Mathematical Statistics
Stockholm University

Prediction of industrial production based on nonlinear time series models

Johan Sjöberg

Examensarbete 2010:3

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se/matstat

Mathematical Statistics
Stockholm University
Examensarbete **2010:3**,
http://www.math.su.se/matstat

Prediction of industrial production based on nonlinear time series models

Johan Sjöberg*

Juni 2010

Abstract

The use of linear time series models to model economic data can be put into question by the commonly held view that many economic phenomena is, in some sense, nonlinear. One example is the business cycle that by many is believed to be asymmetric. The expansions of the economy tend to last longer with moderate growth while recessions are short-lived with steep downturns in economic activity. To deal with nonlinearities, several nonlinear times series models have been developed. In this study, we compare autoregressive integrated moving average (ARIMA) models with the nonlinear models logistic smooth transition autoregressive (LSTAR) and self exiting threshold autoregressive (SETAR). We use monthly data of Swedish industrial production in 28 branches over the period January 1990 to December 2008. Each model type is fitted to successive subsets of each time series. Based on the fitted models we make out of sample predictions of the industrial production over the next 12 months. We compare the relative performance of the models as judged by the mean square error of their predictions. Our results show that in general, ARIMA models outperform both SETAR and LSTAR models. ARIMA models performed better over the whole range of the prediction horizon. When we break down the results by industrial branch we find a few branches where LSTAR and SETAR models appear to perform better.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: johan@sjoberg.se. Supervisor: Gudrun Brattström.

Acknowledgements

I would like to thank Sven Öhlén at Statistics Sweden for suggesting non-linear time series models as the general topic of study for this thesis. Sven also suggested industrial production index as a possible data set on which to perform the study. I would also like to thank my supervisor at Stockholm University, Gudrun Brattström, for support and encouragement.

Contents

1 Introduction

It is a rather widely accepted view that the economy is nonlinear in the sense that major economic variables have nonlinear relationships. In economic theory it is common that for example production and investment functions are specified in terms of nonlinear functions. Business cycle theory is an area in economics where it has long been argued for the presence of nonlinear relationships. John Maynard Keynes argued that recessions are characterized by rapid decline in economic activity while economic expansions are characterized by a more moderate rate of change in economic activity that extends for a longer period of time (as compared to recessions). A business cycle with such characteristics would be asymmetric. Linear models (i.e. autoregressive integrated moving average, ARIMA) lack the ability to generate such asymmetric processes and thus might be a less suitable choice for modeling macroeconomic time series that follows an asymmetric business cycle (Granger and Teräsvirta, 1993). An alternative approach would be to use nonlinear models such as threshold autoregressive models (TAR) or smooth transition autoregressive models (STAR). These models have made positive contribution to the analysis of macroeconomic time series, for examples, see Montgomery, Zarnowitz, Tsay, and Tiao (1998) (TAR) and Teräsvirta, van Dijk, and Medeiros (2005) (STAR).

In this study, time series of Swedish industrial production will be modeled with STAR, TAR and ARIMA models. Based on these models we will make predictions of future values of the time series in order to compare their suitability as models for this type of macroeconomic series as judged by the performance of the predictions. The time series of industrial production, industrial production index (IPI) is published by Statistics Sweden (SCB).

2 Theory

2.1 Mean function, covariance function and stationarity

Let $\{X_t\}$ be a time series and let $E(X_t^2) < \infty$. We define the mean function as:

$$\mu_X(t) = E(X_t)$$

We define the covariance function of $\{X_t\}$ as:

$$\gamma_X(r,s) = \text{Cov}(X_r, X_s) = E\left[(X_r - \mu_X(r))(X_s - \mu_X(s))\right]$$

for all integers r och s.

We can now define a weakly stationary process. $\{X_t\}$ is weakly stationary if $E(X_t^2) < \infty$ and if:

$$\mu_X(t) \tag{1}$$

and

$$\gamma_X(t, t+h) \tag{2}$$

is independent of t for each h (Brockwell and Davis, 2002).

2.2 ARMA and ARIMA models

A widely used group of parametric models in time series analysis is autoregressive moving-average models (ARMA). We define a general ARMA process in accordance with Brockwell and Davis (2002): $\{X_t\}$ is an ARMA(p, q) process if $\{X_t\}$ is weakly stationary and it holds for every t that:

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q} \tag{3}$$

and the polynomials $(1 - \phi_1 z - ... - \phi_p z^p)$ and $(1 + \theta_1 z + ... + \theta_q z^q)$ have no common factors. $\{Z_t\}$ is white noise, i.e. a sequence of uncorrelated stochastic variables with the same expected value (here 0) and variance σ^2, which we will denote $\{Z_t\} \sim \text{WN}(0, \sigma^2)$.

A process $\{X_t\}$ is an autoregressive integrated moving-average, ARIMA(p, d, q) process if

$$Y_t := (1 - B)^d X_t \quad \text{is an ARMA}(p, q)\text{-process.} \tag{4}$$

(Brockwell and Davis, 2002). The operator B is the backward shift operator and its application gives $BX_t = X_{t-1}$, d is a non-negative integer. The case $d = 0$ gives directly that X_t is an ARMA process. We can conclude from the preceding that ARIMA models can be used to model non-stationary time series as long as transformation of the original series according to (4) gives a series that can be modeled as a stationary ARMA process.

2.3 SETAR models

We define a TAR model in accordance with Peña Sánchez de Rivera, Tiao, and Tsay (2001):

$$X_t = \phi_0^{(j)} + \phi_1^{(j)} X_{t-1} + ... + \phi_{p_j}^{(j)} X_{t-p_j} + Z_t \quad \text{if } \gamma_{j-1} \leq X_{t-d} < \gamma_j \tag{5}$$

where

$$Z_t \sim \text{WN}(0, \sigma^2) \quad \text{and} \quad -\infty = \gamma_0 < \gamma_1 < ... < \gamma_{k-1} = \infty$$

X_{t-d} is the threshold variable and γ_j, $j = 0, ..., k-1$ are our thresholds. Since the threshold variable is the series d lagged value this version of TAR is referred to as Self Exciting TAR (SETAR).

Within the k different regimes the model follows separate autoregressive models (AR). An AR model is an ARMA model as defined in (3) where $\theta_j = 0$ for $j = 1, ..., q$. In econometric applications, when the aim is to capture the different dynamics that is believed to govern the economy during expansions and contractions, SETAR models are often restricted to contain two regimes.

2.4 LSTAR models

Criticism of the use of TAR models for modeling macro economic time series was partly concerned with the model's sharp switching between the different regimes. Critics claimed that even though individual economic actors might change their behavior in a dramatic and quick fashion, those changes would not occur simultaneously in time and on an aggregated level more smooth transition between the states of the economy would be more realistic. The LSTAR model was suggested as an alternative model that did not suffer from this shortcoming (Granger and Teräsvirta, 1993).

We define a LSTAR model in accordance with Dijk, Teräsvirta, and Franses (2002), however with the modification that we allow the order of the AR-models to differ (p_1 does not have to equal p_2). The following is our defining equation of the LSTAR model:

$$\begin{aligned} X_t = (\phi_{1,0} + \phi_{1,1} X_{t-1} + ... + \phi_{1,p_1} X_{t-p_1}) + \\ (\phi_{2,0} + \phi_{2,1} X_{t-1} + ... + \phi_{2,p_2} X_{t-p_2}) G(X_{t-d}; \gamma, c) + Z_t \end{aligned} \quad (6)$$

where $Z_t \sim \text{WN}(0, \sigma^2)$. the first subscript i, in $\phi_{i,j}$ indicates regime membership and the second subscript, j, indicates which order the parameter has in its regime. Furthermore we have d, γ and c which are constants where $d > 0$ and $\gamma > 0$.

The function $G(X_{t-d}; \gamma, c) = \{1 + \exp(-\gamma(X_{t-d} - c))\}^{-1}$, i.e. the logistic function which has the properties $G(X_{t-d}; \gamma, c) \in [0, 1]$ and it is also continuous. The character of the logistic function depends on the value of γ. In general, higher values of γ gives a more sharp transition between the two regimes. For high values of γ the logistic function tends towards the indicator function $\mathbb{1}_{[X_{t-d}>c]}$. This means that the LSTAR model in such cases, from a practical point of view, is a SETAR model. The parameter c acts as a threshold value and we can note that $G(X_{t-d} = c; \gamma, c) = 0.5$ (Dijk et al., 2002).

In figure 1 we have plotted the logistic function for different values of γ. Higher values of γ shrink the interval of $X_t - c$ that gives values of $G(X_{t-d}; \gamma, c)$ that are neither close to 1 or 0.

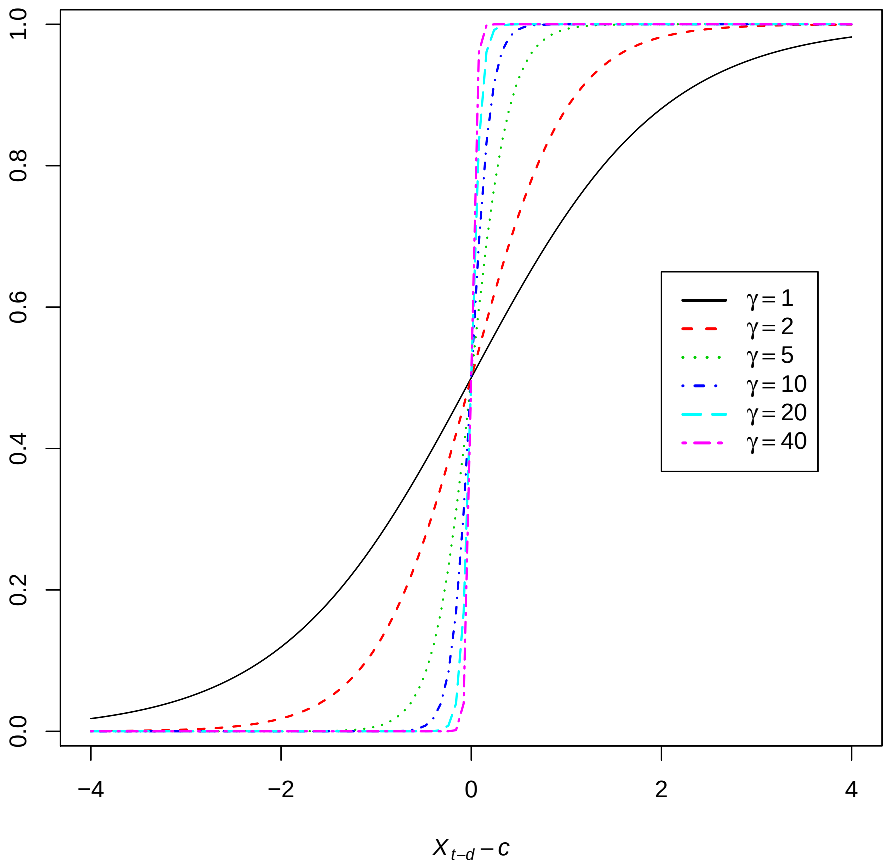


Figure 1: the logistic function for different values of γ

2.5 Methods of prediction

2.5.1 Predictions based on ARMA and ARIMA models

The following description of prediction methods based on ARMA and ARIMA models is mainly based on Brockwell and Davis (2002). Assume that we have observed the n first observations of a stationary time series $\{X_1, ..., X_n\}$. When predicting X_{n+h} , $h > 0$ we will use the best linear predictor by which we intend the linear combination of $1, X_1, ..., X_n$ that minimizes the mean squared error of the prediction. We let $P_n X_{n+h}$ denote the best linear predictor of X_{n+h} based on $1, X_1, ..., X_n$. The predictor takes on the following form

$$P_n X_{n+h} = a_0 + a_1 X_n + ... + a_n X_1$$

and we can decide the coefficients $a_0, a_1, ..., a_n$ by minimizing the function

$$S(a_0, a_1, ..., a_n) = E\left[(X_{n+h} - a_0 - a_1 X_n - ... - a_n X_1)^2\right]$$

We take the partial derivative with respect to each coefficient a_j and put them equal to zero

$$\frac{\partial S(a_0, a_1, ..., a_n)}{\partial a_j} = 0 \ , \ j = 0, 1, ..., n \tag{7}$$

and end up with the following system of equations

$$E\left[X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right] = 0 \tag{8}$$

$$E\left[(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}) X_{n+1-j}\right] = 0 \ , \ j = 1, ..., n \tag{9}$$

We can consider the operator P_n as a special case of a general prediction operator $P(Y|\mathbf{W})$ where Y and $\mathbf{W} = (W_n, ..., W_1)$ are any random variables with finite second moments and known means $\mu = E[Y]$, $\mu_i = E[W_i]$ and known covariances $\text{Cov}(Y, Y)$, $\text{Cov}(Y, W_i)$, $\text{Cov}(W_i, W_j)$. $P(\cdot|\mathbf{W})$ is a function that converts Y into $P(Y|\mathbf{W})$. In the case of P_n we have that $\mathbf{W} = (X_n, ..., X_1)$.

The prediction operator $P(\cdot|\mathbf{W})$ has a number of useful properties that can facilitate the calculation of the best linear predictor.

If we assume that $E[U^2] < \infty$, $E[V^2] < \infty$, $\Gamma = \text{Cov}(\mathbf{W}, \mathbf{W})$ and $\beta, \alpha_1, ..., \alpha_n$ are constants. The prediction operator $P(\cdot|\mathbf{W})$ has (among others) the following three properties (Brockwell and Davis, 2002)

1. $P(\alpha_1 U + \alpha_2 V + \beta|\mathbf{W}) = \alpha_1 P(U|\mathbf{W}) + \alpha_2 P(V|\mathbf{W}) + \beta$
2. $P(\sum_{i=1}^n \alpha_i W_i + \beta|\mathbf{W}) = \sum_{i=1}^n \alpha_i W_i + \beta$
3. $P(U|\mathbf{W}) = E[U]$ if $\text{cov}(U, \mathbf{W}) = 0$

If we now study the defining equation for a stationary AR(p)-model (i.e. an ARMA($p, 0$))

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + Z_t \tag{10}$$

$Z_t \sim \text{WN}(0, \sigma^2)$ and Z_t is uncorrelated with X_s for each $s < t$. For $n > p$ we then apply the prediction operator P_n to both sides of 10 (changing the time index by $t = n + 1$) and receive

$$\begin{aligned} P_n X_{n+1} &= P_n(\phi_1 X_n + ... + \phi_p X_{n+1-p} + Z_n) \\ &= \phi_1 P_n X_n + ... + \phi_p P_n X_{n+1-p} + P_n Z_n \\ &= \phi_1 X_n + ... + \phi_p X_{n+1-p} + P_n Z_n \\ &= \phi_1 X_n + ... + \phi_p X_{n+1-p} \end{aligned} \tag{11}$$

Where we have used, in order, property 1-3 above and thus concluded that the best linear predictor of an AR(p)-model can be determined directly from its defining equation. This result will be reused when we make predictions based on SETAR and LSTAR models.

We will now decide the best linear predictor based on an ARIMA model. We have a time series $\{X_t\}$ and further, according to (4),

$$(1 - B)^d X_t = Y_t, \qquad t = 1, 2, ... \tag{12}$$

where $\{Y_t\}$ is an ARMA(p, q) process. The preceding equation can be rewritten as

$$X_t = Y_t - \sum_{j=1}^{d} \binom{d}{j} (-1)^j X_{t-j}, \qquad t = 1, 2, ... \tag{13}$$

To simplify the notation we will change the indexation of the time t. We assume that we have observed $X_{1-d}, X_{2-d}, ..., X_n$ of the time series $\{X_t\}$ which means that we have observed $Y_1, ..., Y_n$ of the series $\{Y_t\}$. We want to decide $P_n X_{n+h}$, i. e. the best linear predictor of X_{n+h} based on $1, X_{n-d}, ..., X_n$ which is equivalent to deciding the best linear predictor based on $1, X_{1-d}, ...,$ $X_0, Y_1, ..., Y_n$. If we apply the operator P_n to (13) and use the linearity of P_n we get

$$P_n X_{n+h} = P_n Y_{n+h} - \sum_{j=1}^{d} \binom{d}{j} (-1)^j P_n X_{n+h-j} \tag{14}$$

where $t = n + h$.

To decide the best linear predictor we further have to assume that $(X_{n-d}, ..., X_0)$ is uncorrelated with Y_t, $t > 0$. With this assumption we can conclude that $P_n Y_{n+h}$ is a prediction of an ARMA process which we already know how to decide from the beginning of this section. We further note that $P_n X_{n+1-j} = X_{n+1-j}$ for $j \geq 1$. To predict h steps ahead we start with $h = 1$ and decide $P_n X_{n+1}$. We can then decide the right side of (14). For $h = 2$ we repeat the process and use $P_n X_{n+1}$ as if it was the observation of X_{n+1}. By iteration of this process we can decide $P_n X_{n+h}$ for $h \geq 2$.

2.5.2 Predictions based on SETAR models

We determine the one step prediction based on n observations, $P_n X_{n+1}$, by deciding the best linear predictor of the AR model that the process is assumed to follow at time $t = n$. In section 2.5.1 we have shown that the best linear predictor of an AR(p) model is determined by:

$$P_n X_{n+1} = \phi_0 + \phi_1 X_{t-1} + ... + \phi_p X_{t-p} \tag{15}$$

In our case we have SETAR models with two regimes, the threshold variable X_{t-d} and the threshold value γ

$$P_n X_{n+1} = \phi_0^{(j)} + \phi_1^{(j)} X_{t-1} + ... + \phi_{p_j}^{(j)} X_{t-p_j} \tag{16}$$

where j indicates regime membership which is determined by whether or not X_{t-d} is greater than γ.

Predicting h steps ahead, $P_nX_{n+h},\ h > 1$ is achieved by iterating prediction in one step, h times. In each iteration we use the previous prediction as an observation.

2.5.3 Predictions based on LSTAR models

Prediction in one step based on n observations, P_nX_{n+1}, is decided by determining the best linear predictor of the AR model that the process is assumed to follow at time $t = n$. The reasoning is analogous to the one used when deciding prediction for SETAR models. The LSTAR model at time t is assumed to follow an AR($max(p_1, p_2)$) model in accordance with (6):

$$X_t = (\phi_{1,0} + \phi_{1,1}X_{t-1} + ... + \phi_{1,p_1}X_{t-p_1}) + \\ (\phi_{2,0} + \phi_{2,1}X_{t-1} + ... + \phi_{2,p_2}X_{t-p_2})G(X_{t-d}; \gamma, c) + Z_t$$

The one step best linear predictor is known from section 2.5.1:

$$P_nX_{n+1} = (\phi_{1,0} + \phi_{1,1}X_{t-1} + ... + \phi_{1,p_1}X_{t-p_1}) + \\ (\phi_{2,0} + \phi_{2,1}X_{t-1} + ... + \phi_{2,p_2}X_{t-p_2})G(X_{t-d}; \gamma, c)$$

As with SETAR-models we decide prediction in h steps, $h > 1$, by iterating prediction in one step h times and for each new iteration we use the previous prediction as an observation.

3 The Study

3.1 Description of the data

In order to make a comparison of the three models at hand we apply them to data from industrial production index (IPI) which is produced by Statistics Sweden. The times series are seasonally and calendar adjusted. Seasonal effects can be for example holiday shutdown of industries and christmas shopping. Calendar effects are for example effects due to the number of working days during a certain period (SCB, 2010). The reason for using seasonally adjusted data is that there are no established methods of seasonal adjustment based on the nonlinear models studied here.

IPI is constituted of a great number of indices over production in different industrial branches. The indices are produced and published monthly. The indices we use in this study cover the period January 1990 up to december 2008. The indices are classified according to the standard NACE 2002. We are only using indices that are complete over the selected period. We have also made the selection of indices in order to maximize the number of time series without including indices that overlap. To clarify, IPI might include indices for NACE X, Y and Z but also a more highly aggregated index of NACE X, Y and Z together (NACE X-Z). In such a case, we choose to include the indices for X, Y and Z, but not the combined index X-Z because the latter is an aggregate of the former and adds little new information. In total 28 indices are included in the study. Further description of IPI can be found in the appendix (7.2) as well as lists of the included indices (7.2.1) and plots (7.2.2).

Each of our 28 indices constitutes a time series $\{X_t\}_{t=1}^{228}$. From each series we create 128 partial time series according to the following pattern: $\{X_t\}_{t=1}^{101}$, $\{X_t\}_{t=1}^{102}$,...,$\{X_t\}_{t=1}^{228}$. This gives us a total of 3584 partial series. To each of these we will apply our three types of models (ARIMA, SETAR and LSTAR). Based on each type of model we will predict the index value of the 12 coming months and based on theses predictions we can calculate the MSE.

3.2 Transformation of data

When modeling economic time series, it is common practice to transform the series by taking the logarithm. The primary motive for this transformation is stabilization of the variance. A study by Luetkepohl and Xu (2009) investigates under what circumstances log transformation has a positive impact on the accuracy of forecasts. The study shows that log transformation has a positive impact on forecast accuracy if taking logs contributes to more stable variance. It is also shown that forecast accuracy suffers if log transformation is made without achieving a more stable variance. A study of the time series plots in section 7.2.2 gives that the series with NACE code 24.4, 32 and 34 seem to have unstable variance and thus we have chosen to apply the models to the log transformed series in these particular cases.

3.3 Handling of outliers

Extreme observations (outliers) in time series can strongly influence the parameter estimation and other aspects of the modeling process (Peña Sánchez de Rivera et al., 2001). This is something we have to take into account in the study at hand, but we wish to refrain from getting engrossed in the countless theoretical and practical aspects of handling outliers. For our purpose -

which is to make sure that the presence of outliers does not effect the general comparison of the models -the rather simple method outlined in this section will suffice. Observations that are treated as outliers are marked in the plots in section 7.2.2. The study is performed on the outlier adjusted time series.

When adjusting the ARIMA-models we will fix $d = 1$ which in practice means that we apply an ARMA model to $Y_t = X_t - X_{t-1}$. In the same way, we will apply the LSTAR and SETAR models to Y_t. For this reason we can perform outlier adjustment directly on Y_t. An observation will be considered an outlier if

$$\frac{|y_t - \hat{\mu}_{Y_t}|}{\hat{\sigma}^2_{Y_t}} > 2.58$$

where 2.58 is the 99.5-th percentile of the normal distribution, $\hat{\mu}_{Y_t}$ is the sample mean and $\hat{\sigma}^2_{Y_t}$ is the sample variance. An outlier is replaced with a linearly interpolated value according to the following:

$$\hat{y}_t = y_{t-h} + h\,\frac{y_{t+k} - y_{t-h}}{k+h} \tag{17}$$

If an outlier is not preceded nor succeeded by an outlier then $h = k = 1$. When several outliers occur in sequence h is the number of observations since the last non-outlier and k is the number of observations from t to the next non-outlier observation. To illustrate our method in action we have applied it to the index of NACE 10-12. The result is plotted in figure 2.

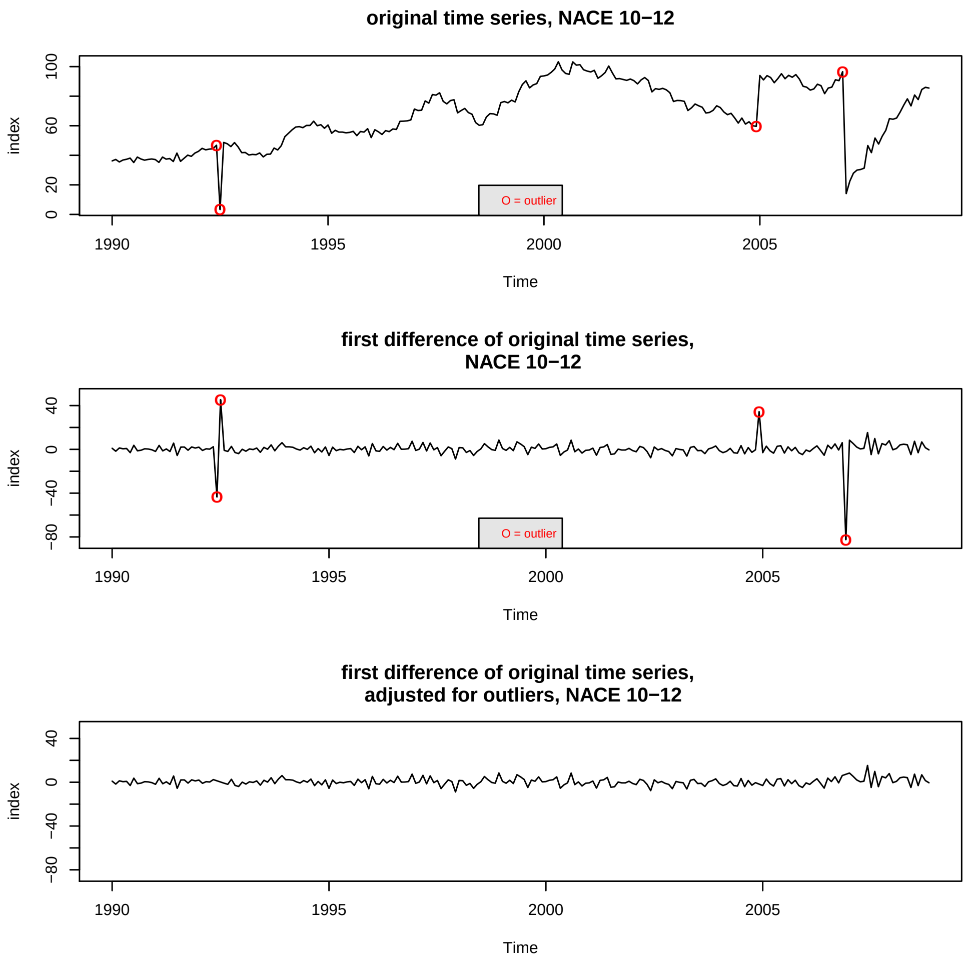


Figure 2: Procedure for handling outlier as applied to index of NACE 10-12

3.4 Model selection - ARIMA

When choosing a model with good properties for prediction, one is not well served by letting p and q be of arbitrarily high order. In general, the white noise variance of a model of higher order will tend to be lower. But the MSE of the predictions depends not only on the white noise variance but also on errors of the parameter estimates. A model of higher order tends to have greater such errors because of the greater number of parameters (Brockwell and Davis, 2002). We will thus need a criteria for model selection that punishes overparameterization of the models. Akaike's information criterion(AIC) is one of several criteria that is developed to serve this purpose. The definition of AIC is $2k - 2\log(L)$ where L is the maximum of the likelihood function and k is the number of parameters in the model. When selecting among models the one with the lowest AIC is preferred. For each series we apply a procedure for automatic model selection that exists in the software R (R Development Core Team, 2009), provided by the package forecast (Hyndman, 2009). The procedure compares ARIMA models of different order where $d = 1$ and p and q is allowed to vary up to maximum 15. The motivation for fixing $d = 1$ is that the we need to take the first difference of the series in order to achieve stationarity (se section 7.3). A more detailed account of the estimation of ARIMA models can be found in section 7.1.1.

3.5 Model selection - SETAR

The model selection process for SETAR models will also be based on AIC. To each series, an automatic model selection process supplied in R through the package tsDyn (Antonio et al., 2008) will be applied. The procedure will compare models where p_1 and p_2 are allowed to vary up to maximum 8 and d is allowed to take on the values 1, 2 and 3. The procedure will only take into account models with a threshold γ that results in a model where the regime with the least number of observations contains at least 15 percent of the total number of observations.

The SETAR models are applied to the outlier adjusted first difference of the original series $Y_t = X_t - X_{t-1}$ where X_t is the original series. Details on the estimation of SETAR models can be found in section 7.1.2.

3.6 Model selection - LSTAR

Consistent with the reasoning in section 3.4 and 3.5 we let the model selection process for LSTAR be guided by AIC to prevent overparameterization. The models will be applied to the outlier adjusted first difference of the original series $Y_t = X_t - X_{t-1}$ where X_t is the original series. A procedure for automatic model selection is provided in R through the package tsDyn (Antonio et al., 2008). All models where $d = 1$ and p_1 and p_2 are allowed to take on the maximum value of 6 are considered. Details of the estimation process for LSTAR models can be found in section 7.1.3.

3.7 Trimming of predictions

In some studies where predictions based on a large number of models are studied, a technique for *trimming* the predictions is applied. Trimming refers to the procedure of replacing an "unreasonable" prediction with a more conservative one according to some rule. Stock and Watson (1998) employed such a technique that we will reuse in this study, If $|P_n Y_{n+h}| \geq \max(|Y_t|)$ it is replaced with $\hat{\mu}_{Y_t}$, where $\hat{\mu}_{Y_t}$ is the sample mean of Y_t. The motivation for trimming is that it is a way to simulate the involvement of humans in the forecasting procedure. A human forecaster would disregard model based predictions that by all reasonable standards is deemed as just plain foolish. There could be situations where the predictions - even though judged as unreasonable - still contains some information about where the series is headed. One could therefore argue that instead of just ignoring the predictions by replacing them with the sample mean we could apply a trimming rule that just moderates the predictions. However, it was clear that in the majority of cases where the trimming procedure came into play in this study, the models had produced predictions that just seemed raving mad.

4 Results

4.1 SETAR versus ARIMA

We let MSE_{ijk} denote MSE of model type i, time series j and prediction horizon k. A comparison of ARIMA and SETAR models shows that in 254 of 336 cases ARIMA models produces predictions with a lower MSE while in 82 cases SETAR models produces better predictions. In figure 3 we have plotted the ratio of $\text{MSE}_{\text{SETAR},jk}/\text{MSE}_{\text{ARIMA},jk}$. A ratio greater than 1 thus means that the ARIMA models made better predictions and a ratio less than 1 means that SETAR models predicted better.

In figure 5, we break down the results by time series (index). From this figure we can se that the relative performance of the models tend to differ between different series. There are 13 series where ARIMA models perform better on every prediction horizon. For 6 of the 28 series SETAR models perform better than ARIMA models on at least half of the predictions horizons. There are no series where SETAR models perform better than ARIMA models over every prediction horizon. In figure 4, we break down the MSE ratios by prediction horizon. From this figure we get the impression that the relative performance of the models does not seem to vary with the length of the prediction horizon.

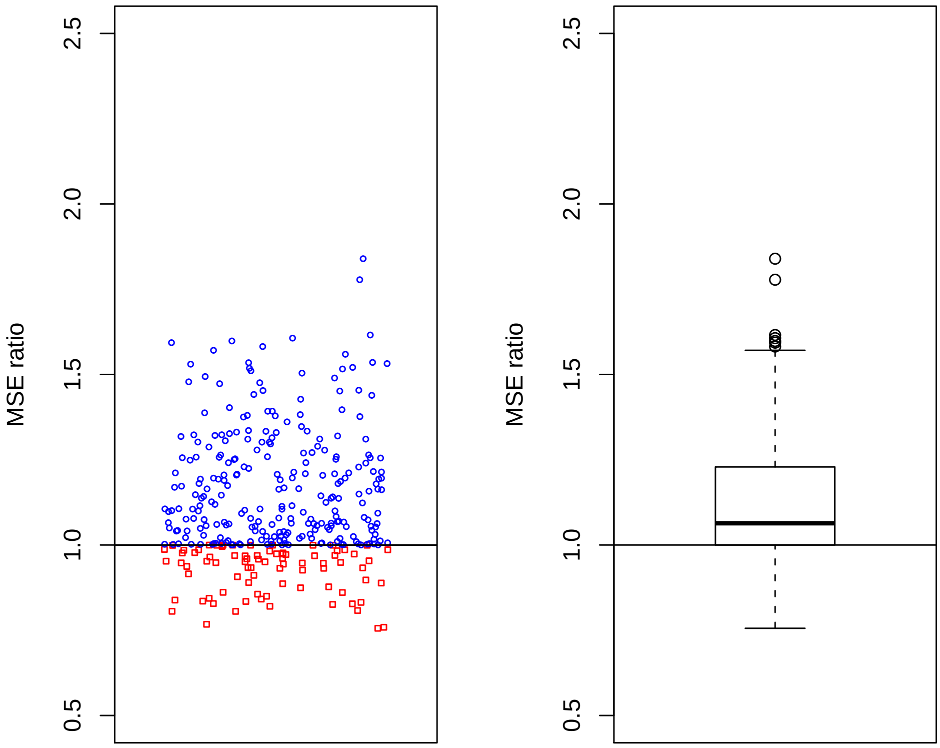


Figure 3: MSE ratio (SETAR/ARIMA)

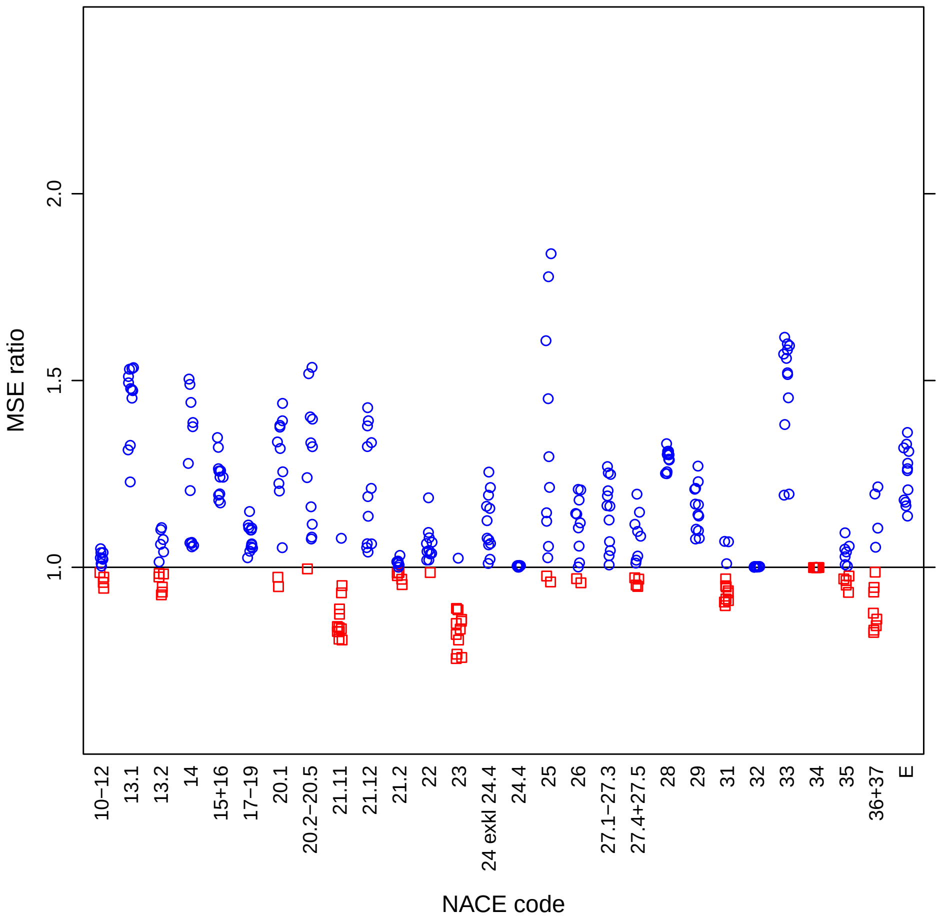


Figure 4: MSE ratio (SETAR/ARIMA), by NACE code

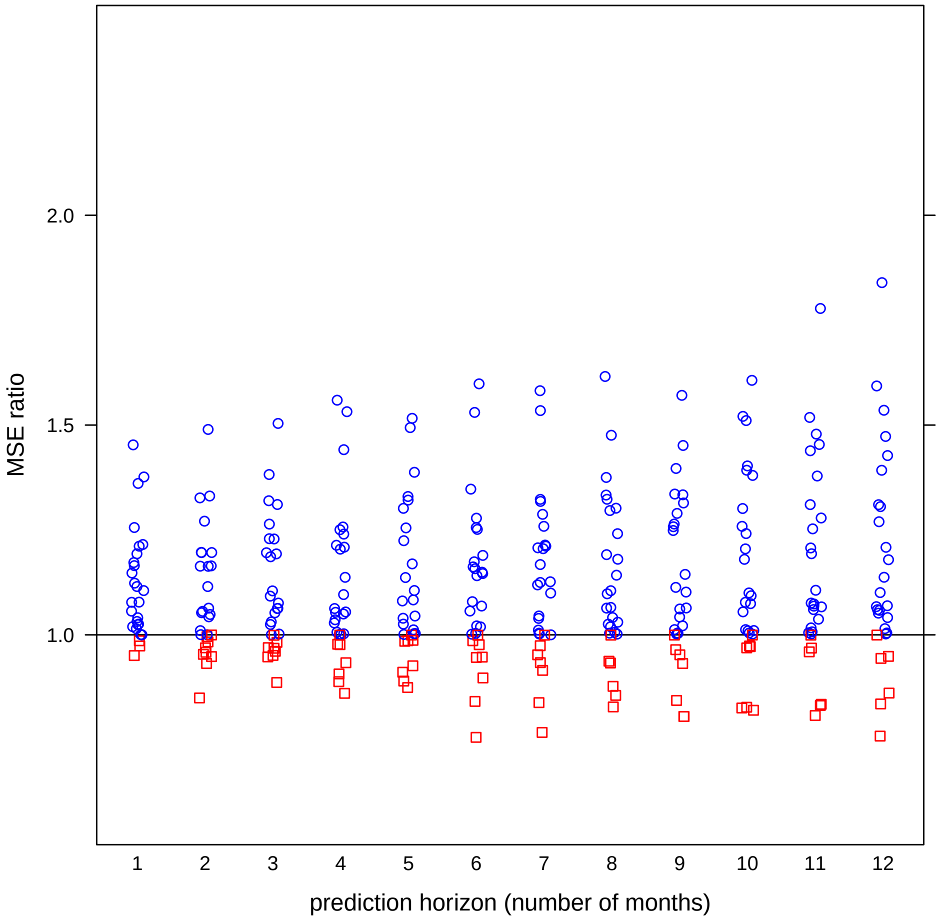


Figure 5: MSE ratio (SETAR/ARIMA), by prediction horizon

The procedure for trimming predictions affected in total 17 SETAR models (out of a total of 3584) as summarized in table 4.1.

Table 1:

NACE code	number of SETAR models affected by trimming
10-12	2
15+16	2
20.1	1
24 excl. 24.4	1
27.4+27	2
29	2
33	2
35	2
E	3

4.2 LSTAR versus ARIMA

In figure 6, we have plotted the ratio $\mathrm{MSE}_{\mathrm{LSTAR},jk}/\mathrm{MSE}_{\mathrm{ARIMA},jk}$. In 219 of 336 instances ARIMA models outperformed LSTAR models. In the remaining 117 cases LSTAR models performed better than ARIMA. In figure 7, we break down the MSE ratios by time series (index). The figure shows that the relative performance of the models vary between the series. For 11 series ARIMA models outperformed LSTAR models on all prediction horizons. For 10 series LSTAR models achieve better predictions on at least half of the prediction horizons. In figure 8, we break down the results by prediction horizon. We can se that on all horizons ARIMA models perform better. The relative performance of the models does not seem to be depending on the prediction horizon.

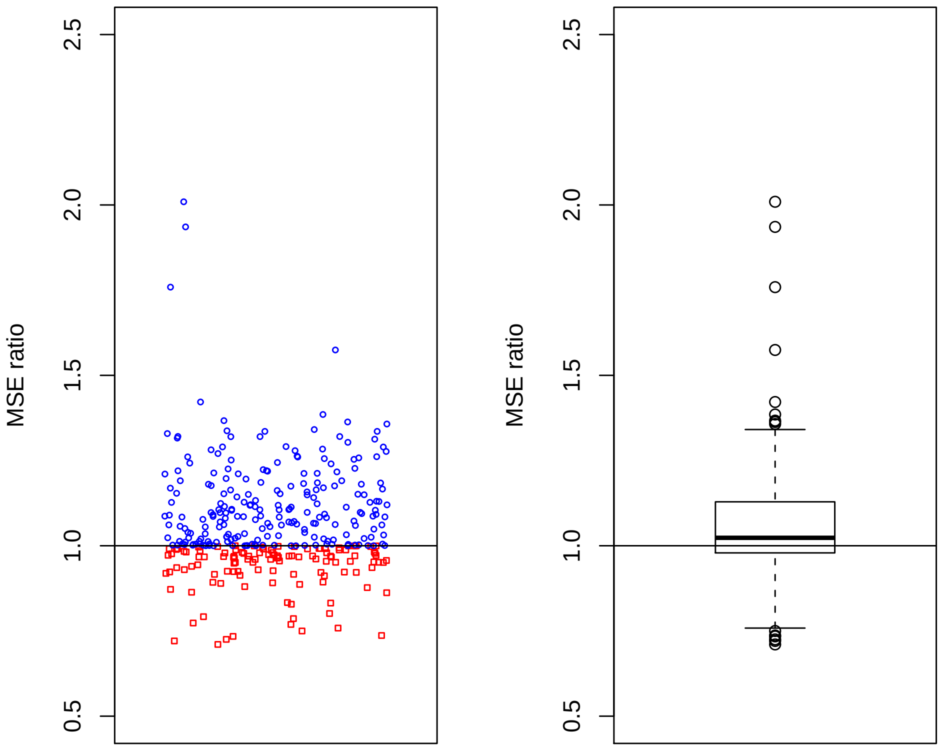


Figure 6: MSE ratio (LSTAR/ARIMA)

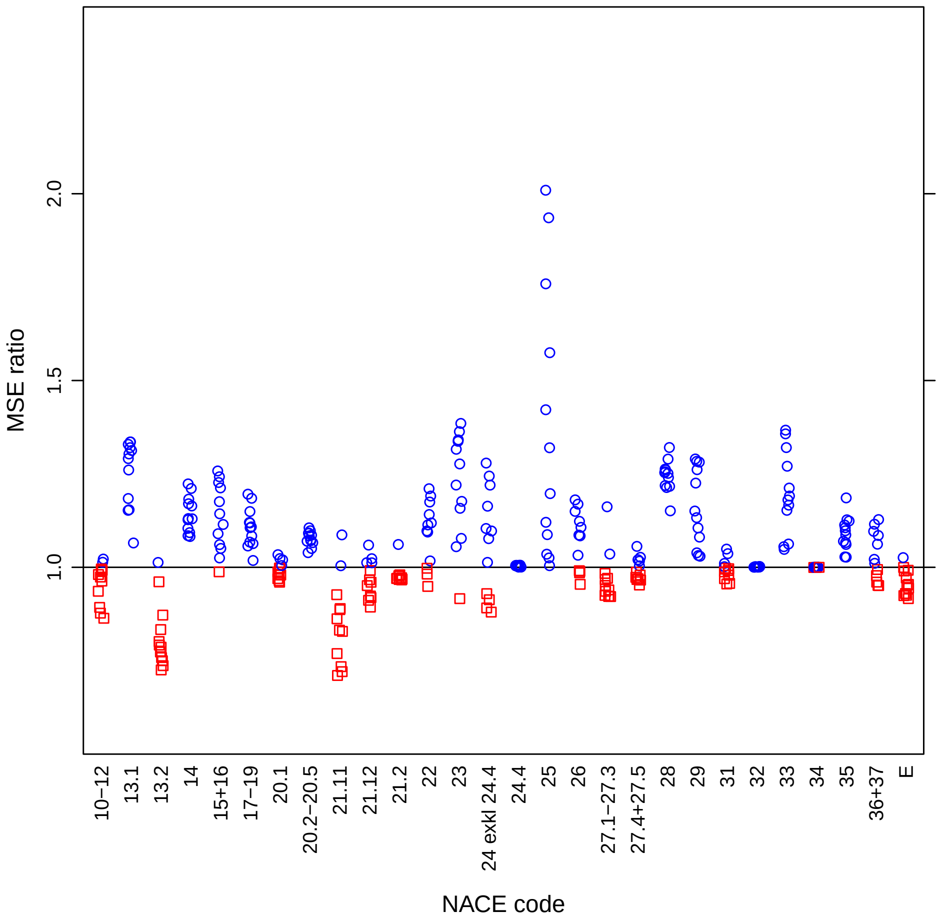


Figure 7: MSE ratio (LSTAR/ARIMA), by NACE code

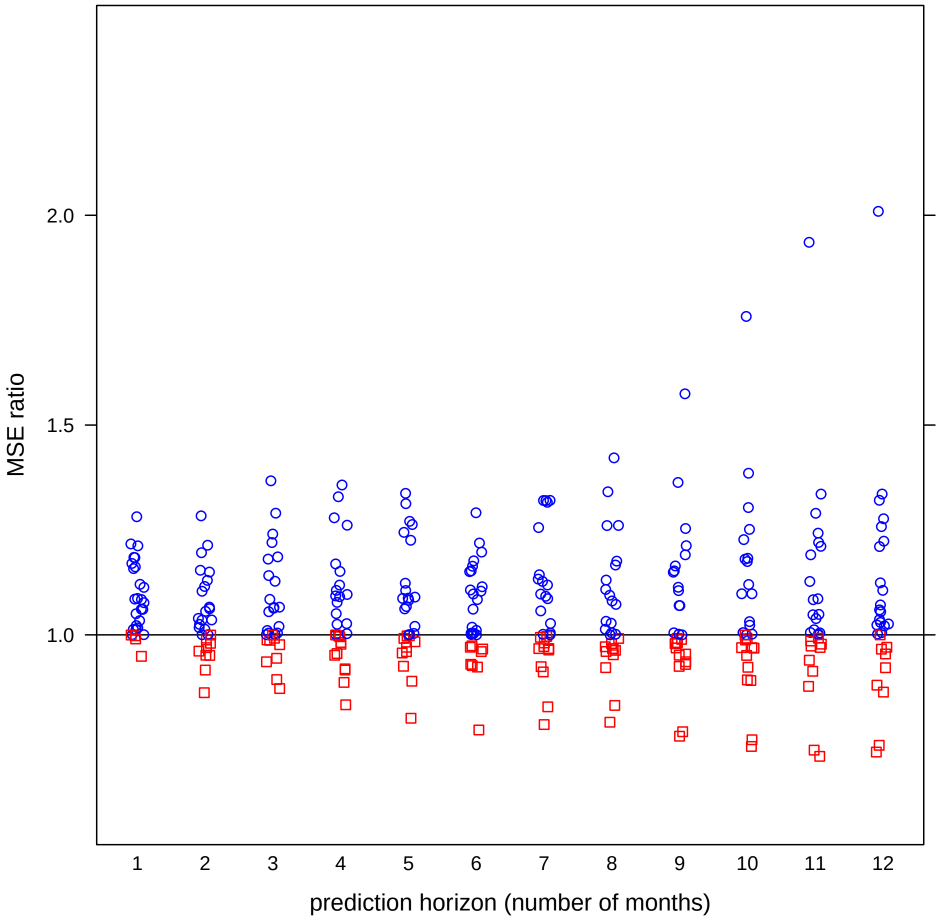


Figure 8: MSE ratio (LSTAR/ARIMA), by prediction horizon

The procedure for trimming predictions affected in total 28 LSTAR models (out of a total of 3584) as summarized in table 2. The ARIMA models were subjected to the same trimming procedure, but no predictions were trimmed.

We finally take a look at the values of γ received when fitting our LSTAR models. In figure 9 boxplots of the γ values for the different indices are displayed. It is quite obvious that in many cases the LSTAR model has values that are so great that from a practical aspect makes it behaves much like a SETAR model. The indices in the study have an accuracy of one tenth of an index point. We recall the logistic function $\{1+\exp(-\gamma(X_{t-d}-c))\}^{-1}$. For example, if $\gamma = 40$ and $X_{t-d} - c$ changes from 0.1 to -0.1 the logistic function changes from approximatively 0.98 to 0.02.

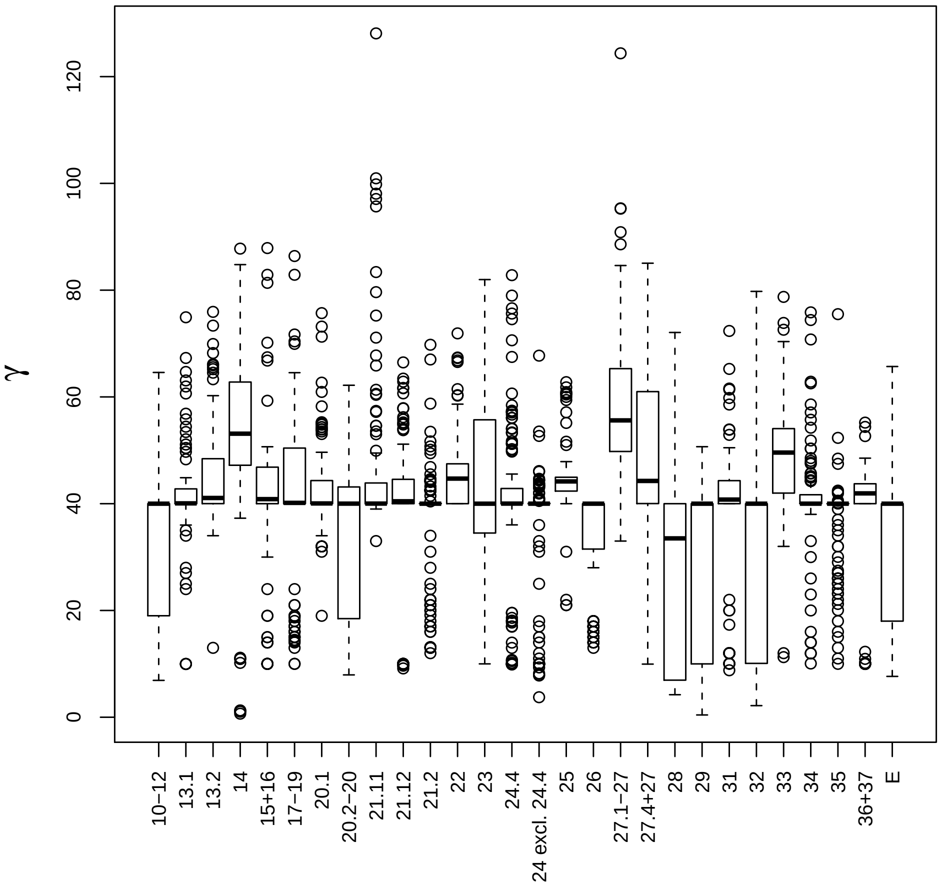


Figure 9: Estimated γ values by NACE code

Table 2:

NACE code	number of LSTAR models affected by trimming
10-12	2
13.2	1
14	1
15+16	1
17-19	1
20.2-20	3
23	2
24 excl. 24.4	3
27.4+27	1
28	1
29	2
33	2
35	1
36+37	1
E	6

5 Conclusions

It is clear that ARIMA models produce better predictions than both SETAR and LSTAR models in this study. Prediction based on ARIMA models perform better over the whole range of the prediction horizon considered (1-12 months). There are individual time series where LSTAR and SETAR models manage to perform better, but for the majority of the 28 series considered, ARIMA models are clearly better. One possible explanation is that for most branches of industry, the development of Swedish production over time is not nonlinear, at least not in a sense that is well captured by LSTAR and SETAR models. There seems to be no reason for using LSTAR and SETAR models as an alternative to ARIMA for modeling Swedish industrial production in general. Someone with an interests in particular industrial branches and with sufficient resources at their disposal might want to consider SETAR or LSTAR models as options or complements to ARIMA models. Careful examination of how the models preform when applied to the actual data could in some cases reveal benefits of choosing a nonlinear model.

6 Tables

Table 3: MSE of ARIMA models by NACE and prediction horizon (number of months)

NACE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
10-12	75.29	131.57	185.49	234.25	282.20	333.89	386.80	432.35	458.03	497.08	543.97	590.45
13.1	7.90	6.95	7.77	11.26	13.76	16.98	22.56	26.81	28.84	39.94	50.93	60.57
13.2	22.24	19.40	23.38	28.11	32.58	43.38	52.68	63.69	80.09	97.22	117.45	143.53
14	17.74	24.14	45.27	61.27	85.13	102.64	119.89	132.96	152.10	170.61	188.15	207.67
15+16	1.32	1.75	2.44	2.80	3.59	4.63	5.27	6.37	7.57	8.98	10.10	11.89
17-19	16.33	19.23	20.72	22.71	28.66	30.11	35.30	41.76	42.39	48.35	52.62	63.81
20.1	5.84	7.06	8.47	11.18	14.16	17.35	21.47	27.34	29.09	32.78	37.93	43.08
20.2-20	10.83	9.41	11.77	18.56	18.68	25.38	33.20	39.82	48.75	53.15	62.45	77.56
21.11	6.29	6.02	8.92	10.34	12.12	14.50	16.70	18.90	21.08	25.02	27.00	32.25
21.12	3.02	3.98	4.57	6.72	7.62	9.04	10.85	13.33	16.55	18.80	20.71	24.38
21.2	5.83	9.37	14.47	20.70	27.03	35.16	43.21	51.08	60.18	68.64	76.46	84.89
22	2.41	3.67	4.58	5.48	7.40	9.25	11.03	13.97	16.49	19.37	21.64	25.74
23	10.83	9.84	10.96	13.58	14.28	16.31	19.72	23.46	25.74	28.25	37.75	39.99
24.4	43.07	51.63	56.84	77.10	105.62	133.45	159.89	202.87	244.65	284.49	312.65	392.14
24 excl. 24.4	9.03	11.12	12.86	15.35	18.47	21.41	25.93	31.40	37.57	43.46	50.65	57.93
25	2.77	3.87	5.43	7.63	10.75	14.77	19.02	24.31	29.15	35.63	43.86	50.25
26	7.22	9.02	10.36	12.11	14.15	19.50	23.65	27.80	31.54	34.90	42.45	49.94
27.1-27	17.88	23.31	24.31	25.55	33.33	40.06	52.00	58.56	63.95	69.12	77.14	89.29
27.4+27	53.07	63.88	83.84	85.69	95.81	107.48	137.34	164.97	179.99	220.16	246.30	275.14
28	3.21	3.76	5.27	7.93	10.79	14.35	18.81	24.79	29.07	35.69	41.94	48.60
29	9.14	11.40	13.41	17.13	21.76	25.57	32.74	39.23	45.50	54.22	62.01	78.18
31	48.19	50.91	56.39	60.44	63.48	67.16	70.86	77.72	84.07	99.79	104.71	114.06
32	131.73	162.89	154.61	166.51	201.15	190.75	203.35	225.03	273.00	298.77	333.31	392.56
33	25.20	28.57	38.65	47.33	47.62	64.52	74.99	83.78	91.56	97.09	103.88	143.09
34	27.33	38.94	50.74	66.76	87.63	101.00	119.13	135.39	153.09	163.98	179.30	200.29
35	10.32	12.12	14.54	16.28	18.05	20.41	23.43	22.13	25.26	28.53	33.08	36.79
36+37	8.75	14.01	17.61	20.56	23.31	27.57	32.16	34.44	37.88	44.29	51.88	58.97
E	2.06	1.86	4.63	5.11	8.78	9.95	13.30	16.27	21.38	25.94	30.52	37.13

Table 4: MSE of SETAR models by NACE and prediction horizon (number of months)

NACE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
10-12	78.06	132.01	179.55	222.15	269.16	318.35	364.99	406.03	427.20	455.76	497.31	540.21
13.1	6.44	6.05	6.73	9.77	12.09	14.32	19.41	22.90	25.29	34.46	46.01	54.93
13.2	22.21	18.97	20.76	25.09	28.18	35.42	42.52	48.43	57.21	67.86	77.00	96.06
14	15.09	18.31	32.65	46.46	66.41	88.67	112.13	141.10	166.42	191.27	213.60	240.03
15+16	1.19	1.50	1.91	2.34	2.97	3.83	4.78	6.04	7.30	8.88	10.52	12.69
17-19	18.87	22.05	20.74	24.09	27.67	26.67	33.94	41.87	43.77	49.23	53.79	67.10
20.1	6.20	7.36	8.21	9.27	11.54	13.27	15.70	19.47	21.38	23.03	26.49	31.67
20.2-20	10.20	9.83	11.66	16.32	19.11	23.68	27.41	32.03	37.35	41.60	44.67	54.12
21.11	6.34	5.57	9.41	10.32	12.33	15.97	16.50	18.98	20.13	22.19	23.75	27.82
21.12	2.94	3.60	3.84	5.81	6.43	7.02	8.16	9.74	12.30	13.81	15.20	18.09
21.2	6.00	9.63	14.60	20.70	26.64	34.45	42.00	49.11	57.58	65.64	72.93	81.16
22	2.24	3.51	4.41	5.20	7.48	9.41	11.81	15.00	17.61	20.81	24.84	29.19
23	12.24	10.61	13.05	17.00	21.47	25.38	33.82	36.76	43.57	47.70	55.20	67.29
24.4	43.07	51.64	56.85	77.12	105.68	133.51	159.98	203.04	244.79	284.74	312.82	392.46
24 excl. 24.4	9.02	10.55	13.15	16.18	18.31	21.52	25.30	29.91	34.18	38.34	43.08	48.09
25	2.77	3.80	5.68	8.01	11.40	15.43	20.68	26.66	31.62	39.00	47.75	54.89
26	7.08	10.82	12.62	14.15	15.70	20.42	22.96	25.13	27.27	29.30	34.65	39.44
27.1-27	17.84	20.74	22.07	24.15	31.37	36.38	44.64	45.32	47.36	52.93	57.86	64.82
27.4+27	47.03	60.49	69.19	80.27	90.24	105.87	131.37	152.54	185.07	219.59	247.53	280.10
28	3.11	3.43	4.98	7.30	10.47	13.98	18.35	24.00	28.26	34.33	41.28	49.17
29	9.67	11.52	14.08	17.87	22.81	25.78	31.76	38.62	45.64	51.91	59.88	70.78
31	50.58	50.94	60.14	63.67	66.67	74.97	77.11	82.22	88.32	97.61	102.78	110.56
32	131.72	162.89	154.63	166.46	201.10	190.59	203.20	224.95	272.98	298.74	333.31	392.48
33	25.61	25.38	38.23	41.19	39.91	46.52	62.60	60.48	69.41	75.36	74.89	94.77
34	27.31	38.93	50.77	66.79	87.69	101.08	119.23	135.48	153.18	164.08	179.43	200.37
35	10.87	12.33	15.79	17.51	19.56	22.18	25.27	24.39	28.01	32.33	37.02	39.72
36+37	7.82	13.07	17.97	21.39	25.08	29.45	34.24	37.70	42.72	51.01	60.99	69.94
E	1.50	1.52	3.32	4.11	6.11	7.88	10.18	13.40	16.14	20.60	23.68	29.08

Table 5: MSE of LSTAR models by NACE and prediction horizon (number of months)

NACE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
10-12	104.57	137.94	192.88	234.91	280.24	337.85	382.36	437.94	468.38	517.19	563.18	613.35
13.1	6.42	14.16	8.51	10.24	14.08	14.12	17.65	22.40	20.64	24.83	25.69	35.20
13.2	22.37	17.19	23.34	28.66	31.78	40.55	49.46	55.97	65.76	78.55	85.29	99.35
14	23.56	17.90	32.35	45.85	63.78	84.06	102.63	126.08	144.37	163.66	176.84	197.02
15+16	1.23	1.87	1.79	2.09	2.72	3.08	3.73	4.61	5.21	6.23	7.02	8.57
17-19	18.60	23.42	23.21	23.61	30.83	28.68	33.25	40.37	38.13	44.59	46.96	61.01
20.1	5.61	9.02	8.35	9.32	9.98	12.72	13.35	17.57	17.72	19.97	20.29	26.98
20.2-20.5	9.70	12.37	11.75	15.46	17.80	22.65	24.56	29.65	34.91	38.58	39.89	51.44
21.11	5.39	8.35	9.37	12.46	11.61	15.00	16.88	18.73	20.17	24.35	25.30	31.11
21.12	2.97	3.67	4.61	6.48	6.61	7.32	8.77	10.11	12.65	13.75	14.92	18.39
21.2	6.82	12.41	15.74	21.83	26.02	35.24	40.60	48.21	55.42	64.81	67.59	78.37
22	2.53	3.54	4.14	5.45	8.32	8.37	10.39	13.91	15.88	17.48	20.48	23.94
23	11.04	16.80	15.19	14.33	15.14	17.86	20.94	21.31	25.89	24.76	27.86	37.91
24.4	53.78	60.92	60.54	103.12	98.28	156.86	153.44	214.31	219.39	254.28	278.60	349.52
24 exkl 24.4	9.05	10.16	12.82	15.99	17.71	21.34	24.19	28.17	32.96	36.49	40.14	46.16
25	2.64	4.14	6.28	8.74	12.05	15.31	18.94	23.85	26.88	30.38	34.45	38.34
26	6.96	9.92	11.24	11.40	13.24	17.74	19.83	23.24	26.04	27.95	33.89	39.69
27.1-27.3	16.59	20.99	28.23	25.05	32.39	35.76	39.16	43.21	45.75	47.15	50.29	60.75
27.4+27.5	46.56	55.67	72.34	83.02	91.94	103.88	129.03	144.87	169.02	204.70	232.67	274.78
28	3.35	4.32	5.41	7.52	10.71	13.26	16.21	20.29	22.86	26.60	29.35	34.03
29	8.89	12.03	12.52	19.45	21.80	25.32	29.19	39.95	41.67	50.79	54.95	68.72
31	47.65	47.46	57.11	58.93	67.79	69.00	73.32	77.85	81.86	91.39	89.77	102.87
32	150.08	217.92	223.35	219.34	239.84	237.13	225.63	252.57	350.76	399.66	342.11	382.49
33	22.67	31.90	28.10	35.54	33.68	42.68	46.39	52.82	58.05	66.19	68.25	92.14
34	40.92	46.90	57.12	63.34	73.92	87.56	102.19	116.57	130.90	140.69	155.86	178.34
35	10.78	12.14	14.45	16.36	18.53	20.66	23.59	21.92	24.46	28.53	33.22	37.42
36+37	10.90	14.86	18.83	21.79	23.36	25.93	29.79	32.71	34.41	39.27	45.61	53.42
E	1.32	2.75	4.24	4.98	7.81	9.60	12.36	15.55	18.11	22.64	24.58	30.24

7 Appendix

7.1 Estimation

7.1.1 Estimation ARIMA parameters

In section 7.3 we concluded that the original series X_t shows clear signs of non-stationarity. The first difference of the original series $Y_t := (1-B)^d X_t$, $d = 1$ on the other hand shows signs of stationarity and we will therefore estimate ARIMA$(p, 1, q)$ models to our time series which by definition means that we adjust ARMA(p, q) models to the first difference of the original series. All ARMA(p, q) models where p and q are allowed to take on the maximum value of 15 are considered. The parameters are estimated according to the description found in Brockwell and Davis (2002) by maximum likelihood. For our time series $\mathbf{X} = (X_1, ..., X_n)$ we assume a multivariate normal distribution with the density function

$$f_{\mathbf{X}}(\mathbf{X}) = (\frac{1}{2\pi})^{n/2} \frac{1}{\sqrt{\det(\Gamma)}} \exp\{(\mathbf{X}-\mu)'\Gamma^{-1}(\mathbf{X}-\mu)\} \tag{18}$$

where Γ is the autocovariance matrix

$$\Gamma = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdot\;\cdot\;\cdot & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdot\;\cdot\;\cdot & \gamma(n-2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \gamma(n-1) & \gamma(n-2) & \cdot\;\cdot\;\cdot & \gamma(0) \end{pmatrix}$$

The autocovariance matrix can be expressed in terms of the (unknown) model parameters and its entries are determined by the order of the ARMA model. The parameters are estimated in order to maximize the likelihood function. Even if the assumption of the multivariate normality of $\mathbf{X}$ does not quite seem to hold, it is still appropriate to estimate the parameters by maximizing 18 (Brockwell and Davis, 2002).

7.1.2 Estimation SETAR parameters

For every threshold γ we get two AR models. The parameters of the two models are estimated by linear regression. We have $\mathrm{y}_j = (y_{j_1}, ..., y_{j_{n_j}})'$ where y_j are the observations of the time series that belong to regime j. Further, we have the matrix $\mathbf{X}_j' = (\mathbf{x}_{j_1}, ..., \mathbf{x}_{j_{n_j}})'$ and $\mathbf{x}_{j_i} = (y_{j_i-1}, y_{j_i-2}, ..., y_{j_i-p_j})'$, p_j denotes the order of the AR model being estimated in regime j and n_j denotes the total number of observations in regime j. In our case we only study SETAR models with two regimes and thus j=1, 2.

In the final step we estimate the two linear regression models.

$$\mathbf{y}_j = \mathbf{X}_j \mathbf{\Phi}_j + \epsilon_j \ , \quad j = 1, 2$$

7.1.3 Estimation LSTAR parameters

The parameters of the LSTAR model are estimated by nonlinear least squares as described in Dijk, Teräsvirta, and Franses (2002). We have a vector of parameters $\theta = (\Phi_1, \Phi_2, \gamma, c)$ where $\Phi_1 = (\phi_{1,0}, ..., \phi_{1,p_1})$ and $\Phi_2 = (\phi_{2,0}, ..., \phi_{2,p_2})$.

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \sum_{i=1}^{n} (y_i - F(x_i; \theta))^2$$

$F(x_i; \theta)$ is given by the definition of the model as in 6 (excluding the white noise term):

$$F(x_i; \theta) = (\phi_{1,0} + \phi_{1,1} X_{t-1} + ... + \phi_{1,p_1} X_{t-p_1}) + (\phi_{2,0} + \phi_{2,1} X_{t-1} + ... + \phi_{2,p_2} X_{t-p_2}) G(X_{t-d}; \gamma, c)$$

7.2 Description of industrial production index (IPI)

Industrial production index (IPI) is produced by Statistics Sweden (SCB). The index is published monthly and gives information about the total swedish industrial production as well as industrial production in different industrial branches. There are several different kinds of sources for the index. About 75 percent is based on data over sales which is gathered from the survey *Konjunkturstatistik för industrin*. The data in this survey is collected from 2200 companies that are sampled from the business register at SCB. The method used to collect this data is stratified random sampling. About five percent is based on information about worked hours which is gathered by a survey at SCB that collects data on the wages in the private sector. Finally, about 20 percent is based on information about production volume gathered from selected companies and/or branch organizations (SCB, 2009).

7.2.1 List of indicies included in the study

NACE 2002	Name
10-12	Mining and quarrying of energy producing materials
13.1	Iron ore mines
13.2	Other metal ore mines
14	Other mines and quarries
15+16	Food product, beverage and tobacco industry
17-19	Textile industry, industry for wearing and Tanneries
20.1	Saw-mills and planing-mills; wood impregnation plants
20.2-20.5	Industry for wood, products of wood; except furniture
21.11	Industry for pulp
21.12	Industry for paper and paperboard
21.2	Industry for articles of paper and paperboard
22	Publishers and printers; other industry for recorded media
23	Industry for coke, refined petroleum products and nuclear fuel
24.4	Industry for pharmaceuticals, medicinal chemicals and botanical products
24 excl. 24.4	Manufacture of chemicals and chemical products except pharmaceuticals, medicinal chemicals
25	Industry for rubber and plastic products
26	Industry for other non-metallic mineral products
27.1-27.3	Manufacture of basic iron and steel and of ferro-alloys, tubes and other first processing of iron
27.4+27.5	Manufacture of basic precious and non-ferrous metals and casting of metals
28	Industry for fabricated metal products, except machinery and equipment
29	Industry for machinery and equipment n.e.c.
31	Industry for electrical machinery and apparatus n.e.c.
32	Industry for radio, television and communication equipment and apparatus
33	Industry for medical, precision and optical instruments, watches and clocks
34	Industry for motor vehicles, trailers and semi-trailers
35	Industry for other transport equipment
36+37	Other manufacturing industry n.e.c.
E	Electricity, gas and water works

7.2.2 Plots of IPI

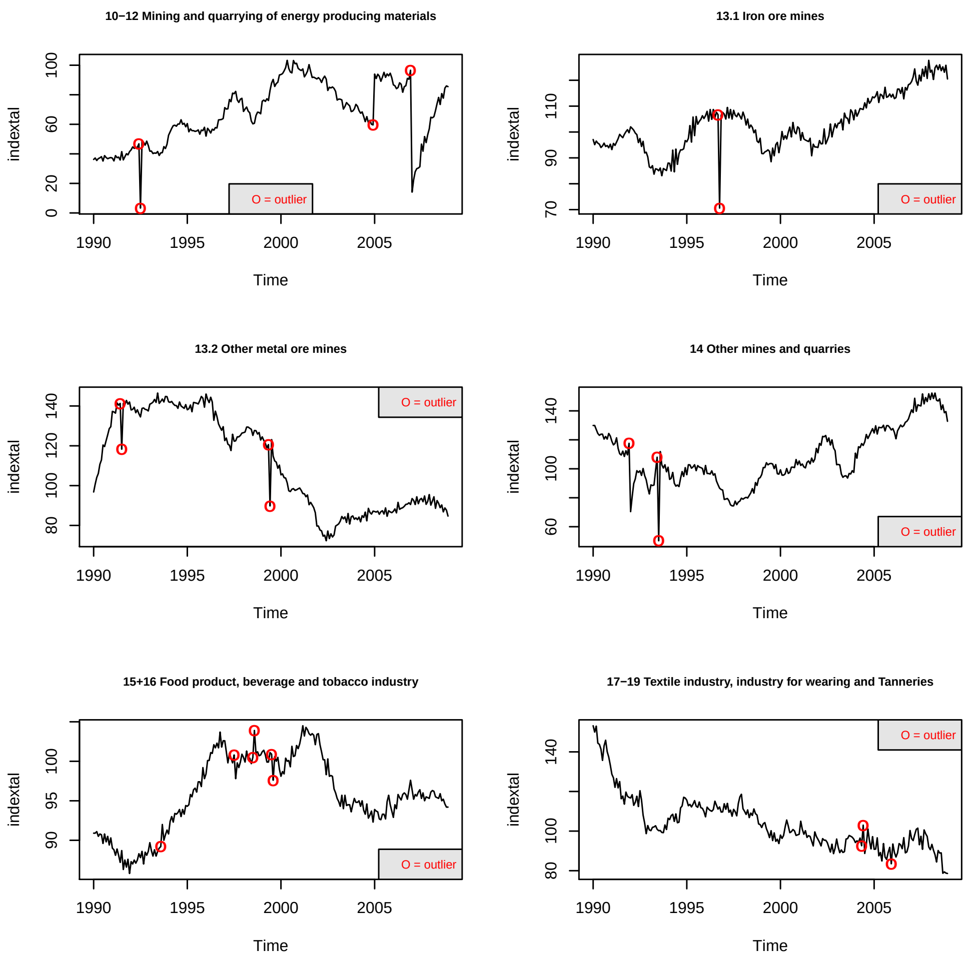


Figure 10:

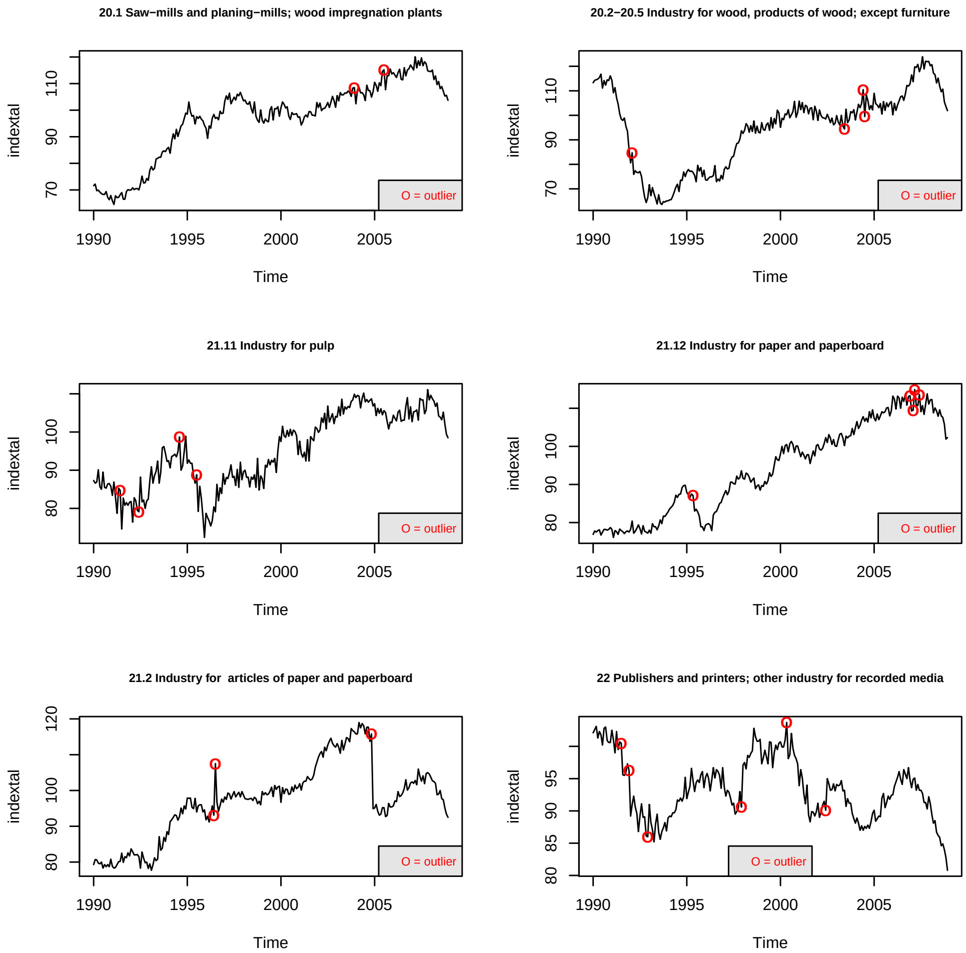

20.1 Saw−mills and planing−mills; wood impregnation plants
20.2−20.5 Industry for wood, products of wood; except furniture
21.11 Industry for pulp
21.12 Industry for paper and paperboard
21.2 Industry for articles of paper and paperboard
22 Publishers and printers; other industry for recorded media
indextal
Time
O = outlier
1990
1995
2000
2005

Figure 11:

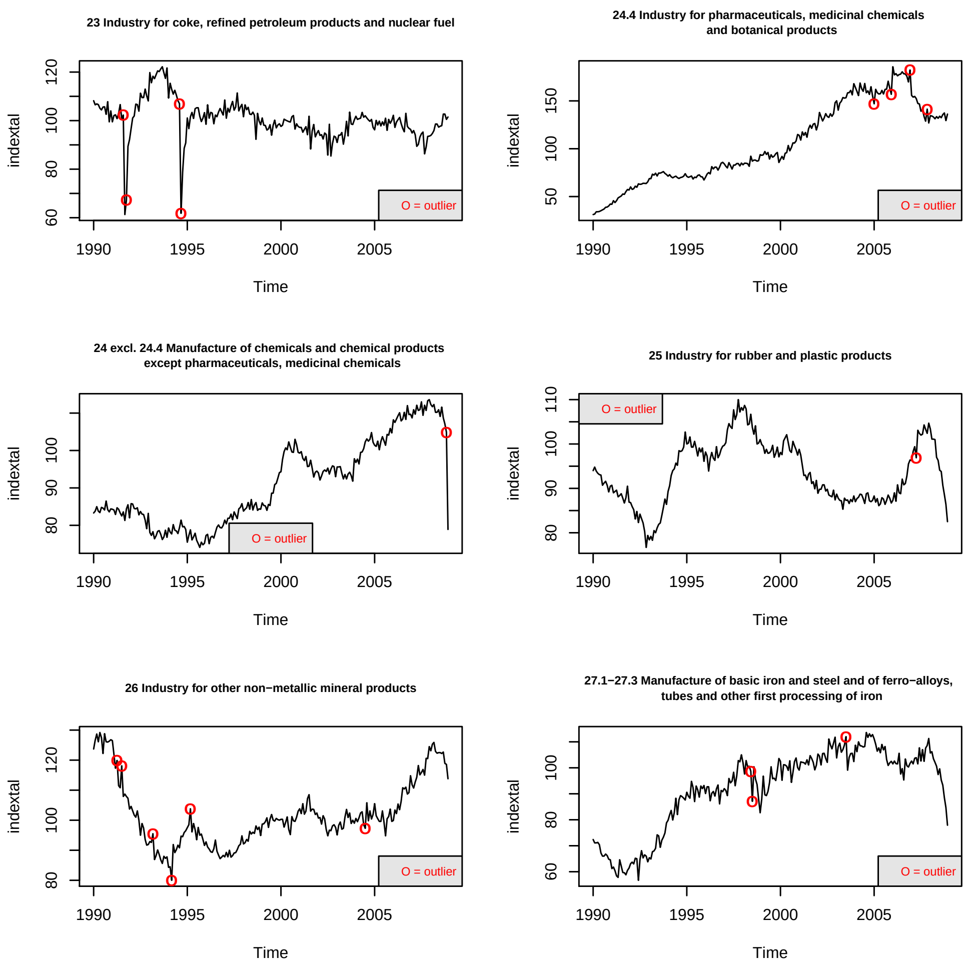

23 Industry for coke, refined petroleum products and nuclear fuel
24.4 Industry for pharmaceuticals, medicinal chemicals
and botanical products
24 excl. 24.4 Manufacture of chemicals and chemical products
except pharmaceuticals, medicinal chemicals
25 Industry for rubber and plastic products
26 Industry for other non−metallic mineral products
27.1−27.3 Manufacture of basic iron and steel and of ferro−alloys,
tubes and other first processing of iron
indextal
Time
O = outlier
1990
1995
2000
2005

Figure 12:

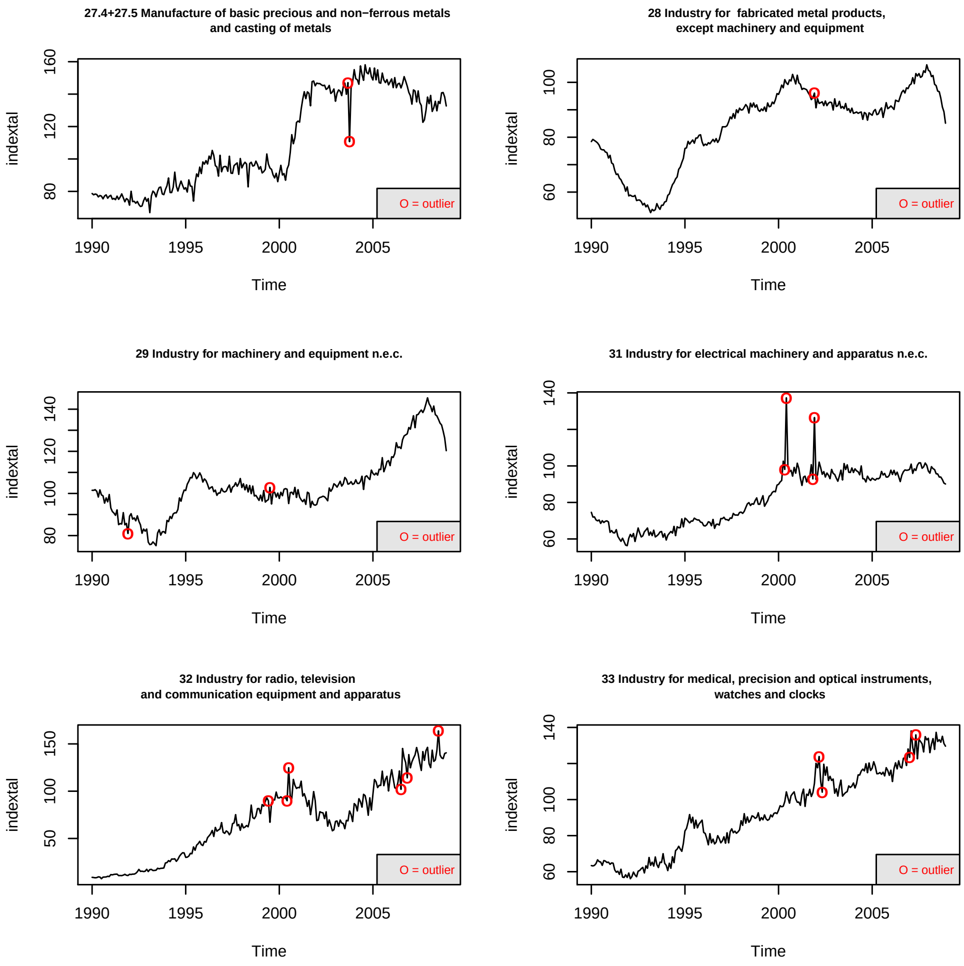

27.4+27.5 Manufacture of basic precious and non−ferrous metals
and casting of metals
28 Industry for fabricated metal products,
except machinery and equipment
29 Industry for machinery and equipment n.e.c.
31 Industry for electrical machinery and apparatus n.e.c.
32 Industry for radio, television
and communication equipment and apparatus
33 Industry for medical, precision and optical instruments,
watches and clocks
indextal
Time
O = outlier
1990
1995
2000
2005

Figure 13:

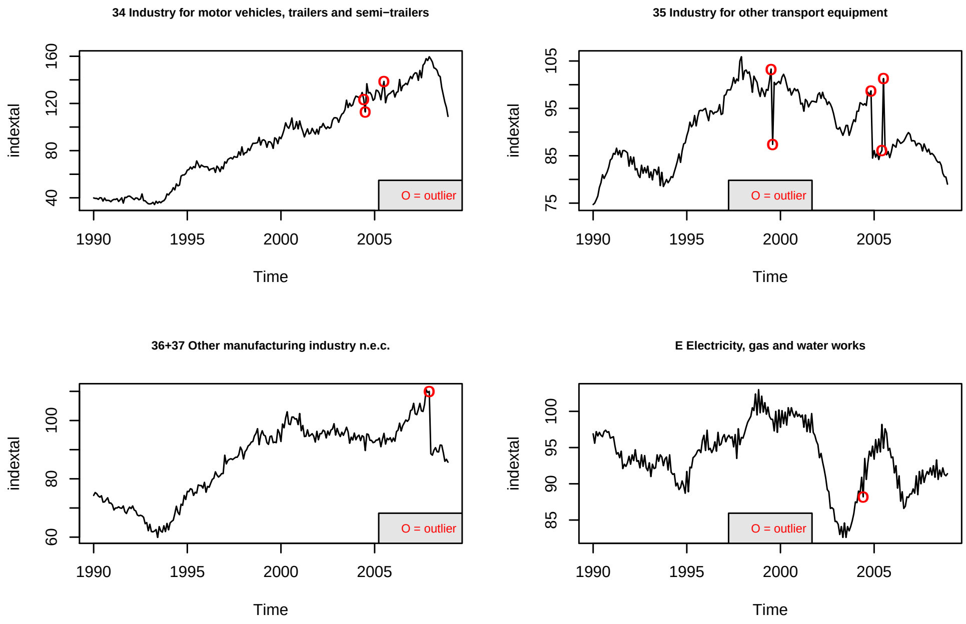

34 Industry for motor vehicles, trailers and semi−trailers
indextal
40
80
120
160
O = outlier
1990
1995
2000
2005
Time
35 Industry for other transport equipment
indextal
75
85
95
105
O = outlier
1990
1995
2000
2005
Time
36+37 Other manufacturing industry n.e.c.
indextal
60
80
100
O = outlier
1990
1995
2000
2005
Time
E Electricity, gas and water works
indextal
85
90
95
100
O = outlier
1990
1995
2000
2005
Time

Figure 14:

7.3 Plots of ACF of indices before and after taking the first difference

As we concluded in section 2.1 ARIMA models can be used to model non-stationary time series. We also want to apply LSTAR and SETAR models to series that are stationary. A common way of determining if a time series meets the conditions of stationarity is to study the autocorrelation function. We have already defined the autocovariance function (ACVF) (2). The autocorrelation function (ACF), which we denote $\rho(\cdot)$, is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \tag{19}$$

(Brockwell and Davis, 2002). We use data to estimate the ACF. If $x_1, ..., x_n$ are observations from a time series we estimate the mean of the series by the average

$$\bar{x} = n^{-1} \sum_{t=1}^{n} x_t.$$

and we have the following estimation of the AVCF

$$\hat{\gamma}(h) := n^{-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n.$$

and thus the estimation of the ACF becomes

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad -n < h < n.$$

A common sign of non-stationarity is that the ACF is decreasing slowly as a function of h (Brockwell and Davis, 2002). The transformation of the series by taking the first difference if often sufficient to achieve stationarity of the transformed series $Y_t = X_t - X_{t-1}$. In our case, the ACF of X_t decreases slowly while the ACF of the transformed series Y_t does not show any sign of this behavior so typical of non-stationarity. Thus it seems reasonable to draw the conclusion that Y_t is stationary, but not X_t. In the figures below the ACF of X_t and Y_t are plotted for 28 time series in the study.

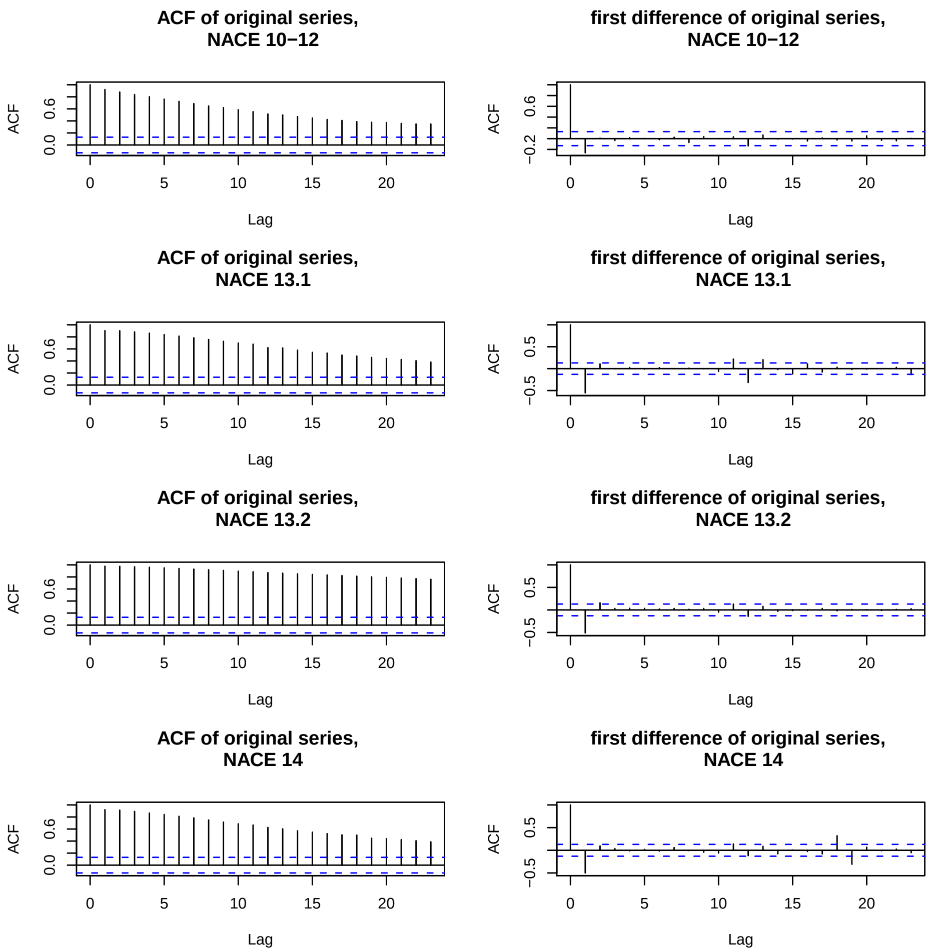

ACF of original series,
NACE 10−12
first difference of original series,
NACE 10−12
ACF of original series,
NACE 13.1
first difference of original series,
NACE 13.1
ACF of original series,
NACE 13.2
first difference of original series,
NACE 13.2
ACF of original series,
NACE 14
first difference of original series,
NACE 14
ACF
Lag

Figure 15:

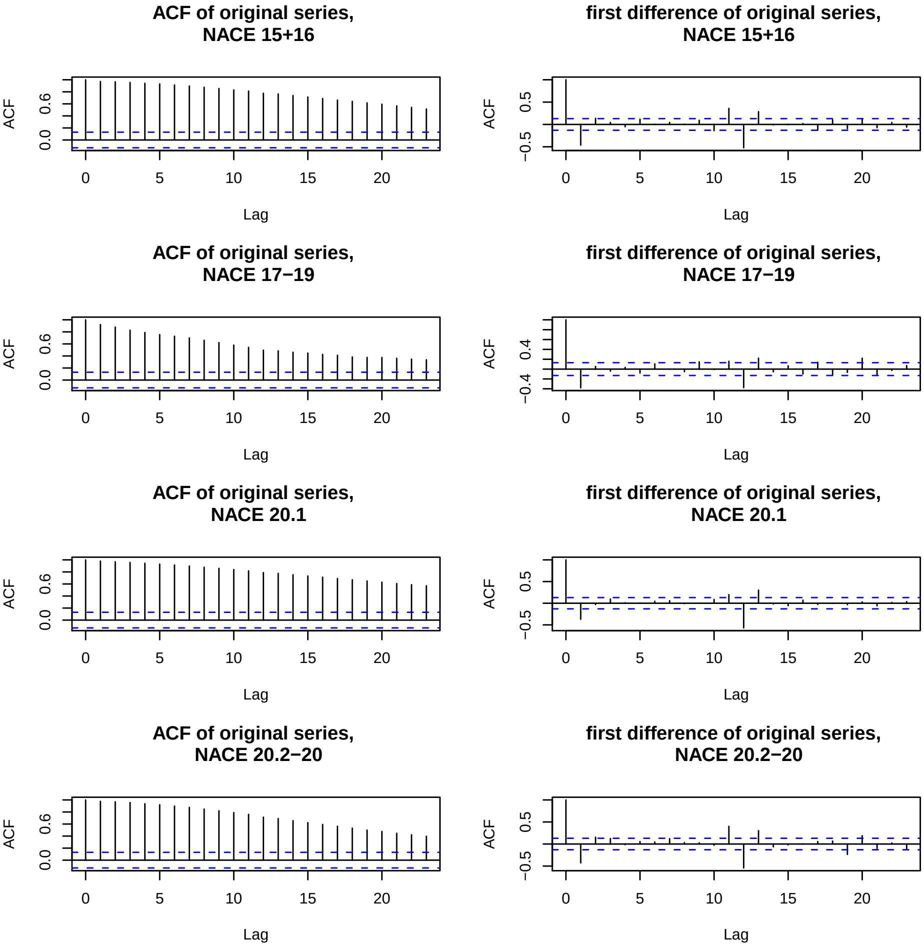

ACF of original series,
NACE 15+16
first difference of original series,
NACE 15+16
ACF of original series,
NACE 17−19
first difference of original series,
NACE 17−19
ACF of original series,
NACE 20.1
first difference of original series,
NACE 20.1
ACF of original series,
NACE 20.2−20
first difference of original series,
NACE 20.2−20
ACF
Lag

Figure 16:

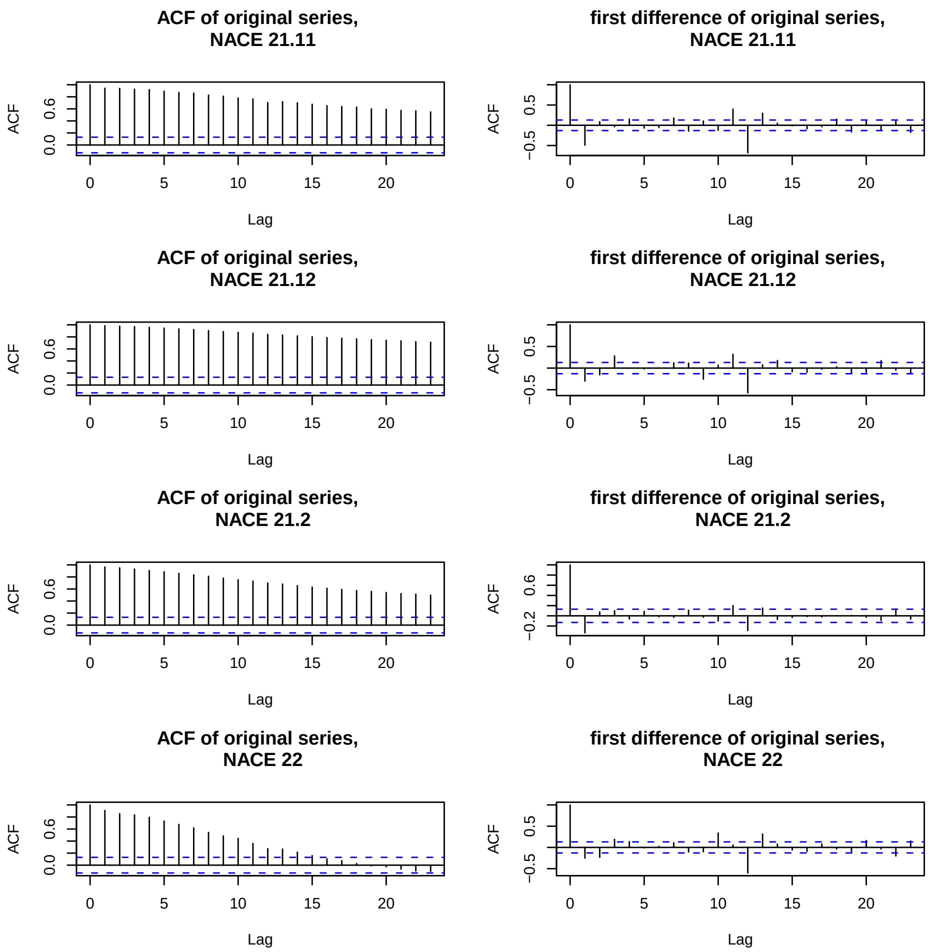

ACF of original series,
NACE 21.11
first difference of original series,
NACE 21.11
ACF of original series,
NACE 21.12
first difference of original series,
NACE 21.12
ACF of original series,
NACE 21.2
first difference of original series,
NACE 21.2
ACF of original series,
NACE 22
first difference of original series,
NACE 22
ACF
Lag

Figure 17:

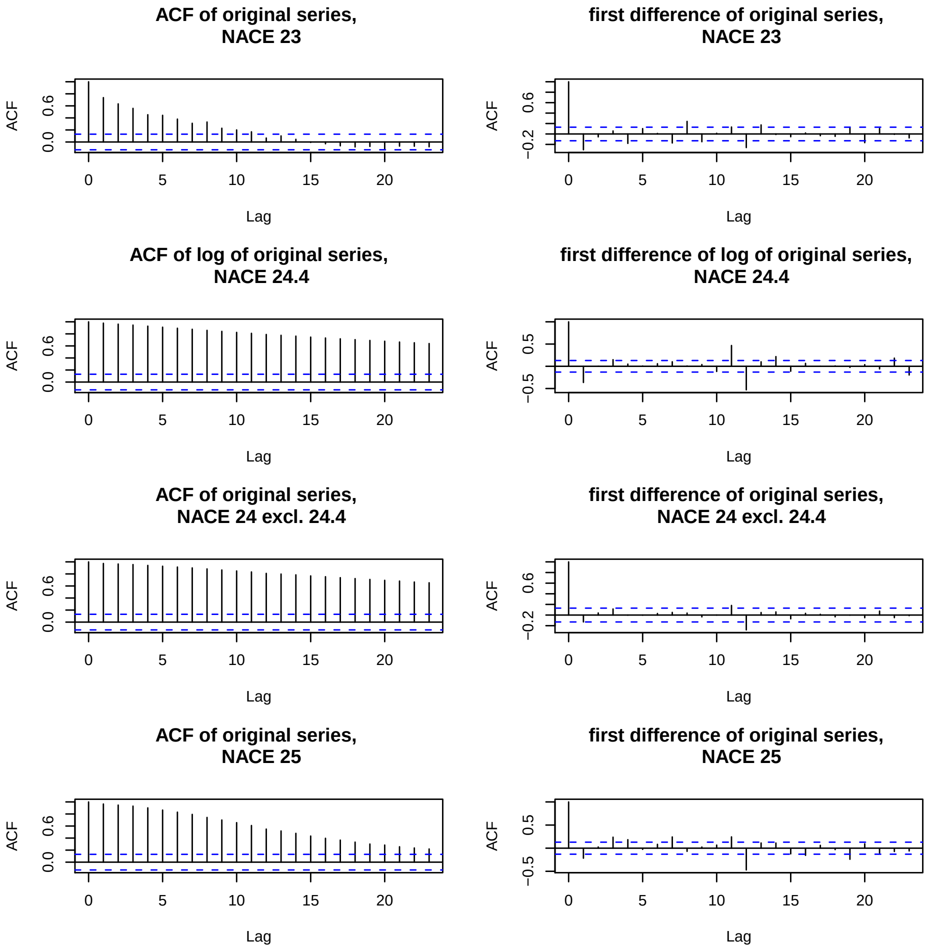
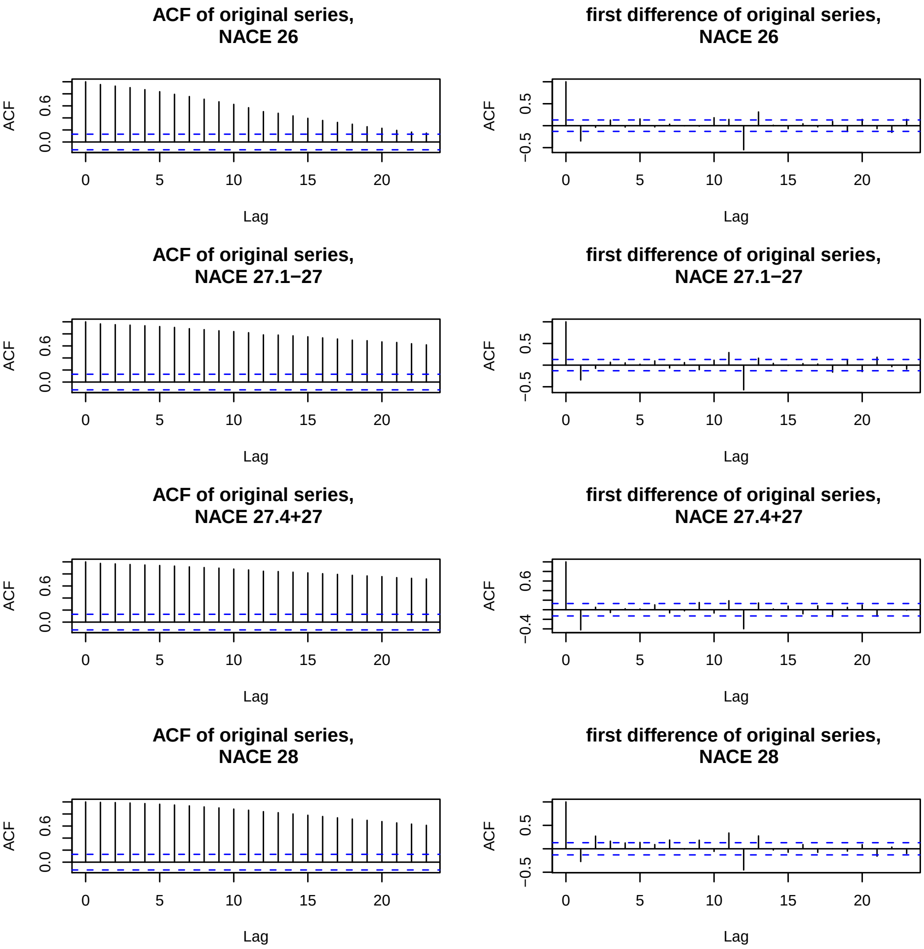

ACF of original series, NACE 23
first difference of original series, NACE 23
ACF of log of original series, NACE 24.4
first difference of log of original series, NACE 24.4
ACF of original series, NACE 24 excl. 24.4
first difference of original series, NACE 24 excl. 24.4
ACF of original series, NACE 25
first difference of original series, NACE 25
ACF
Lag

Figure 18:



ACF of original series,
NACE 26
first difference of original series,
NACE 26
ACF of original series,
NACE 27.1−27
first difference of original series,
NACE 27.1−27
ACF of original series,
NACE 27.4+27
first difference of original series,
NACE 27.4+27
ACF of original series,
NACE 28
first difference of original series,
NACE 28
ACF
Lag

Figure 19:

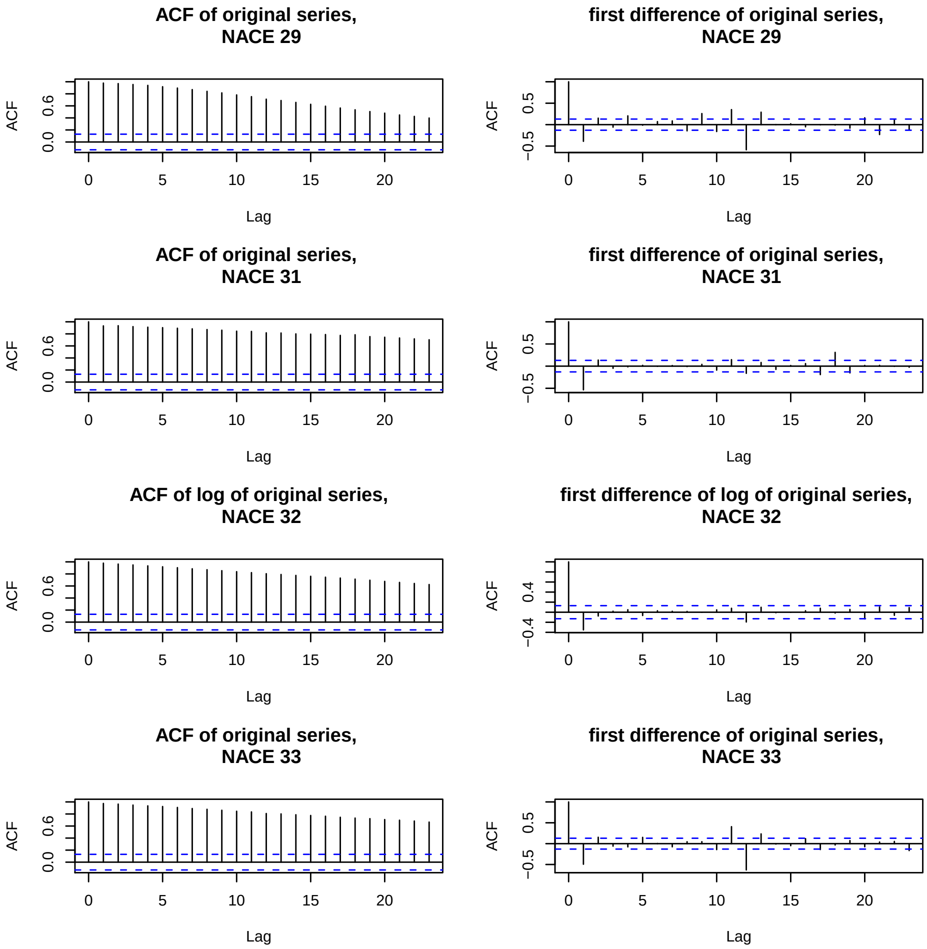

ACF of original series,
NACE 29
first difference of original series,
NACE 29
ACF of original series,
NACE 31
first difference of original series,
NACE 31
ACF of log of original series,
NACE 32
first difference of log of original series,
NACE 32
ACF of original series,
NACE 33
first difference of original series,
NACE 33
ACF
Lag

Figure 20:

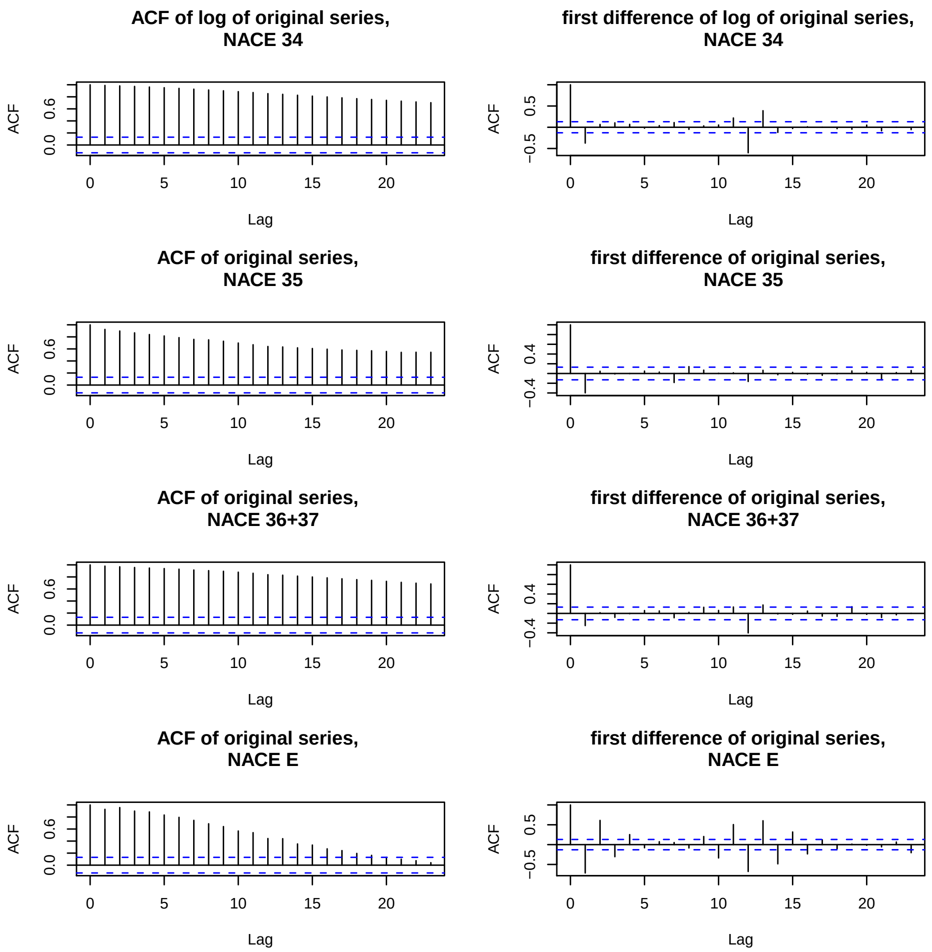

ACF of log of original series,
NACE 34
first difference of log of original series,
NACE 34
ACF of original series,
NACE 35
first difference of original series,
NACE 35
ACF of original series,
NACE 36+37
first difference of original series,
NACE 36+37
ACF of original series,
NACE E
first difference of original series,
NACE E
ACF
Lag

Figure 21:

References

Fabio Di Narzo Antonio, Jose' Luis Aznarte, and Matthieu Stigler. *tsDyn: Time series analysis based on dynamical systems theory*, 2008. R package version 3.1-90.

Peter J. Brockwell and Richard A. Davis. *Introduction to time series and forecasting*. Springer, New York, 2. ed. edition, 2002. ISBN 0-387-95351-5 (alk. paper).

Dick van Dijk, Timo Teräsvirta, and Philip Hans Franses. Smooth transition autoregressive models - a survey of recent developments. *Econometric Reviews*, 21(1):1–47, 2002. URL http://econpapers.repec.org/RePEc:taf:emetrv:v:21:y:2002:i:1:p:1-47.

Clive William John Granger and Timo Teräsvirta. *Modelling nonlinear economic relationships*. Oxford Univ. Press, Oxford, 1993. ISBN 0-19-877319-6 (inb.).

Rob J Hyndman. *forecast: Forecasting functions for time series*, 2009. URL http://CRAN.R-project.org/package=forecast. R package version 1.24.

Helmut Luetkepohl and Fang Xu. The role of the log transformation in forecasting economic variables. CESifo Working Paper Series CESifo Working Paper No. 2591, CESifo Group Munich, 2009. URL http://econpapers.repec.org/RePEc:ces:ceswps:_2591.

Alan L. Montgomery, Victor Zarnowitz, Ruey S. Tsay, and George C. Tiao. Forecasting the u.s. unemployment rate. *Journal of the American Statistical Association*, 93(442):478–493, 1998. ISSN 01621459. URL http://www.jstor.org/stable/2670094.

Daniel Peña Sánchez de Rivera, George C. Tiao, and Ruey S. Tsay. *A course in time series analysis*. Wiley, New York, 2001. ISBN 0-471-36164-X ;.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. URL http://www.R-project.org. ISBN 3-900051-07-0.

SCB. Industriproduktionsindex. http://www.scb.se/statistik/NV/NV0402/_dokument/NV0402_BS_2008.doc, mars 2009.

SCB. Seasonal adjustment. http://www.scb.se/Pages/Standard____35726.aspx13/12010, januari 2010.

James H. Stock and Mark W. Watson. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. NBER

Working Papers 6607, National Bureau of Economic Research, Inc, June 1998. URL http://ideas.repec.org/p/nbr/nberwo/6607.html.

T. Terasvirta and H. M. Anderson. Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, 7:S119–S136, 1992. ISSN 08837252. URL http://www.jstor.org/stable/2284987.

Timo Teräsvirta, Dick van Dijk, and Marcelo C. Medeiros. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21(4):755 – 774, 2005. ISSN 0169-2070. doi: DOI:10.1016/j.ijforecast.2005.04.010. URL http://www.sciencedirect.com/science/article/B6V92-4GMJ8X2-1/2/c960be792af9af88f57deb2e3c9f4885. Nonlinearities, Business Cycles and Forecasting.