

Analysis and forecasting of the Swiss Rent Index Time Series, 1993 to 2018

Zully Faralli, Marc Spörri

May 2, 2018

1 Introduction

describe aims of your project and give an overview of the report. describe briefly the methods you have used to analyze the data.

copied from the proposal: The main goal of the project is to draw inference from the rent index time series and to predict their future values and a future trend. In order to do that we have first to transform the data in a stationary time series, either by fitting a polynomial trend, either by differencing. Afterwards we are going to fit a model to our data. The model will be used to describe and interpret our data as well as to predict future values.

2 Data Description

Describe the data and the context in which it has been gathered and explain. Give some general references about the data.

copied from the proposal: The data is provided by the Swiss Federal Statistical Office, implying 100 quarterly observations of the Swiss rent index beginning by the 2nd quarter of the year 1993 and ending by the first quarter of the year 2018. The rent index measures the inflation of the rented dwellings in Switzerland. It is the most significant partial index of the Swiss consumer's price index, representing a weighted share of 13 percent of this index. The data are collected quarterly, based upon a stratified random survey sampling of around 10'000 lessors. The time series' first observation (2nd quarter of the year 1993) will be the reference value which is set to a base index of 100 and represents the weighted average rent at this time in Switzerland (OFS, 2016: p. 20-23).

3 Data Analysis and Transformation into a Stationary Time Series

Describe your models carefully, give references, justify your choices. You can use the guidelines provided below.

1. Do you transform the data? If yes, give the transformation you used and explain this choice.
2. Does the series show a trend and a seasonality? If yes, describe how you model them. Check the plot of the residuals.
3. Comment the autocorrelation and the partial autocorrelation of the residual series obtained after removing the trend and seasonality.
4. Model the residuals with an appropriate model. If you consider different models, explain strengths and weaknesses of each model.
5. Comment the diagnostic plots of the model(s) you chose. you can plot the standardized residuals, the normal qqplot, the ACF and PACF of the residuals, and the p-value of the Ljung-Box-Tests (Ljung and Box, 1978) for different lags to assess stationarity, independence and normality of the residuals.
6. By means of your model, make some predictions and give confidence intervals for the prediction. you can illustrate it by showing some plots.

The first step in any analysis of a time series is to plot the data in order to identify discontinuities or a sudden change of level in the series (book p. 23). As we can see from figure 3.1 there seems to be a clear positive trend.

In addition it may be advisable as well to analyze the series by breaking it into homogeneous segments (Brockwell and Davis, 2002, p. 23) . Let's have a look at the segmented plots 1993 to 2009 and 2009 to 2018 in figure 3.2 and figure 3.3. We can see a slightly slower increase in the segment from 2009 to 2018 indicating that in the years from 1993 to 2009 the growth in rental prices was higher than in the years in the second's segment which begins from 2009. that can be explained by the big baisse in the early 90ies, starting from a lower initial point and having better conjunctural perspectives the increase was stronger, whilst from 2009 on the growth in rental prices slowed down, which can be very well explained by the US subprime crises beginning in the year 2008 followed by a long-taking global recession. However there are no sudden changes in level or huge outliers, therefore we can start to find a transformation into a stationary time series for our whole data.

To get a stationary time series, which means we want to generate a noise sequence whose covariances are not depending on time (book p. 35) and expectation means which are constant, we have to eliminate any trend and/or seasonal components in our data. A trend is obviously visible from figure 3.1, however we cannot be that sure if it doesn't exist a seasonal

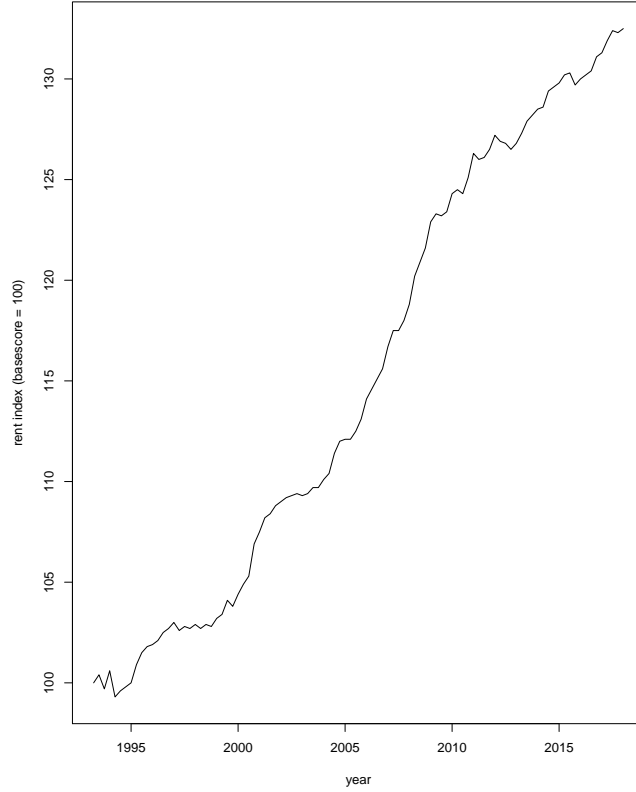


Figure 3.1: Swiss rent index, years 1993 to 2018

component either. Therefore in the next step we estimate a possible seasonal component by linear regression.

3.1 Check for Seasonality

Since the data are collected quarterly we use a linear regression model with $d=4$ dummy predictors, ergo a dummy for every season (LN1-2, p. 36). As we can see from figure 3.4 none of the estimated seasonal coefficients are significant. We can be sure now that there are no seasonal impacts on our data, hence what remains is to eliminate the trend. In the absence of a seasonal component our model becomes the following (Brook Davis p. 24):

$$X_t = m_t + Y_t, t = 1, \dots, n, \text{ where } EY_t = 0 \quad (1)$$

3.2 Trend Elimination by fitting polynomial models

As we have shown above in order to model our data we don't have to get rid of seasonal components, nevertheless we have to get rid of the obvious trend.

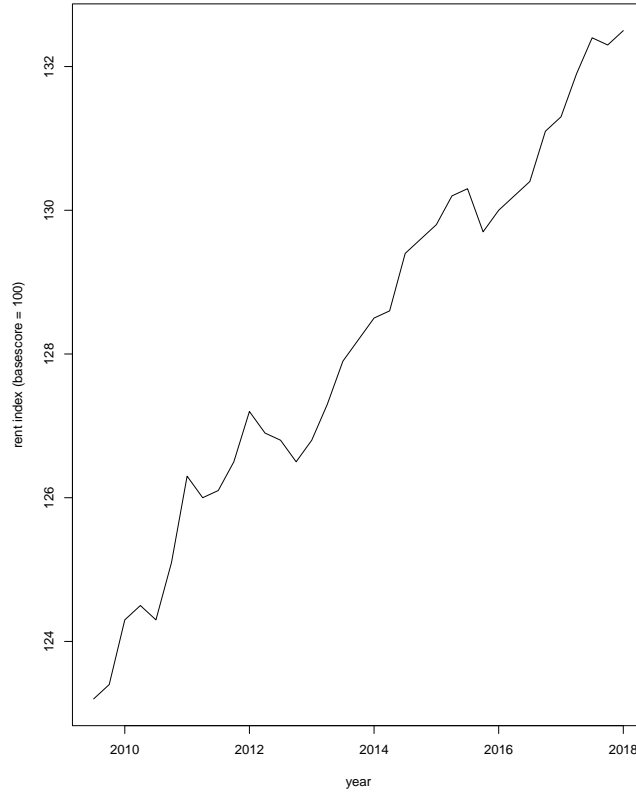


Figure 3.2: Swiss rent index, years 2009 to 2018

Whilst the time series 3.1 indicates a probable polynomial trend, especially a linear one, we are starting by fitting different polynomial trends by ordinary least squares estimation. We are fitting a linear trend, a quadratic trend, a cubic trend as well as a logarithmic trend. The latter helps to transform a potentially exponential increase in the rent index into a linear trend, even though at first sight the time series 3.1 looks more like a linear than an exponential trend, we are going to test for a logarithmic trend either.

The coefficients in all 4 models are highly significant and with an R^2 -Value of 98 to 99 percent (meaning that the models would be able to explain up to 98percent (99percent respectively) of the variance), the linear, the quadratic and the cubic model would explain our data extremely well which can be well seen on the cubic model's example in 3.5. The logarithmic model shows a bit less explanatory power with an R^2 -Value of 0.72, even though its coefficient is highly significant as well. However we have to be careful with the interpretation of the p-values since this regression models assume independence of the observations whereas our purpose of the summary statistics of

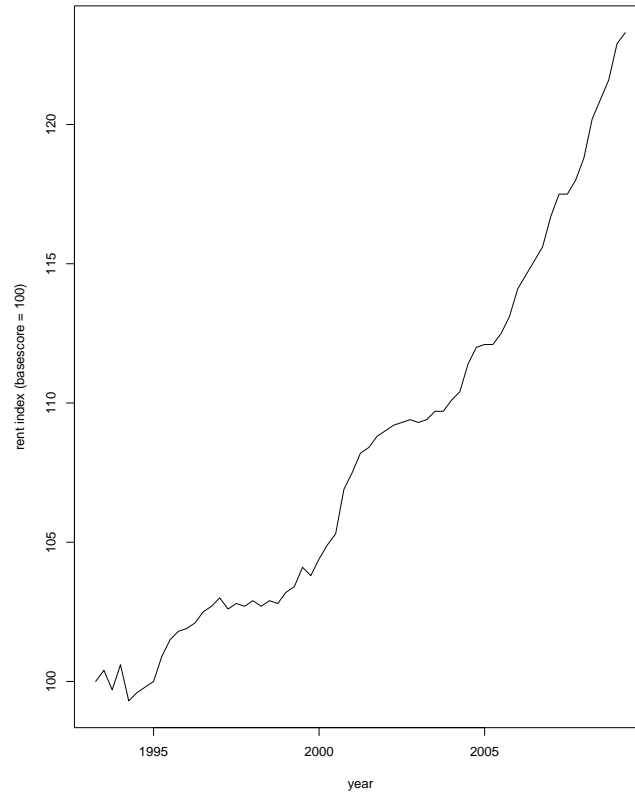


Figure 3.3: Swiss Rent index, years 1993 to 2009

```

Residuals:
    Min       1Q   Median       3Q      Max
-15.73 -11.35  -2.31   11.27   17.41

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      115.728     2.254   51.341  <2e-16 ***
season.model2nd quarter    -1.072     3.188   -0.336    0.737
season.model3rd quarter    -0.740     3.188   -0.232    0.817
season.model4rd quarter    -0.496     3.188   -0.156    0.877
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.27 on 96 degrees of freedom
Multiple R-squared:  0.001251, Adjusted R-squared:  -0.02996
F-statistic: 0.04009 on 3 and 96 DF, p-value: 0.9892

```

Figure 3.4: summary statistics of the fitted seasonal model

fitted linear models is a different one: to use regression in order to construct a stationary time series (Brooks Davis p. xy).

```

Residuals:
    Min       1Q   Median       3Q      Max
-1.8119 -0.7479 -0.1490  0.7887  1.9046

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.009e+02  3.869e-01  260.685 < 2e-16 ***
t            -1.104e-01  3.301e-02   -3.345  0.00117 **
t2           1.097e-02  7.575e-04   14.484 < 2e-16 ***
t3          -6.767e-05  4.931e-06  -13.723 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9313 on 96 degrees of freedom
Multiple R-squared:  0.9932,    Adjusted R-squared:  0.993
F-statistic: 4661 on 3 and 96 DF,  p-value: < 2.2e-16

```

Figure 3.5: summary statistics of the fitted cubic model

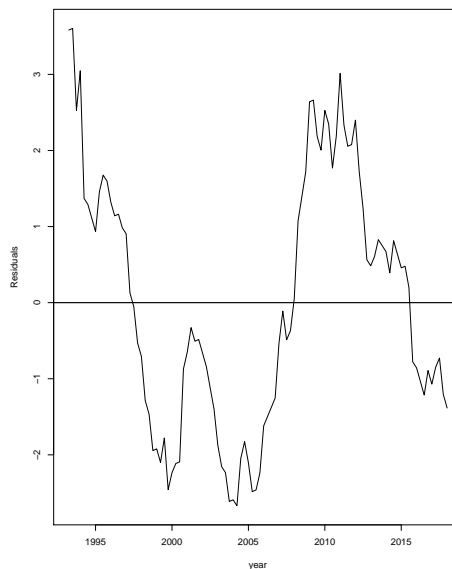


Figure 3.6: Residuals of a fitted linear model

3.2.1 Diagnostics of the fitted polynomial trends

Unfortunately the plots for polynomial trend elimination don't show any stationarity, as we can derive from the four figure 3.6, figure 3.7, figure 3.8 and figure 3.9. The four plots show large covariances and their residuals are evidently depending on time t .

Furthermore from the ACF-plot in figure 3.10, exemplifying the cubic model, we can see that there are a lot of significant lags outside the 95-percent-confidence bounds of $\pm 1.96\sqrt{rtn}$. These bars don't die out quickly. So the Rho_h can be a useful indicator of nonstationarity (book p.21). Hence, we can conclude our covariances depend on time and therefore the series is non-stationary. By the way, similar time-depending patterns

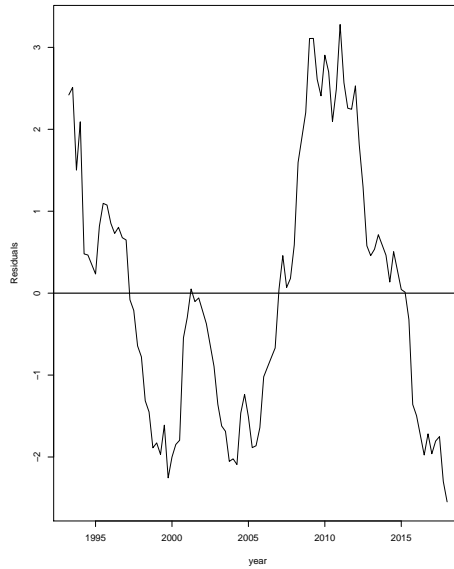


Figure 3.7: Residuals of a fitted quadratic model

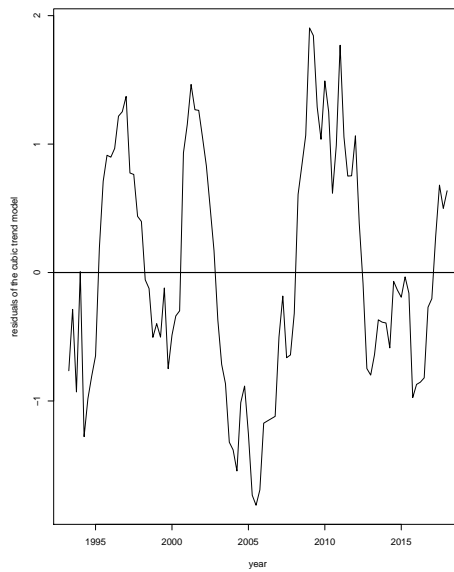


Figure 3.8: Residuals of a fitted cubic model

hold true for the autocovariance-functions of our other polynomial trend models: the linear, the quadratic and the logarithmic model. In addition we get numerical evidence from three different tests checking

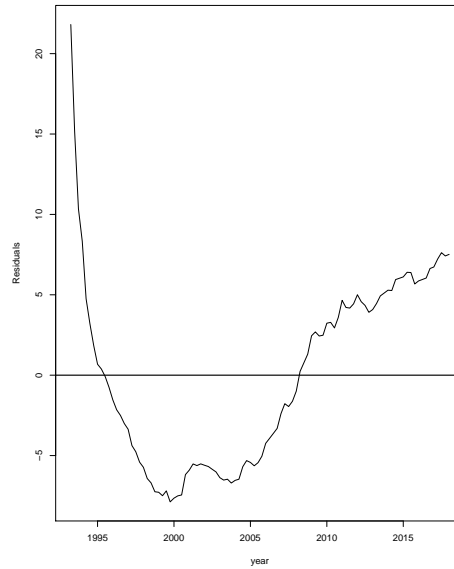


Figure 3.9: Residuals of a fitted logarithmic model

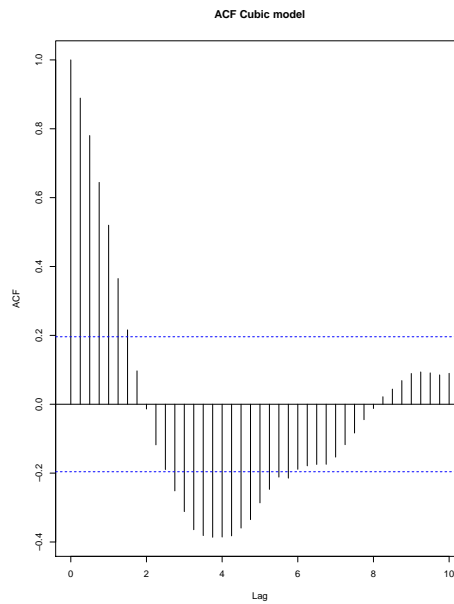


Figure 3.10: Sample autocorrelation function of the cubic model

for stationarity or independence of the observations, respectively. First the Ljung-Box test examines whether there is significant evidence for non-zero correlations at lags 1-40. Small p-values (i.e., less than 0.05) suggest that

the residuals are independent. The Ljung-Box Hypothesis of independence has to be rejected for all four polynomial models so far (TODO show the 4 p-values). The Augmented Dickie-Fuller test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test provide as well strong evidence for non-stationary for all 4 polynomial models (TODO show the 4 p-values). Even though we have to keep in mind the low power of Dickie-Fuller Test for our purposes given the fact this test assumes a AR-Process, about which we cannot be sure.

3.3 Trend Elimination by differencing

3.3.1 Diagnostics of the once-differenced series

Since it is not possible to obtain a stationary time series by fitting polynomial trends we have to use different methods. Hence, we try to eliminate the trend by differencing the data at different lags (1, 2, 3 and 4) in order to generate a noise sequence and therefore get a stationary series (brook davis p.35).

The following figure 3.11, figure 3.12, figure 3.13 and figure 3.14 show the differenced series derived from the quarterly rents for each lag = 1, 2, 3 and 4. For the first two models we were successfully able to generate

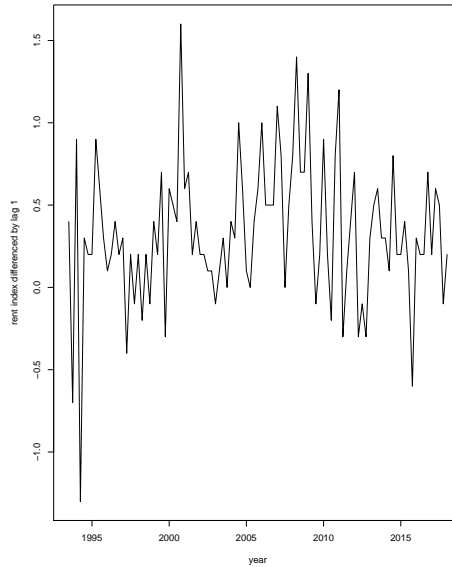


Figure 3.11: once-differenced rent index

stationary residuals which don't depend on time. However figure 3.11 shows mostly patterns of a white noise which is not what we intended since in case of a white-noise-process prediction quality is poor (Expectation of $x_n + 1$ would be 0 in case of a mean-centered series). This white-noise patterns

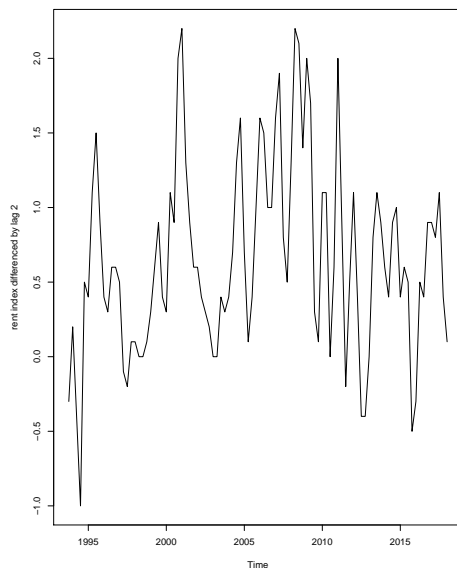


Figure 3.12: twice-differenced rent index

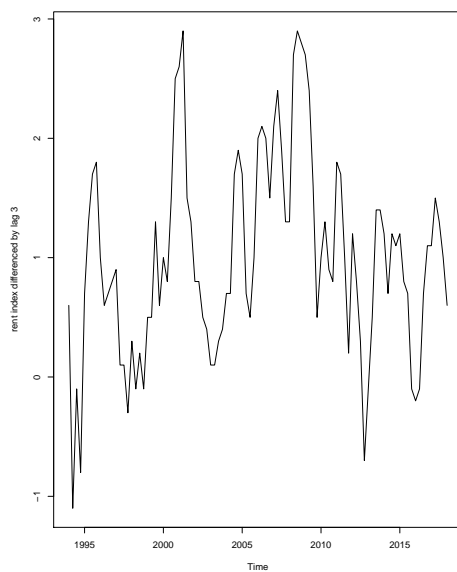


Figure 3.13: triple-differenced rent index

is as well indicated by the sample autocorrelation function of the once-differenced series where all lags up to 40 fall well within the bounds of $\pm 1.96\sqrt{98}$ (brooks davis p.39) (see figure 3.15) . To conclude, the

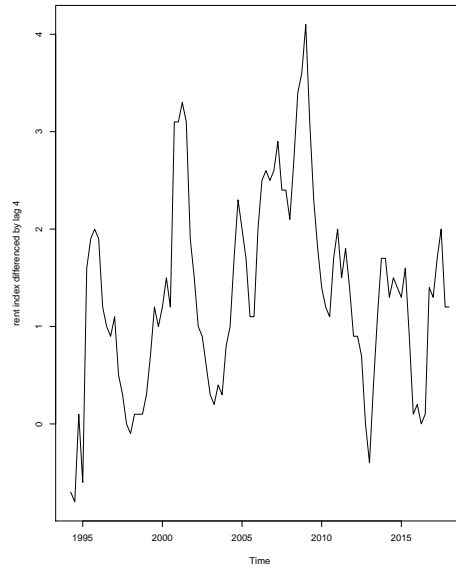


Figure 3.14: quadruple-differenced rent index

once-differenced series is too versatile and contains too much randomness to be able to get good predictions. The Ljung-Box-Test (Ljung-Box 1978, see

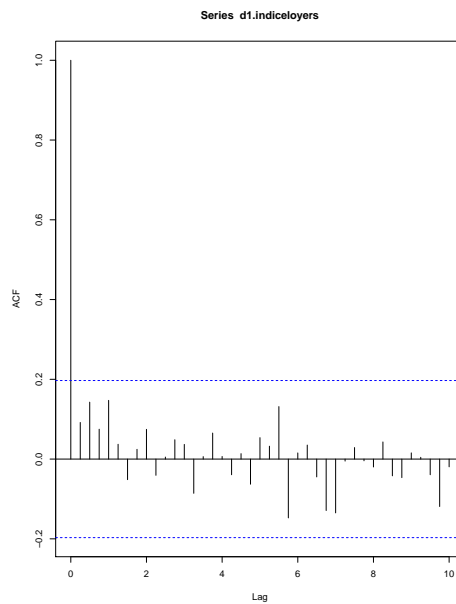


Figure 3.15: auto-correlation function of the once-differenced series

brooks davis p. 36) indicates a p-value of 0.35, meaning that we cannot reject

the iid-Hypothesis H_0 that there is independence between the observations. Therefore The Ljung-Box-Test emphasizes furthermore the independence of the values and the white-noise-patterns of the once-differenced series. Since white-noise would mean $E0$ and Variance σ^2 no more modelling would be necessary.

3.3.2 Diagnostics of the twice-differenced series

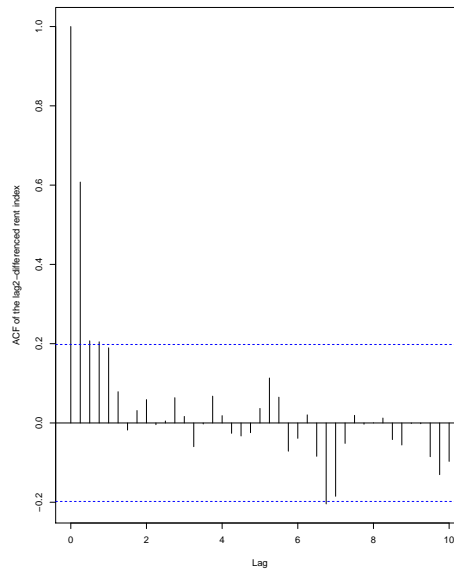


Figure 3.16: auto-correlation function of the twice-differenced series

In order to get better predictions and forecasting power than with pure white-noise-residuals seen in the once-differenced series, we now go for a twice-differenced model. Indeed our twice-differenced model doesn't show anymore iid-noisy patterns, indicating by the sample ACF (see figure 3.16 which shows at least a correlation at lag1 and some other lags are touching the $\pm 1.96\sqrt{n} = 100$ - 95percent-confidence-bounds. The ACF of the twice-differenced model shows only a few significant lags which die out quickly. This is a good sign, meaning that on one hand we have correlations which are needed to do prediction on the other hand the covariance is not too large to be time-dependant and therefore we can conclude our series is stationary. Cross-validation with a test data set considering only the last 35 of the 100 observations (see figure 3.19 above) confirms furthermore the non-stationary character of the twice-differenced-model `Box.test(d2.indiceloyers.test, lag=2, type="Ljung-Box")` p-value of 0.033 suggests the data are (time-)independent `adf.test(d2.indiceloyers.test,`

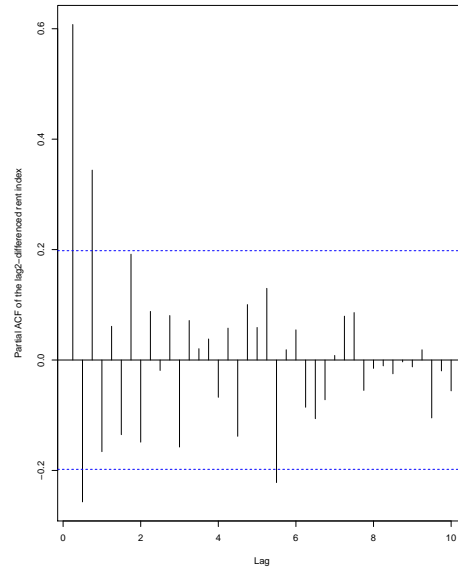


Figure 3.17: Partial auto-correlation function of the twice-differenced series

alternative = "stationary", k=2) H0 of non-stationarity cannot be rejected which is not a problem, that could be due to lack of observations `kpss.test(d2.indiceloyers.test)` with a p-value > 0.05 the test data seems to be stationary as well the data of the first segment's (1993 to 2009) twice-differenced series show independent observations as well (see figure 3.18), however we can see a slightly slower increase in the segment from 2009 to 2018 indicating that in the years from 1993 to 2009 the growth in rental prices was higher than in the years in the second's segment which begins from 2009. that can be explained by the big baisse in the early 90ies, starting from a lower initial point and better conjunctural perspectives the increase was stronger, whilst from 2009 on the growth in rental prices slowed down, which can be very well explained by the US subprime crises beginning in the year 2008 followed by a long-taking global recession.

3.3.3 Diagnostics of the triple-/quadruple-differenced series

Differencing by lag 3 and 4 is not appropriate neither since there is too much autocorrelation showed by the ACF a lot of bars outside the $\pm 1.96\sqrt{98}$ - bounds.

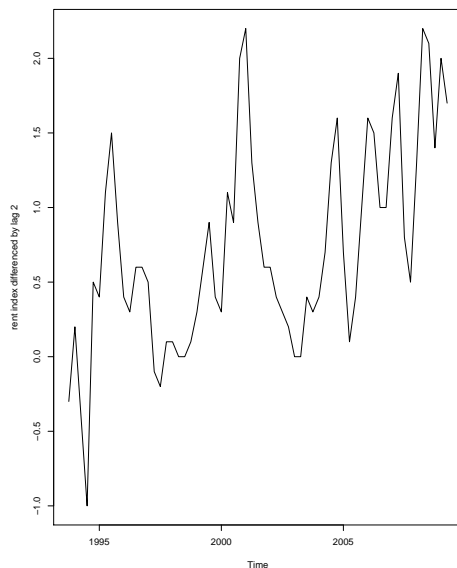


Figure 3.18: twice-differenced series from 1993 to 2009

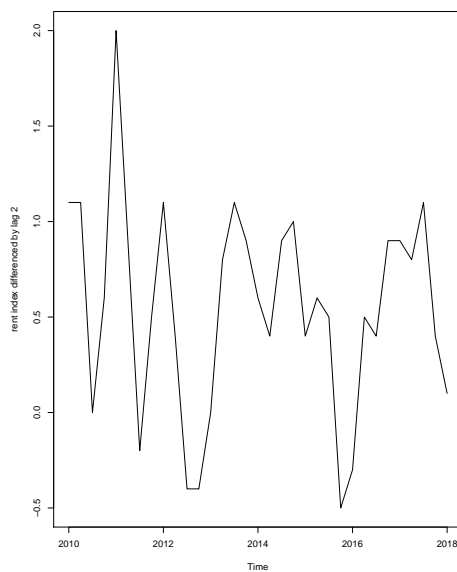


Figure 3.19: twice-differenced series from 2009 to 2018

4 Fitting and Testing the Model

from L6/7 p. 25. Therefore our model to rely on in order to transform our data into a stationary series should be the twice-differenced series,

since the once-differenced series is too white-noisy and the triple/quadruple-differenced series show too large covariances (time-dependence/nonstationarity). From the visualization of the sample ACF and sample PACF of the twice-differenced series (see figure 3.16 and figure 3.17, respectively) we can guess an AR(1) model as a suitable model for our data, since the ACF-plot shows only the first bar outside the borders (besides the following three touching them which we consider to give us too little evidence to justify an AR-model of order more than 1. Our sample PACF $\hat{\alpha}(h)$ is significantly different than zero for all h smaller or equal than 2 and becomes negligible for h larger than 2, indicating that a suitable model for the data could be $p=2$ so an AR(2) process. This is further emphasized by the plot where the sample PACF values beyond lag 2 are approximately i.i.d. with a $N(0, 1/n)$ distribution. Therefore in our supposed AR(2) process roughly 95 percent of the sample PACF values beyond lag 2 should be within the bound $\pm 1.96/\sqrt{n}$ which is in fact true according to the plot (no single bar after $p=2$ is outside the bound $\pm 1.96/\sqrt{n}$). sample PACF satisfies $|\hat{\alpha}(h)| > 1.96/\sqrt{n}$ for $h \leq p$ and $|\hat{\alpha}(h)| < 1.06/\sqrt{n}$ for $h > p$, which suggest an AR($p=2$) model for our data. In other words: The fact that all of the PACF values beyond lag 3 fall within the bounds suggests the possible suitability of an AR($p=3$) model for the mean corrected data set $X_t = S_t - 46.93$. One simple way to estimate the parameters ϕ_1 ϕ_2 ϕ_3 and σ^2 of such a model is to require that the ACVF of the model at lags 0, 1 and 2 should match the sample ACVF at those lags. Substituting the sample ACVF values $\hat{\gamma}_0$ $\hat{\gamma}_1$ $\hat{\gamma}_2$ for γ_0 γ_1 and γ_2 in the first three equations of (3.2.5) and (3.2.6) and solving for ϕ_1 ϕ_2 ϕ_3 and σ^2 gives the fitted model $X_t - 1.318x_{t-1} + 255x_{t-2} + s_{t-3} = Z_t$ with $hZ_t \sim WN$ (this method of model fitting is called Yule-Walker estimation) (Brock Davis p. 99)

Furthermore the choice of an AR(p) model is underlined by the fact that the ACF plot shows an exponentially decreasing pattern of $|\rho(h)|$ and PACF patterns with $p=3$ "spikes". Exponential decrease in AR and $\alpha(h)=0$ for $h > p=3$ in PACF are typical properties for an AR(p)-process.

In addition to the visualised evidence provided above further analytical evidence for an AR(1)-process is provided by the preliminary prediction by using the Yule-Walker-Algorithm for estimating roughly the AR-coefficients (ϕ). As we can see in table xy only the first estimated coefficient is a strong coefficient, all the others are close to zero and therefore not significant, emphasizing our assumption to go on with an AR(1)-Model. The program used from the ITSMR package will search through all the Yule-Walker AR(p) models, $p=0, 1, \dots, 27$, selecting the one with smallest AICC value. (explain AICC, corrected AIC) The minimum-AICC Yule-Walker AR model turns out to be the one defined by (eq 5.1.14) with $p=1$ and AICC value 74.541

In addition the estimated coefficient by Yule-Walker-Algorithm guarantees us that our model is causal. We can see that as well in table x which

shows indicating that the unit root is (slightly) outside the unit circle, with an value of 1.023714.

causality of a process (brock davis p. 85)

HannanRissanen-Algorithm shows strongly varying coefficients $\theta(q)$ for different fixed numbers of coefficients which further underlines our assumption that an AR(p)-model suits our data well instead of an MA(q) or ARMA(p,q) process.

we can see bringing a q into the game doesn't bring much, this is aswell confirmed by the AICC which is 41 and heigher than 34

BrockwellDavis p. 193 a root near 1 of the autoregressive polynomial suggests that the data should be differenced before fitting an AMA model, whereas a root near 1 of the moving-average polynomial indicates that the data were overdifferenced.

5 Prediction of future values

6 Discussion

Discuss your results, the strengths and the limitations of your own analysis.

7 Conclusion

Conclude the report. Sketch further analyses that you could carry out if you would have more time.

8 quotation examples

Hello World. Mister Resnick wrote a beautiful book about Harry, see (Resnick, 1992). If you want something more fancy, try Baddeley (2007). By looking at the plot we cannot definitely exclude that there might exist some seasonal patterns. To detect if a seasonality effect is really present in the series, we test it by modelling the seasonality and add it in the linear model.

A good numerical optimization algorithm is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. This algorithm requird an initial value which we obtained by the Hannam-Rinnanen-Alg. The BFGS-algorithm is a local optimization algorithm, meaning that it does not fid the global in minimum in general but instead a local minimum where the gradient is null. that's why we need a rough estimate by hannan-rinnanen to give te BFGS-algorithm a good inital vaue to get more secure that it will find the the global minimum and not only a loclal minimum instead.

References

- Baddeley, A. (2007). Spatial point processes and their applications. In Weil, W., editor, *Stochastic Geometry*, Lecture Notes in Mathematics, pages 1–75. Springer Berlin Heidelberg.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer New York, New York.
- Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65:297–303.
- Resnick, S. I. (1992). *Adventure in Stochastic Processes*. Birkhäuser, Boston.