Model Selection for Nonlinear Time Series

Sebastiano Manzan

Center for Nonlinear Dynamics in Economics and Finance (CeNDEF), Faculty of Economics

and Econometrics, and Tinbergen Institute, University of Amsterdam

Roetersstraat 11, 1018 WB Amsterdam, The Netherlands

Tel.: +31 20 5254248, Fax: +31 20 5254349

e-mail: `manzan@fee.uva.nl`

Abstract

We investigate the finite-sample performance of model selection criteria for local linear regression by simulation. Similarly to linear regression, the penalization term depends on the number of parameters of the model. In the context of nonparametric regression we use a suitable quantity to account for the Equivalent Number of Parameters as previously suggested in the literature. We consider the following criteria: Rice T, FPE, AIC, Corrected AIC and GCV. To make results comparable with other data-driven selection criteria we consider also Leave-Out CV. We show that the properties of the penalization schemes are very different for some linear and nonlinear models. Finally, we set up a goodness-of-fit test for linearity based on bootstrap methods. The test has correct size and very high power against the alternatives investigated. Application of the methods proposed to macroeconomic and financial time series shows that for some of them there is evidence of nonlinearity.

1 Introduction

Recently, the application of nonlinear methods to economic and financial data has gathered increasing interest. The seminal work of Hamilton (1989) on markov switching models and the successful application to US GNP pointed out the relevance of considering nonlinear effects. The range of nonlinear models used has widened rapidly and an account of the various approaches and results can be found in Granger (2001). An alternative way to model the nonlinearities in the data is to use nonparametric regression techniques. Their main advantage is the flexibility in capturing the dependence structure in the data without relying on a specific parametric family. However, a major disadvantage is that they require long time series to deliver reliable results and this availability is not always the case for macroeconomic time series. The flexibility of nonparametric methods allows to use it as a first tool of analysis for the presence of dependence, either linear or nonlinear. If the nonparametric methods find significant nonlinear structure, then the analysis can proceed by specifing which parametric model fits the data best. In this perspective, nonparametric techniques should be interpreted as complementary to parametric methods and not competing. The curse of dimensionality affects the reliability of these methods in small and moderate sample sizes that are typically available in economics. Balancing of these contrasting issues has lead to few applications of nonparametric regression methods to economic time series. Some selected references are Diebold and Nason (1990) to weekly exchange rates, Mizrach (1992) to daily exchange rates and Pagan and Schwert (1990) to estimate the conditional variance of stock prices.

In a linear regression framework, the model selection step involves the choice of the number of lags to include in the regression. In addition to the lag, in nonparametrics another element that characterizes the models is the bandwidth, that is the amount of smoothing that is used. In this article we compare the finite sample performances of the criteria in selecting the nonparametric model. The paper is organized as follows: in section (2) we describe the local linear smoother and the selection criteria we use to choose order and bandwidth. Section (3) shows simulation results concerning various linear

and nonlinear autoregressive models. In section (4) a goodness-of-fit test for linearity is proposed and the appropriateness of the selection criteria will be emphasized. Finally, in section (5) we investigate the presence of nonlinearity in some macroeconomic time series.

2 Method

Assume $\{x_t\}_{t=-p}^{n+1}$ is a univariate stationary time series generated by the following nonlinear autoregressive model of order p

$$x_{t+1} = m(X_t) + \epsilon_{t+1}$$

where $m(x)$ is a function of unknown form, $X_t = (x_t, ..., x_{t-p+1})'$ denotes the p-dimensional vector of lagged values of the time series and ϵ_t is an $i.i.d.$ disturbance term with mean 0 and variance σ^2. This general form encompasses the AR model as well as many nonlinear time series models like thresold autoregressive (TAR) and exponential autoregressive (EXPAR). We estimate $m(x)$ using nonparametric regression techniques. In particular, we adopt the local linear approach proposed by Cleveland and Devlin (1988) and Fan and Gijbels (1996). The local linear estimator using nearest-neighbors bandwidth is known in the literature as LOESS and it minimizes

$$\sum_{t=1}^{n} \{x_{t+1} - \alpha - \beta(X_t - x)\}^2 K\left(\frac{\| X_t - x \|}{d_k(x)}\right) \tag{1}$$

where $\widehat{m}(x) = \hat{\alpha}$, $K(\cdot)$ is the tricube kernel defined as

$$K(u) = \begin{cases} (1 - u^3)^3 & \text{for } 0 \leq u < 1 \\ 0 & \text{otherwise,} \end{cases}$$

$\| \cdot \|$ indicates the euclidean distance and

$$d_k(x) = \begin{cases} \| X_{x(k)} - x \| & \text{for } 0 < h \leq 1 \\ \| X_{x(n)} - x \| h^{\frac{1}{p}} & \text{for } h > 1 \end{cases}$$

where $X_{x(k)}$ denotes the k-th nearest neighbor of x and k is the integer part of hn. The bandwidth h can be interpreted as the parameter that regulates the smoothness of the

local linear fit. The smoothing scheme can be described as follows: if X_t is among the k nearest neighbors of the design point x it receives a positive weight given by the tricube kernel $K(\cdot)$ otherwise it receives a null weight. A practical advantage of nearest neighbors bandwidths with respect to fixed bandwidths is that they deliver more reliable and stable variances of the fit in regions where the data are sparse. If we let the bandwidth $h \to \infty$ we are including all the points in the regression and the local linear model approaches the linear AR model.

Here we will consider data-drive (or automatic) methods to select h, that is choosing the bandwidth that is optimal with respect to some error criteria. If we would choose h such that it minimizes the RSS (Residual Sum of Squares)

$$\frac{1}{n} \sum_{t=1}^{n-1} \{x_{t+1} - \widehat{m}_h(X_t)\}^2 w(X_t) \tag{2}$$

where $w(\cdot)$ is a weight function, we would trivially achieve a minimum for $h \to 0$ because it implies $\widehat{m}(X_t) \to x_{t+1}$. One way to solve this problem is the Leave-One-Out Cross-Validation method, that minimizes the following function

$$CV(h) = \frac{1}{n} \sum_{t=1}^{n} \{x_{t+1} - \widehat{m}_{h,-t}(X_t)\}^2 w(X_t) \tag{3}$$

where $\widehat{m}_{h,-t}(x)$ indicates the fitted value obtained by excluding the t-th observation. In a time series context, Härdle and Vieu (1992) proved that this method gives asymptotically optimal h. An alternative way to solve the problem is to multiply the RSS by a penalization factor that corrects for h too small. These methods are inspired by selection criteria used for linear models that choose the order p that minimizes

$$SC(p) = \log RSS + \phi(\pi) \tag{4}$$

where the first term indicates the goodness-of-fit of the model and the second term penalizes the inclusion of more parameters given by $\pi = p + 1$. The extension of this approach to a nonparametric regression framework is straightforward because they are linear smoothers. We can express the fitted regression function as

$$\widehat{y} = Hy$$

where $y = (x_2, ..., x_{n+1})'$ and H is the $n{\times}n$ hat matrix that depends only on lagged values. Similarly to the linear case, we can define the number of parameters involved in the regression by

$$\pi = \frac{tr(H)}{n}$$

if the bandwidth tends to ∞ then $tr(H)$ will be approximately equal to $p+1$ but for $h \to 0$ it will approach n, the case in which we fit as many parameters as data points.

The extension to the nonparametric case of the criterion in (4) is

$$SC(h, p) = \log RSS + \phi(\pi) \qquad (5)$$

where the dependence of the criterion on the bandwidth, in addition to the lag order, is taken into account. In this way we quantify the complexity of the model implied by both the choice of the bandwidth and the number of lags. Considering more lags and smaller bandwidths increases π and $\phi(\pi)$ will increase the penalization on the goodness of fit measure. π is also called in the literature the Equivalent Number of Parameters (ENP) by Cleveland and Devlin (1988) to stress the analogy with linear regression.

There are many proposals concerning the form of the penalization function $\phi(\cdot)$. We will consider here the most frequently used:

1. Akaike Information Criteria (AIC): $\phi(\pi) = 2\pi$

2. Corrected AIC (AICC): $\phi(\pi) = \frac{1+\pi}{1-\pi-\frac{2}{n}}$

3. Final Prediction Error (FPE): $\phi(\pi) = \log\left(\frac{1+\pi}{1-\pi}\right)$

4. Generalized Cross-Validation (GCV) : $\phi(\pi) = -2\log\left(1-\pi\right)$

5. Rice T (T): $\phi(\pi) = -\log\left(1-2\pi\right)$

A discussion of these criteria can be found in Härdle (1990). A bias corrected version of AIC, indicated as AICC, has been recently proposed by Hurvich *et al.* (1998). Selection criteria are used in nonparametric regression to select order or bandwidth and in theoretical work these two problems are kept separate. Some previous work on lag selection using

nonparametric regression are Tschernig and Yang (2000), Tjøstheim and Auestad (1994) and Cheng and Tong (1991); for bandwidth selection see Hurvich *et al.* (1998), Yao and Tong (1998) and Härdle and Vieu (1992).

In this article we adopt the practical point of view of the time series analyst that has to select both bandwidth and lag order. Instead of considering the problems of choosing the optimal bandwidth and order separately, we jointly select these 2 parameters by minimization of Equation (5). In the following paragraph we will compare via simulation the finite-sample performance of the different selection criteria.

3   Simulation

We simulate 1000 samples of length 100 for each of the models. We consider a maximum lag of 4 and the bandwidth h is varied from 0.20 to 1 at steps of 0.02. The choice of the lower bound to search for the optimal bandwidth is somewhat arbitrary and taking it small can be unfair with respect to criteria that are affected by undersmoothing. However, the purpose of this section is to investigate the performance of the criteria for reasonable enough h. Increasing the order p we consider more lags in the regression function instead of the procedure adopted by Tjøstheim and Auestad (1994) that performs a specification search for the lags to be included in the estimation.

For all the models the noise term u_t is distributed standard normally. The optimal bandwidth, h^{opt}, is selected by minimizing the Average Squared Error (ASE) defined as

$$\frac{1}{n}\sum_{t=1}^{n}\left[\widehat{m}_h(X_t) - m(X_t)\right]^2 w(X_t)$$

over h, whereas the order is assumed to be known. We will also report the average of the ratio of the ASE calculated for the different selection criteria and for the optimal bandwidth.

AR(1)   The AR(1) process is defined as

$$x_{t+1} = 0.6x_t + u_{t+1}.$$

The top panel of Figure 1 presents the smoothed density of the log-ratio of the ENP (π) selected by the criterion to the optimal number of parameters. This plot gives information about the behaviour of the criteria in selecting the model with respect to the optimal choice: for positive values of the log-ratio the selection criteria is using more parameters than optimal, so it is affected by overfitting; instead, if negative values of the log-ratio occurs, the criteria is affected by underfitting, that is, it selects a too parsimonious model than optimal. Overfitting is caused by undersmoothing, that is a bandwidth too small than optimal, and/or by an order too large. The middle plot in Figure (1) shows the log-ratio of the bandwidth selected by the criteria with respect to the optimal one and the table gives the frequencies of order selections.

It emerges clearly that AIC and FPE have the tendency to overfit, in the sense of using too many parameters in the regression with respect to the optimal amount. The overfit of AIC and FPE is caused both by undersmoothing, too frequent selection of small bandwidth, and by frequent selection of high orders (than optimal). For this model, AIC is more likely to select lag 4 (49% of the times) than the true order 1 (only 36%). The other criteria deliver very similar results even though AICC and T are closer to the optimal choice both in terms of bandwidth and order selected. They also have a lower average ASE than GCV and CV. However, there seems to be some overfitting also for these criteria as can be seen in the skew of the distribution in the top panel. This is caused by the order selection, where also for AICC and T near 23% of the times an order higher than the true one is selected. This is probably a finite-sample result that might disappear if larger sample sizes are considered.

AR(1)-GARCH(1,1)   If we allow for heteroskedasticity in the disturbance of the AR(1) process we get

$$x_{t+1} = 0.6x_t + u_{t+1}, \qquad u_{t+1} \sim N(0, h_{t+1})$$

$$h_{t+1} = 1 + 0.1u_t^2 + 0.8h_t$$

Considering heteroskedastic innovations in the AR(1) process doesn't change dramat-

icaly our previous conclusions. Figure (2) shows that AIC and FPE select too small h and too large p with respect to the optimal amount. This causes the distribution of the ratio of the ENP to have a second mode in the right tail of the smoothed density. In 75% of the simulation AICC selected the correct order similarly to T. A little less precise are GCV that achieves results similar to CV. As pointed out before, AIC and FPE select the correct order in only 34% and 55% of the times.

TAR(1) The model is

$$x_{t+1} = -0.5x_t I(x_t \leq 1) + 0.6x_t I(x_t > 1) + u_{t+1}.$$

Figure 3 confirms that also for this model AIC and FPE tend to overfit. In this case the problem arise mainly in the selection of the bandwidth that is often too small than optimal. Also in terms of mean ASE it is clear that they have worse performances than the other criteria. T, AICC and GCV have similar behaviour and the distribution is slightly skewed. This can be explained by a tendency to undersmooth clear in the middle plot of Figure 3. T and AICC select in 94% of the cases order 1 whereas GCV 90% of the times. T, AICC and GCV have results comparable to CV.

EXPAR(1) The model is

$$x_{t+1} = \{0.5 + 1.3\exp(-0.5x_t^2)\}x_t + u_{t+1}.$$

Figure 4 shows that most of the remarks previously made hold also for the EXPAR model with dependence in the first lag. By looking at the mean ASE the best performing criteria are AICC, T and GCV; worse performances are achieved by FPE and even worse by AIC. The distribution of the ratio of bandwidths is skewed to the right even though this is not the case for the distribution of the ratio of ENP. It is probably the case that when a higher (than optimal) order is chosen it is also associated with a higher h. In this way the higher ENP due to the order are compensated by the lower ENP due to the bandwidth.

EXPAR(2)   The model is

$$x_{t+1} = \{0.5 + 1.3 \exp(-0.5 x_{t-1}^2)\} x_t + u_{t+1}$$

this model has the same structure of the previous one but the dependence is now spread in the first 2 lags. In this way we test to what extent the criteria are able to identify the second order. The bottom table in Figure (5) shows that all the criteria recognize the existence of dependence also in the second lag; T and AICC select the second lag 95% of the times, GCV and CV in around 86% instead of AIC that chooses with almost the same frequency the second and forth lag. Looking at the distribution of the ratio of ENP it appears that GCV is slightly biased toward overfitting instead of T and AICC. Again, AIC and FPE are affected by undersmoothing and higher lag selection also for this model.

4   A test for linearity

We set up a test for linearity based on the comparison of the goodness-of-fit of the parametric and nonparametric autoregression. Recently, Cai *et al.* (2000) and Lee and Ullah (2002) adopted a similar testing strategy. The model selection is based on the criteria examined above for both the parametric and nonparametric regressions. The null hypothesis we want to test is

$$H_0 : E(x_{t+1}|X_t) = X_t'\theta$$

$$H_1 : E(x_{t+1}|X_t) = m(X_t)$$

where θ is the coefficients vector of the AR(p) model and $m(\cdot)$ is an unspecified functional form. Let RSS^P and RSS^{NP} denote the parametric and nonparametric RSS, respectively, defined as

$$RSS^P = \frac{1}{n} \sum_{t=1}^{n} \{x_{t+1} - X_t\hat{\theta}\}^2$$

$$RSS^{NP} = \frac{1}{n} \sum_{t=1}^{n} \{x_{t+1} - \widehat{m}_h(X_t)\}^2 w(X_t)$$

the test statistic is defined as

$$B = \frac{RSS^P - RSS^{NP}}{RSS^{NP}}$$

9

To evaluate the significance of the test statistic we use bootstrap methods. In order to account for the heteroskedasticity observed in many economic time series, we resample the residuals of the nonparametric regression using the wild bootstrap approach proposed by Liu (1988). The procedure is as follows:

1. calculate the test statistic, B, for the original time series.

2. generate bootstrap innovations, u^*_{t+1}, from the centered fitted residuals of the non-parametric regression, $\hat{u}_{t+1}$, as

$$u^*_{t+1} = \begin{cases} a(\hat{u}_{t+1} - \bar{u}) & \text{with probability } p = (\sqrt{5} + 1)/(2\sqrt{5}) \\ b(\hat{u}_{t+1} - \bar{u}) & \text{with probability } 1 - p \end{cases}$$

where $a = -(\sqrt{5} - 1)/2$ and $b = (\sqrt{5} + 1)/2$.

3. generate iteratively a new bootstrap time series as

$$x^*_{t+1} = X^*_t \hat{\theta} + u^*_{t+1}$$

where X^*_t is a p-dimensional vector of lagged values.

5. calculate the test statistic B^* on the bootstrap time series. We use the same order (linear and nonlinear) and bandwidth selected for the original time series.

6. Repeat steps (1) and (2) M times.

7. Calculate the one sided p-value as

$$\hat{p} = \frac{1 + \#\{B^* > B\}}{1 + M}$$

and reject if $\hat{p} < \alpha$, where α denotes the significance level.

We perform a one-sided test because deviations from the null hypothesis are expected to occur for positive values of the test statistic. The consistency of the bootstrap procedure derive from the fact that the residuals of the nonparametric regression are always consistent both under the null and the alternative. See Cai *et al.* (2000) for details.

We simulate 1000 samples of size 100 and the number of bootstrap replications set to 199 for the models examined in section (3). Table 1 shows that the test has good size properties also in the presence of heteroskedasticity that is known to introduce size distorsion. Under the alternative examined the test has high power: for the AICC selection criteria it is 96% against TAR(1), 85% against EXPAR with dependence in the first lag and 91% when the dependence is in the second lag.

5 Empirical Applications

The simulations and the properties of the test for linearity suggest that AICC, T and GCV deliver reasonable results in a wide range of dependence structures, such as linear or nonlinear models. The sample size of 100 used in the simulations guarantees that the results are reliable for small and moderate samples. We apply the methods described to investigate the presence of nonlinearities in some macroeconomic and financial time series. Given the time series tested are longer than the sample size used in the simulations we will extend the lag search to 6 for quarterly data. Unless differently specified we analyze the growth rate of the variables. We will show the $SC(h, p)$ in Equation (5) only for the AICC selection criterion and inference for the other criteria is shown in tables. To make comparisons easier we plot the $SC(h, p)$, normalized by the log-variance of the time series, for both the parametric and nonparametric regression.

We considered the following series for the US economy:

GNP GNP (Gross National Product) in real terms and seasonally adjusted from the first quarter of 1947 to the third quarter of 2001. This time series has been successfully modelled using nonlinear models such as TAR models by Potter (1995)and Tiao and Tsay (1994) and markov-switching models by Hamilton (1989). The middle plot in Figure 6 compares the $SC(h, p)$ of the nonparametric regression to that of a linear AR model (the dotted line). The normalized error of the nonparametric regression using AICC or T criteria achieves its minimum at $h = 0.86$ and the goodness-of-fit test rejects at 10% the

null hypothesis of linearity. T and AICC give a similar answer selecting a moderate h and pointing to dependence in the first 5 lags. GCV selects a smaller bandwidth and an higher lag order and has a smaller p-value.

Industrial Production from first quarter of 1947 to the end of 2001. The data are seasonally adjusted. In Figure (7) it is clear that the nonparametric regression lowers consistently the error of the linear regression. T and AICC select lags up to 5 and $h = 0.74$. GCV selects a smaller bandwidth whereas CV an higher one. The test for linearity strongly reject the null hypothesis for all the selection criteria analyzed. For this time series we can conclude that there is robust evidence of nonlinear structure.

Unemployment Rate from first quarter 1947 to last quarter 2001. The data are seasonally adjusted. For previous nonlinear analysis of this time series see Montgomery *et al.* (1998). Table (8) shows that the selection criteria give a similar answer by rejecting the null of linearity at 1% significance level. However, AICC and T selects a slightly higher bandwidth than GCV. It is also to notice that CV does not detect significant nonlinear structure for this time series.

T-bills Monthly 3-months Treasury Bills interest rates from 1950 to the end of 2001. Also for this time series Figure (9) shows that the error of the nonparametric regression significantly improves over the error of the linear model. The optimal lag for T and AICC indicates a dependence among the last 6 quarters instead of CV that indicates a 2 quarters dependence. The bandwidth selected is 0.66 for AICC, 0.70 for T and 0.32 for GCV. The linearity test is rejected for all criteria except for CV where the p-value is 0.10.

S&P 500 Index from the first quarter of 1947 to the third quarter of 2001. The $SC(h, p)$ shows that the error of the nonparametric regression does not improve with respect to the parametric regression. Both curves fluctuate around 1 and suggest lack of dependence in the time series. The test for linearity, robust to heteroskedasticity in the time series, confirms that there is no evidence to reject the null hypothesis. The lack of dependence,

12

linear or nonlinear, in the conditional mean dynamics of financial time series is a well-known result. Our test based on nonparametric regression methods confirms that there is no evidence to suggest linearity of the conditional mean of the process.

DM/\$ exchange rate from first quarter 1974 to the end of 2001. Also in this case T and AICC reach a similar conclusion as for S&P500 returns, that is $h = 1$ and $p = 3$. Instead, GCV and CV select a smaller bandwidth of 0.82 and 0.88, respectively. The test does not reject the null of linearity except for GCV that has a p-value of 0.09. These results confirm that also for exchange rates there is no evidence of mean dependence in the series.

Yen/\$ exchange rate for the same period of DM/\$. Also in this case there is no evidence of nonlinearity both from the shape of the $SC^T(h, p)$ plot and from the linearity test. The error curve of the parametric and nonparametric regression are above 1 with the nonparametric one converging but never improving the linear one. The criteria unanimously identify h and p equal to 1.

6 Conclusion

The use of nonparametric regression as a tool to detect nonlinearities in time series could be very useful to economist: as a first step in the analysis of time series it guarantees flexibility in the dependence structure. In addition, we have shown that they are reasonably accurate in moderate samples. In this article we have shown that selecting bandwidth and lag order could be reliably carried out by using AICC or T that are not affected by significant problems of overfitting instead of AIC and FPE. In addition, a goodness-of-fit test based on the comparison of parametric and nonparametric regression is a powerful test to detect deviations from the linearity assumption. If the linearity test rejects the null hypothesis, further steps could be to apply a battery of parametric linearity test to identify which nonlinear dependence structure is more suitable to explain the time series dynamics. The application of nonparametric autoregression to macroeconomic time series shows that there
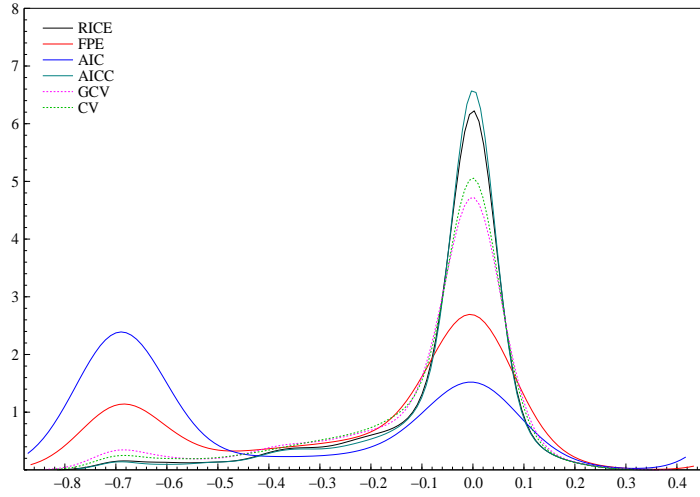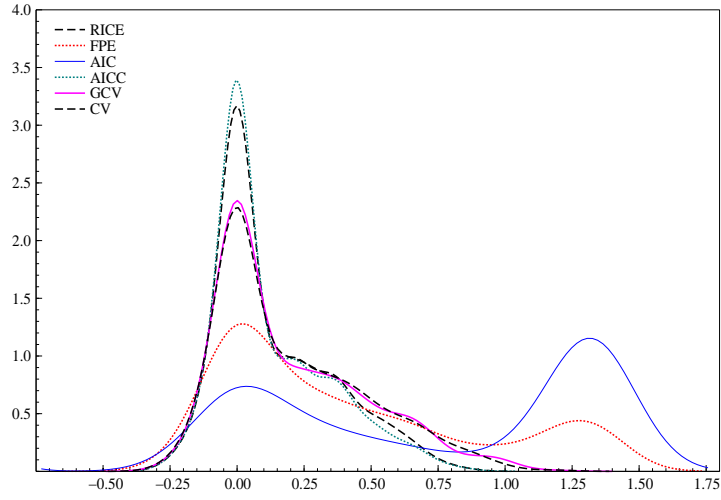
is statistical evidence of nonlinearity for some of them. Growth rates of US real GNP, Industrial Production and Unemployment Rate reject the null hypothesis of linearity when using T and AICC as selection criteria. However, the financial time series do not show any evidence for dependence with the exception of the returns on US T-Bills for which there is strong evidence of nonlinearity.

References

Cai, Z., Fan, J. and Yao, Q. (2000). Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, **95,** 941–956.

Cheng, B. and Tong, H. (1991). On consistent nonparametric order determination and chaos. *Journal of the Royal Statistical Society* B, **54,** number 2, 427–449.

Cleveland, W.S. and Devlin, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, **83,** number 403, 596–610.

Diebold, F.X. and Nason, J.A. (1990). Nonparametric exchange rate prediction? *Journal of International Economics*, **28,** 315–332.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman & Hall.

Granger, C. W. (2001). Overview of nonlinear macroeconometric empirical models. *Macroeconomic Dynamics*, **5,** number 4, 466–481.

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

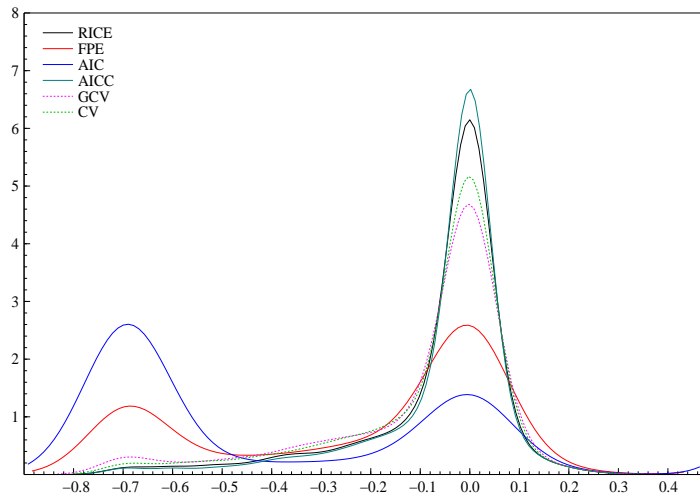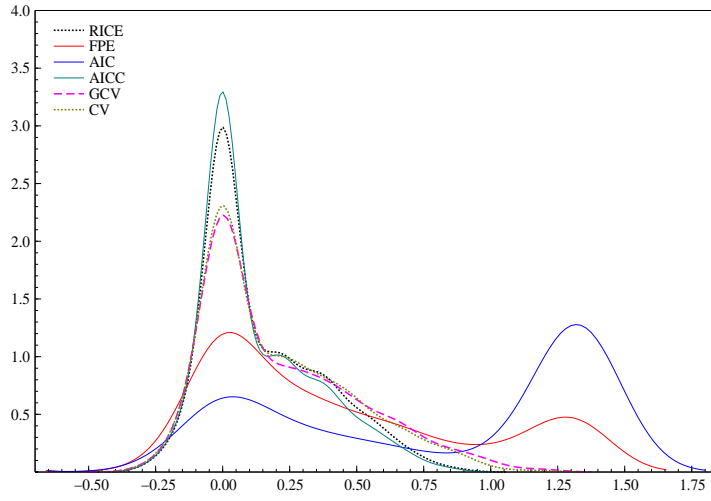Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.

Härdle, W. and Vieu, P. (1992). Kernel regression smoothing of time series. *Journal of Time Series Analysis*, **13,** number 3, 209–224.

Hurvich, C.M., Simonoff, J.S. and Tsai, C.L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike Information Criterion. *Journal of the Royal Statistical Society* B, **60,** 271–293.

Lee, T.-H. and Ullah, A. (2002). Nonparametric bootstrap tests for neglected nonlinearity in time series regression models. *Journal of Nonparametric Statistics* , forthcoming.

Liu, R.Y. (1988). Bootstrap procedure under some non-i.i.d. models. *Annals of Statistics*, **16,** 1696–1708.

Mizrach, B. (1992). Multivariate nearest neighbors forecast of ems exchange rates. *Journal of Applied Econometrics*, **7,** S151–S163.

Montgomery, A.L., Zarnowitz, V., Tsay, R.S. and Tiao, G.C. (1998). Forecasting the U.S. unemployment rate. *Journal of the American Statistical Association*, **93,** number 442, 478–493.

Pagan, A.R. and Schwert, W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, **45,** 267–290.

Potter, S. (1995). A nonlinear approach to US GNP. *Journal of Applied Econometrics*, **10,** 109–125.

Tiao, G.C. and Tsay, R.S. (1994). Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting*, **13,** 109.

Tjøstheim, D. and Auestad, B.H. (1994). Nonparametric identification of nonlinear time series: Selecting significant lags. *Journal of the American Statistical Association*, **89,** number 428, 1410–1419.

Tschernig, R. and Yang, L. (2000). Nonparametric lag selection for time series. *Journal of Time Series Analysis*, **21,** number 4, 457–487.

Yao, Q. and Tong, H. (1998). Cross-validatory bandwidth selections for regression estimation based on dependent data. *Journal of Statistical Planning and Inference*, **68,** 387–415.
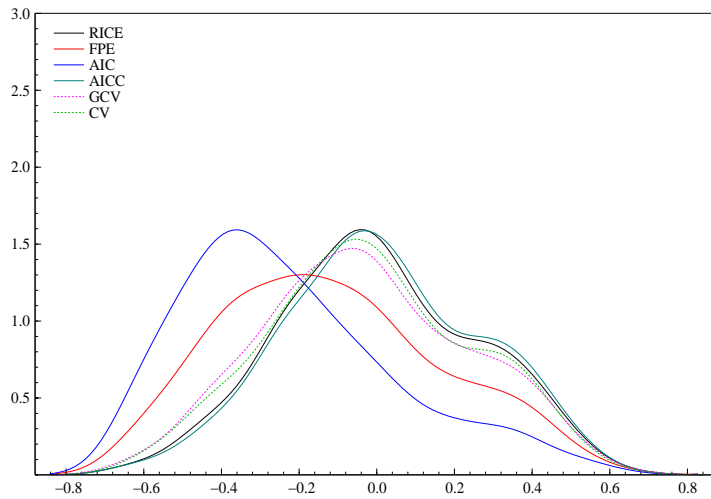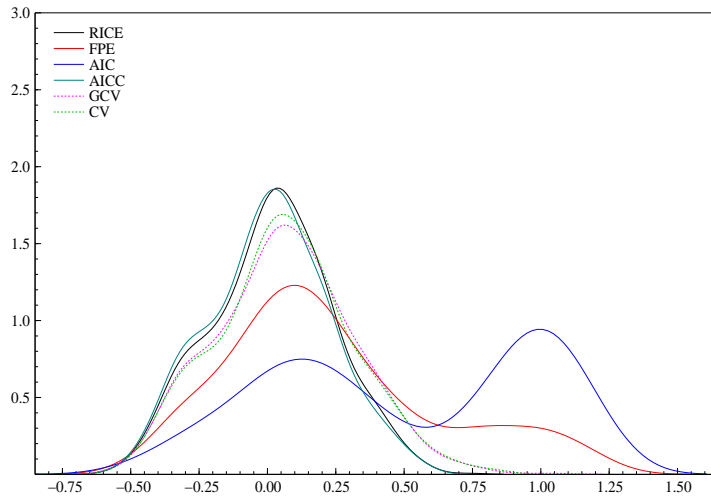
	T	FPE	AIC	$AICC$	GCV	CV
$p = 1$	0.77	0.55	0.36	0.78	0.72	0.68
$p = 2$	0.15	0.15	0.09	0.15	0.17	0.17
$p = 3$	0.05	0.10	0.06	0.05	0.07	0.09
$p = 4$	0.03	0.20	0.49	0.02	0.04	0.06
$\overline{ASE}$	2.49	5.62	8.43	2.43	3.10	2.86

Figure 1: *AR(1) model: smoothed densities of the log of π^{crit}/π^{opt} (top), smoothed densities of the log of h^{crit}/h^{opt} (middle) and mean ASE ratio and frequency of selection of the order (bottom).*

	T	FPE	AIC	$AICC$	GCV	CV
$p = 1$	0.74	0.55	0.34	0.75	0.70	0.68
$p = 2$	0.16	0.18	0.11	0.16	0.17	0.18
$p = 3$	0.08	0.11	0.08	0.07	0.09	0.09
$p = 4$	0.02	0.16	0.47	0.02	0.04	0.05
$\overline{ASE}$	3.00	6.06	8.31	2.89	3.31	3.27

Figure 2: *AR(1)-GARCH(1,1) model: smoothed densities of the log of π^{crit}/π^{opt} (top), smoothed densities of the log of h^{crit}/h^{opt} (middle) and mean ASE ratio and frequency of selection of the order (bottom).*
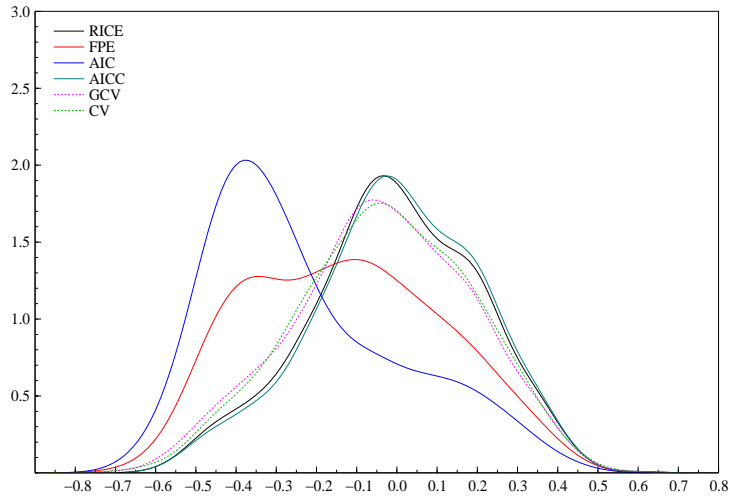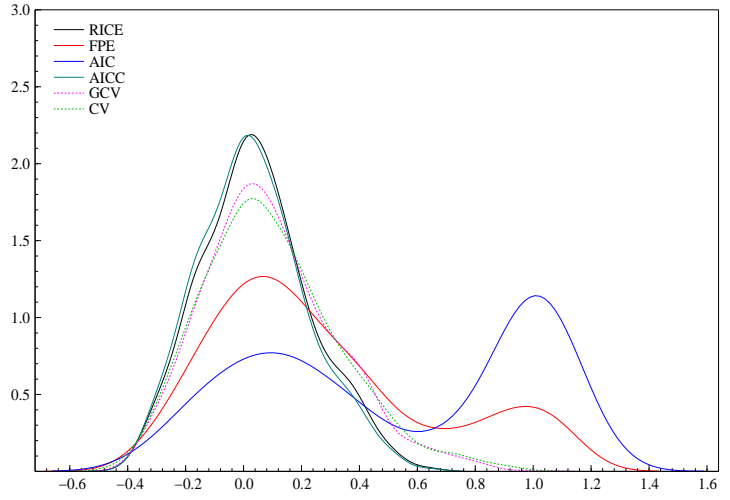
	T	FPE	AIC	$AICC$	GCV	CV
$p = 1$	0.94	0.74	0.49	0.94	0.90	0.88
$p = 2$	0.06	0.10	0.06	0.06	0.08	0.09
$p = 3$	0.00	0.05	0.05	0.00	0.02	0.01
$p = 4$	0.00	0.11	0.40	0.00	0.00	0.01
$\overline{ASE}$	1.31	1.42	1.52	1.31	1.35	1.34

Figure 3: *TAR(1) model: smoothed densities of the log of π^{crit}/π^{opt} (top), smoothed densities of the log of h^{crit}/h^{opt} (middle) and mean ASE ratio and frequency of selection of the order (bottom).*
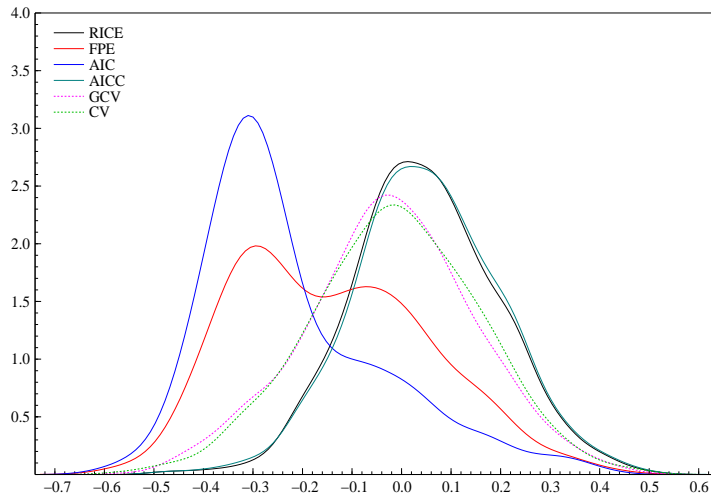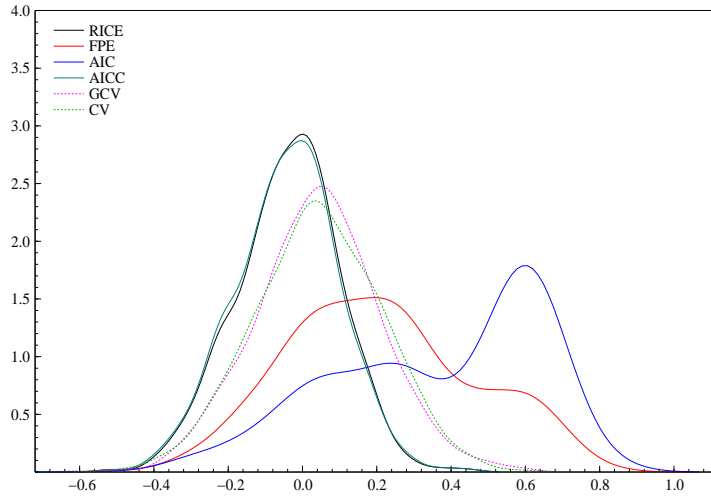
	T	FPE	AIC	$AICC$	GCV	CV
$p = 1$	0.89	0.68	0.46	0.91	0.84	0.82
$p = 2$	0.09	0.14	0.09	0.08	0.12	0.14
$p = 3$	0.01	0.06	0.05	0.01	0.03	0.03
$p = 4$	0.01	0.12	0.40	0.00	0.01	0.01
$\overline{ASE}$	1.43	1.69	1.91	1.42	1.47	1.47

Figure 4: *EXPAR1 model: smoothed densities of the log of π^{crit}/π^{opt} (top), smoothed densities of the log of h^{crit}/h^{opt} (middle) and mean ASE ratio and frequency of selection of the order (bottom).*

	T	FPE	AIC	$AICC$	GCV	CV
$p=1$	0.00	0.00	0.00	0.00	0.00	0.00
$p=2$	0.95	0.68	0.45	0.96	0.89	0.86
$p=3$	0.05	0.16	0.12	0.04	0.10	0.12
$p=4$	0.00	0.16	0.43	0.00	0.01	0.02
$\overline{ASE}$	1.13	1.30	1.39	1.14	1.17	1.16

Figure 5: *EXPAR2 model: smoothed densities of the log of π^{crit}/π^{opt} (top), smoothed densities of the log of h^{crit}/h^{opt} (middle) and mean ASE ratio and frequency of selection of the order (bottom).*

	T	$AICC$	GCV	CV
$AR(1)$	0.06	0.06	0.07	0.06
$AR(1) - GARCH(1,1)$	0.07	0.07	0.08	0.08
$TAR(1)$	0.96	0.96	0.96	0.96
$EXPAR(1)$	0.84	0.85	0.85	0.83
$EXPAR(2)$	0.91	0.91	0.95	0.95

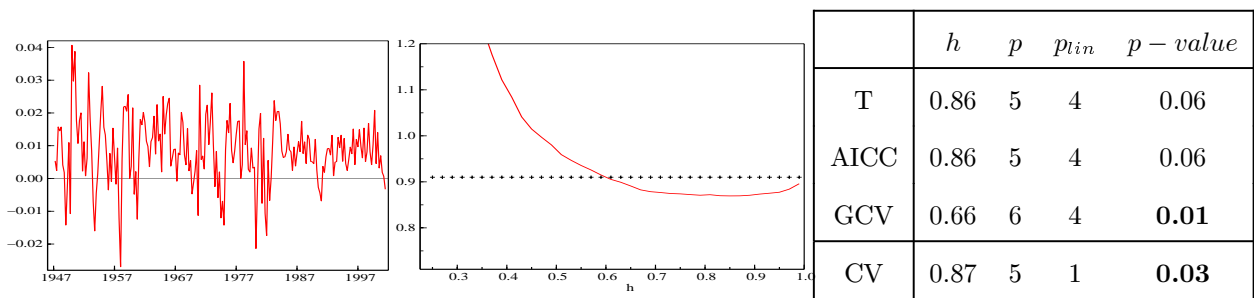**Table 1**: *Frequency of rejection of the null hypothesis of linearity at 5% significance level.*

	h	p	p_{lin}	$p-value$
T	0.86	5	4	0.06
AICC	0.86	5	4	0.06
GCV	0.66	6	4	**0.01**
CV	0.87	5	1	**0.03**

Figure 6: *Plot of the time series, $SC(h,p)$ and parameters chosen with the p-value of the linearity test for US GNP.*

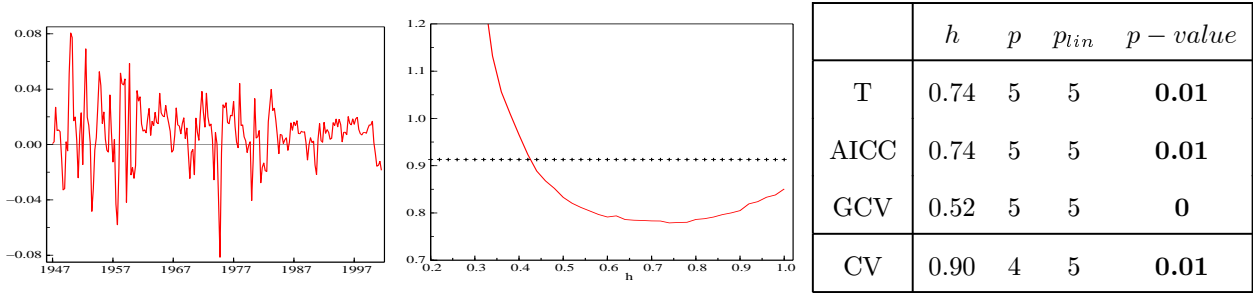|	h	p	p_{lin}	p − value
T	0.74	5	5	**0.01**
AICC	0.74	5	5	**0.01**
GCV	0.52	5	5	**0**
CV	0.90	4	5	**0.01**

Figure 7: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for US Industrial production.*
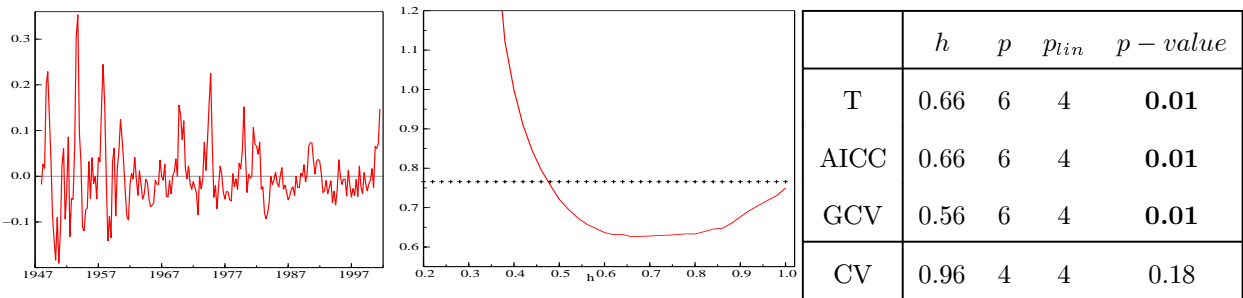

	h	p	p_{lin}	p − value
T	0.66	6	4	**0.01**
AICC	0.66	6	4	**0.01**
GCV	0.56	6	4	**0.01**
CV	0.96	4	4	0.18

Figure 8: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for the unemployment rate.*

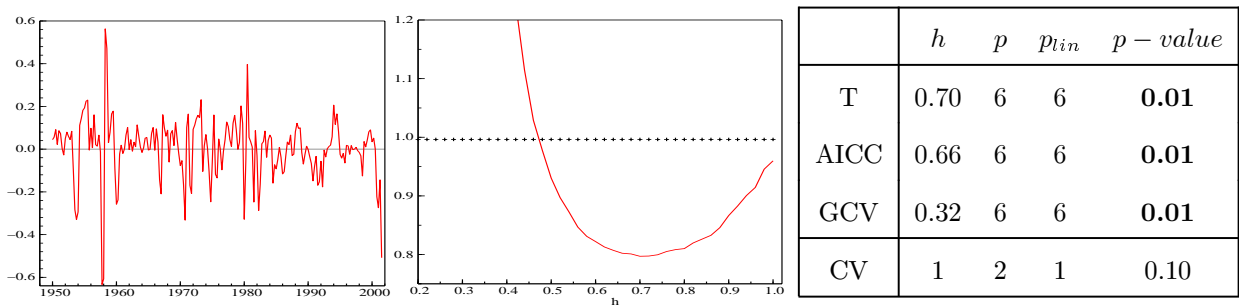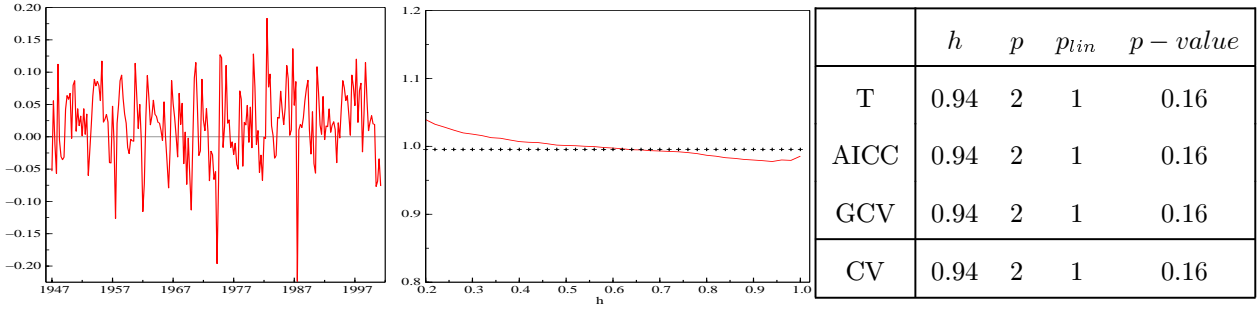|	h	p	p_{lin}	p − value
T	0.70	6	6	**0.01**
AICC	0.66	6	6	**0.01**
GCV	0.32	6	6	**0.01**
CV	1	2	1	0.10

Figure 9: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for T-bills.*

Figure 10: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for S&P500.*
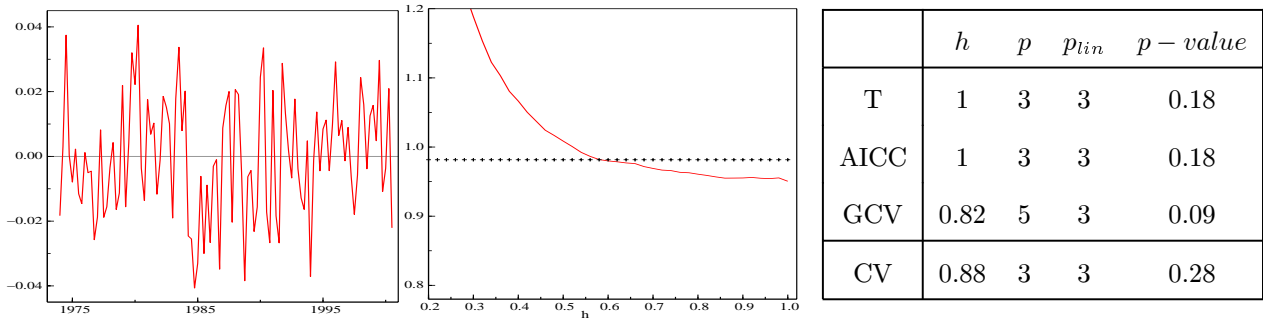
	h	p	p_{lin}	$p-value$
T	0.94	2	1	0.16
AICC	0.94	2	1	0.16
GCV	0.94	2	1	0.16
CV	0.94	2	1	0.16

Figure 11: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for DM/US exchange rate.*

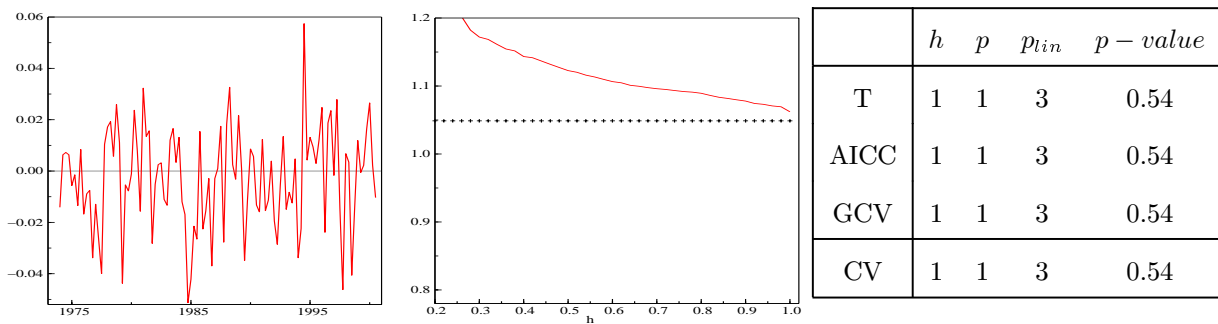|	h	p	p_{lin}	$p-value$
T	1	3	3	0.18
AICC	1	3	3	0.18
GCV	0.82	5	3	0.09
CV	0.88	3	3	0.28

Figure 12: *Plot of the time series, $SC(h, p)$ and parameters chosen with the p-value of the linearity test for Yen/US exchange rate.*

	h	p	p_{lin}	$p-value$
T	1	1	3	0.54
AICC	1	1	3	0.54
GCV	1	1	3	0.54
CV	1	1	3	0.54

24