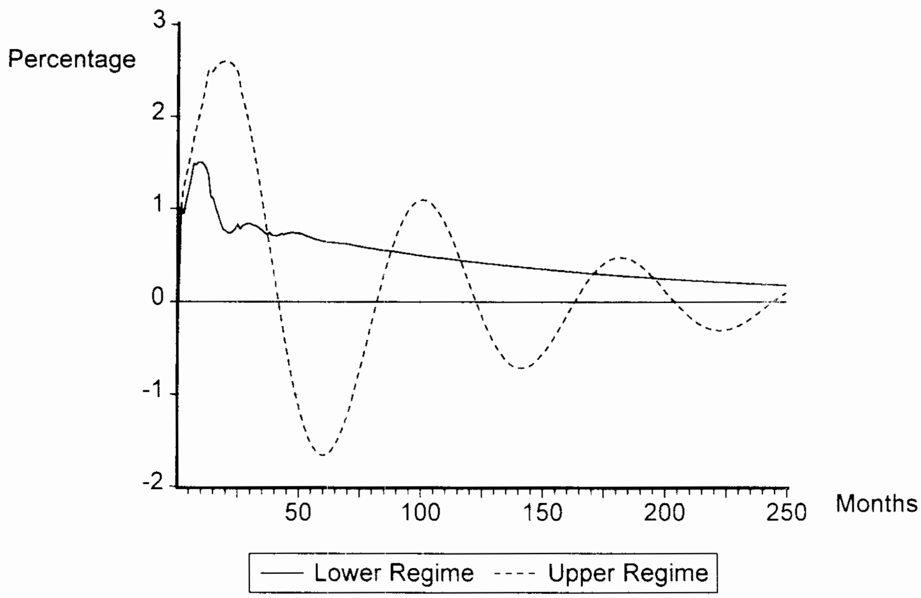

Figure 6.2.4. - U.S. Unemployment Rate and Transition Function from ESTAR Model