

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

Zully Faralli, Marc Spörri

May 10, 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.

These changes are due to the global economic environment. Otherwise there is nothing unusual about the time plot and there appears to be no need to do any data adjustments. There is no evidence of changing variance, so we will not do a Box-Cox transformation.

In this section we want to produce a noise sequence with no apparent

deviations from stationarity, meaning we want to transform our data in such a way that covariances between the observations are not depending on time and that we get zero mean expectation and constant variances (Brockwell and Davis, 2002, pp. 14–23).

However, the objective is not to get a pure white noise sequence with no dependence among the residuals, since there would no further modelling to be done except to estimate their mean (which would be zero) and variance. The objective is too get a stationary series with some few significant dependence among the residuals anyhow, so we can look for a more complex stationary time series model for the noise that accounts for the dependence. Since dependence means that past observations of the noise sequence can assist in predicting future values this allows us to get a better prediction quality than with a pure white noise sequence (Brockwell and Davis, 2002, p. 35).

In order to get our noise sequence we have to eliminate any trend and/or seasonal components from our data. In the next few subsections we will fit several models to get a stationary series. Afterwards we check our fitted models first for stationary by tests like the Augmented Dickie-Fuller-Test for the null hypothesis of a unit root of a univariate time series (with the alternative hypothesis of stationarity by consequence) (Dickey and Fuller, 1979) and the Kwiatkowski-Phillips-Schmidt-Shin-Test (KPSS-Test) for the null hypothesis that the observations are trend stationary (Kwiatkowski et al., 1992). Furthermore we test the estimated noise sequences numerically by examining some simple tests for checking the hypothesis that the obtained residuals are values of independent and identically distributed random variables (as mentioned above: this is not our objective, otherwise our prediction work would be simply done by estimating their mean and variance). In a final step we check them visually by means of the autocorrelation-function plot as well as the partial auto-correlation-function in order to get sure that we are not modelling a white noise sequence on the one hand, and to get an idea of the orders of p and q , respectively for a possible ARMA(p, q)-model (Brockwell and Davis, 2002, pp. 83–110). Once we have found a good model to transform our data in a stationary series we can begin to estimate its parameters in section 4.

As we can see in figure 3.1 a trend is obviously visible, however we cannot be that sure if it doesn't exist a seasonal component either. Therefore in the next step we check first whether they are possible seasonal components, afterwards we will fit several models to get rid of the trend.

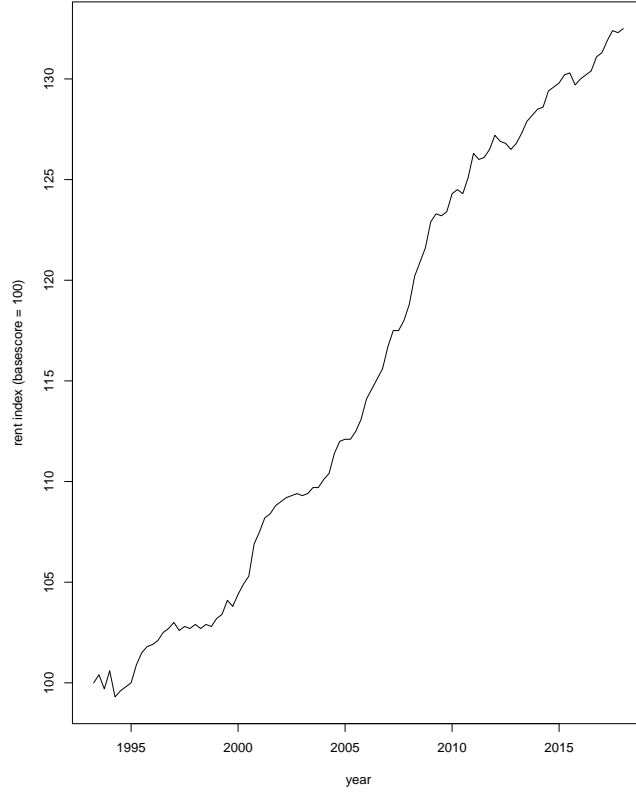


Figure 3.1: Swiss rent index, years 1993 to 2018

3.1 Check for Seasonality

Since the data are collected quarterly we use a linear regression model with $d=4$ dummy predictors to check for significant seasonal coefficients, meaning a dummy for every quarter (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 (Brockwell and Davis, 2002, 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 in the above subsection, we don't have to get rid of seasonal components, nevertheless we have to get rid of the obvious trend. Whilst the time series 3.1 indicates a probable polynomial trend, especially

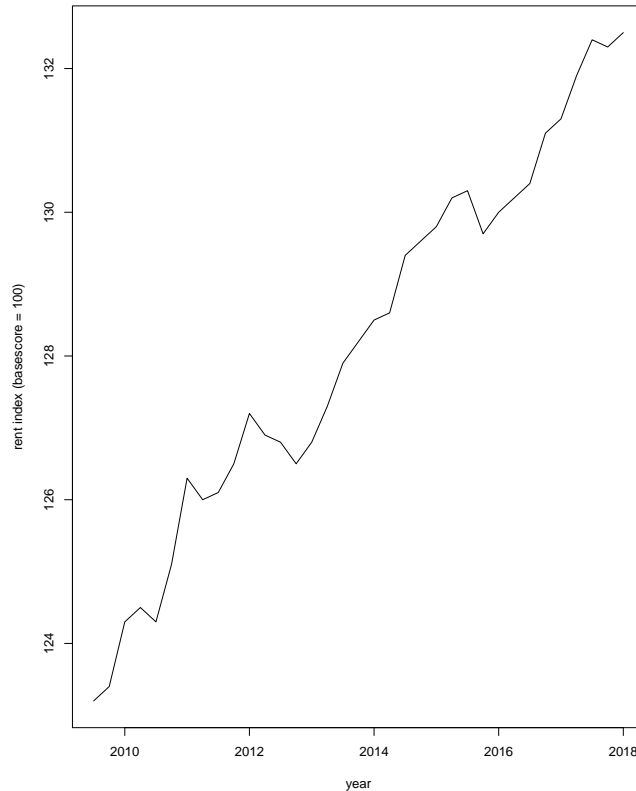


Figure 3.2: Swiss rent index, years 2009 to 2018

a linear one, we are starting by fitting different polynomial trends by ordinary least squares estimation. We are going to fit 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. This can be well seen at 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 fitted linear models is a different one: to use residuals in order to construct

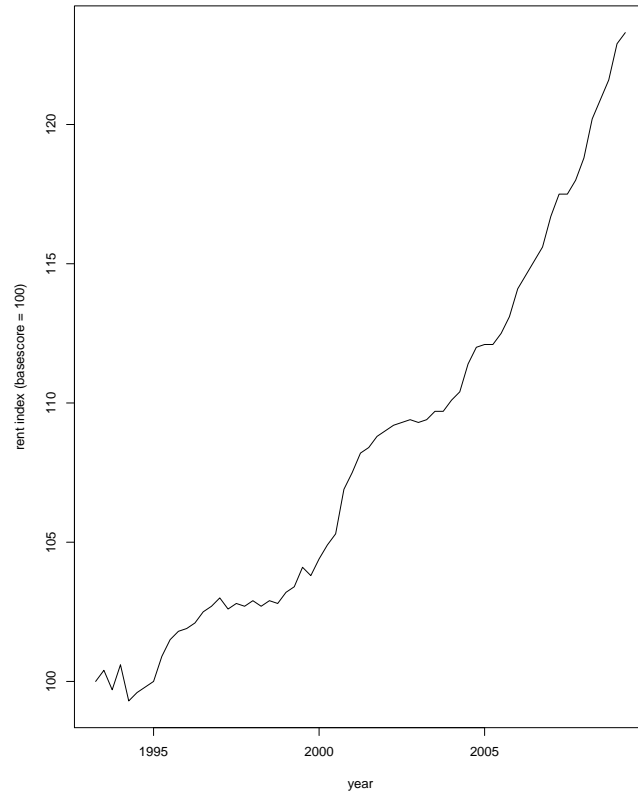


Figure 3.3: Swiss Rent index, years 1993 to 2009

```

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

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

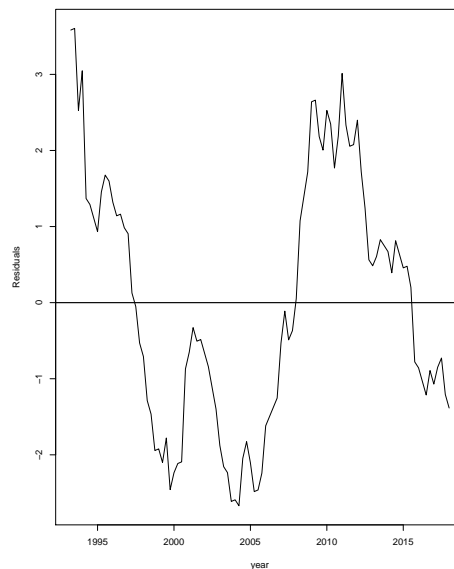


Figure 3.6: Residuals of a fitted linear model

3.2.1 Diagnostics of the fitted polynomial trends

Unfortunately the plots for polynomial trend elimination don't show any stationarity, as the series wanders up and down for quite some periods. 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 , meaning non-stationary.

Furthermore from the ACF-plot (figure 3.10) and the PACF-plot (figure ??) exemplified on the cubic model, we can see that there are a lot of significant lags outside the 95-percent-confidence bounds of $\pm 1.96\sqrt{n}$. These bars don't die out quickly. So the Rho_{hat} can be a useful indicator of non-stationarity (Brockwell and Davis, 2002, p. 21). Hence, we can conclude

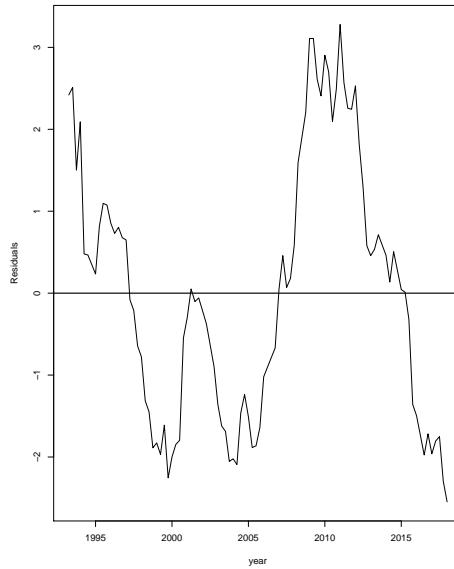


Figure 3.7: Residuals of a fitted quadratic model

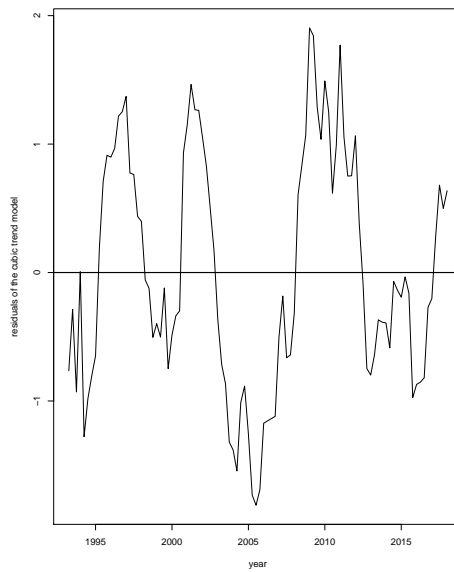


Figure 3.8: Residuals of a fitted cubic model

our covariances depend on time and therefore the series is non-stationary. similar time-depending patterns hold true for the autocovariance-functions of our other polynomial trend models: the linear, the quadratic and the

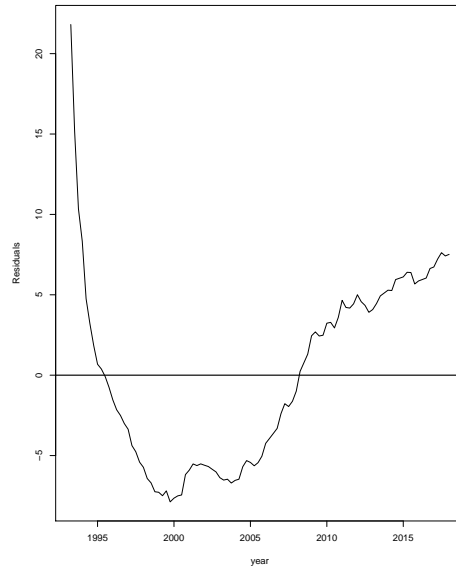


Figure 3.9: Residuals of a fitted logarithmic model

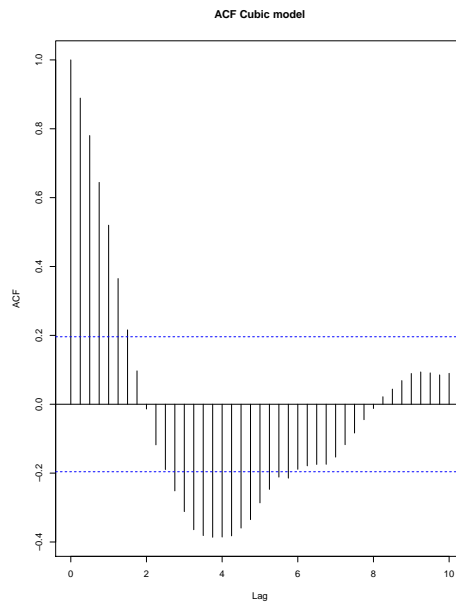


Figure 3.10: Sample autocorrelation function of the cubic model

logarithmic model.

In addition we get numerical evidence from three different tests checking 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 too sure yet.

The following figure ?? shows an overview over the tests considering non-stationarity and independence of the observations.

Since the fitted polynomial models doesn't help us to give a stationary series we will consequently try to differentiate our data in the next subsection.

3.3 Trend Elimination by differencing

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 (Brockwell and Davis, 2002, 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 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 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 $\pm 1.96\sqrt{100}$ (brooks davis p.39) (see figure 3.15). To conclude, the once-differenced series is too versatile and contains too much randomness to get good predictions. The Ljung-Box-Test (Ljung and Box, 1978) 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 E_0 and Variance σ^2 no more modelling would be necessary which is not what we want as we evoked in the introduction to this section

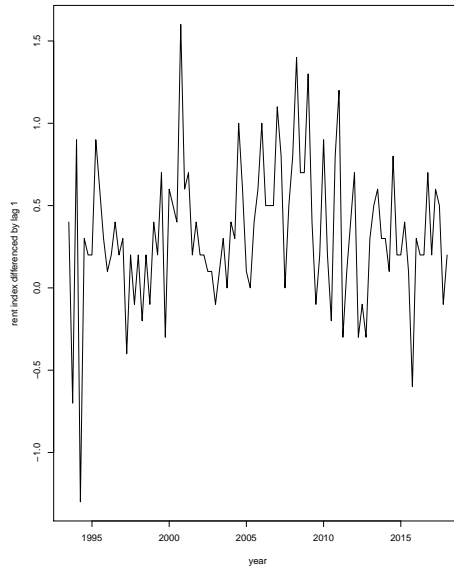


Figure 3.11: once-differenced rent index

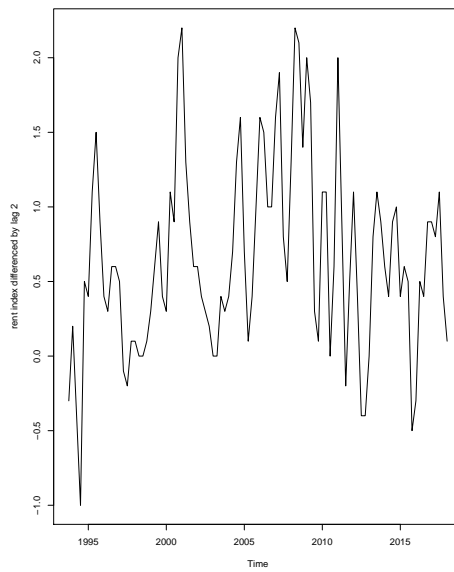


Figure 3.12: twice-differenced rent index

3.3.1 Diagnostics 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

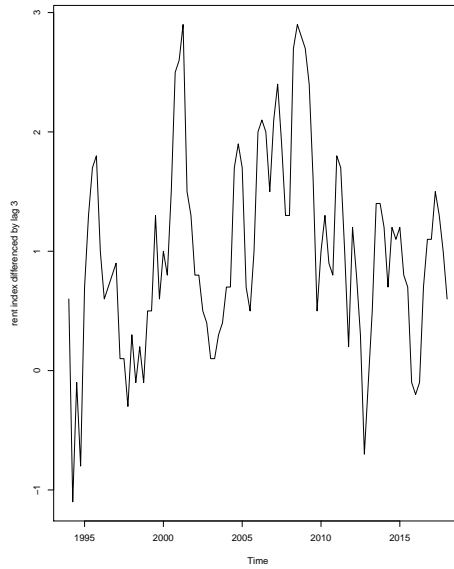


Figure 3.13: triple-differenced rent index

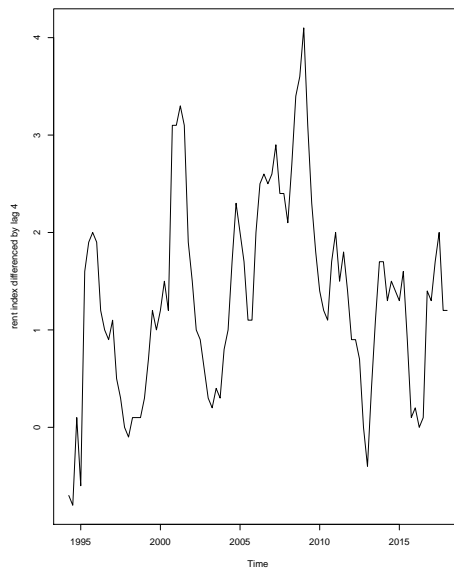


Figure 3.14: quadruple-differenced rent index

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 touch-

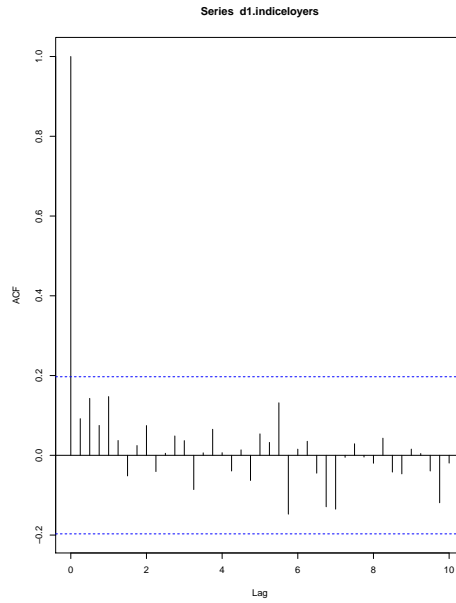


Figure 3.15: auto-correlation function of the once-differenced series

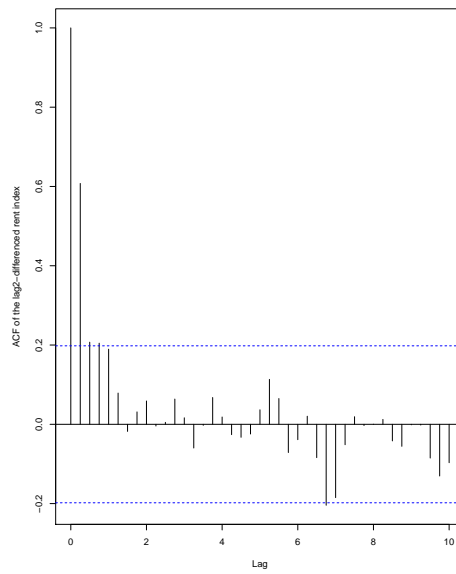


Figure 3.16: auto-correlation function of the twice-differenced series

ing 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

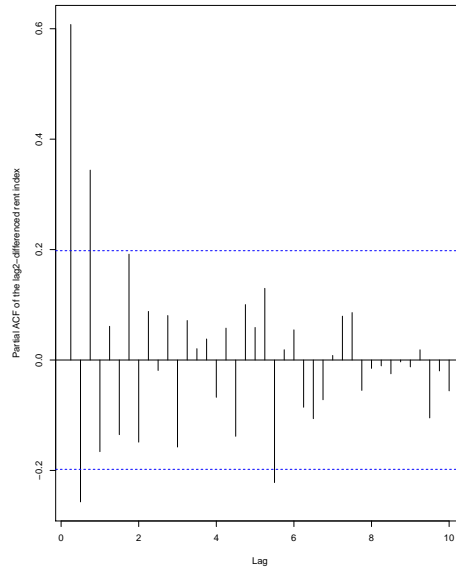


Figure 3.17: Partial auto-correlation function of the twice-differenced series

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, alternative = "stationary", k=2)` H_0 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 aswell 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.

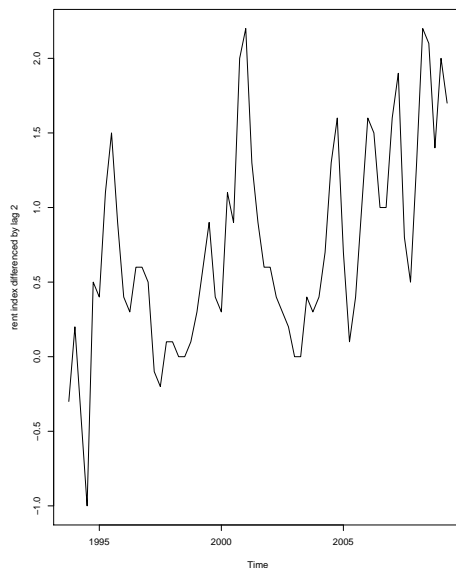


Figure 3.18: twice-differenced series from 1993 to 2009

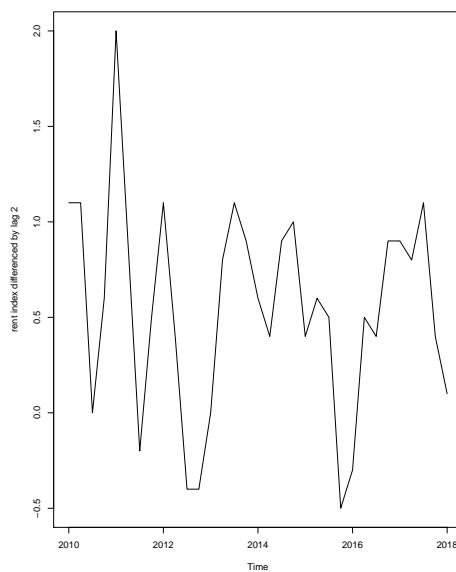


Figure 3.19: twice-differenced series from 2009 to 2018

3.3.2 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 $\pm 1.96\sqrt{100}$

- bounds.

To conclude, our transformation 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). Finally to continue fitting and testing a model we center our twice-differenced series by subtracting the mean in order to...

4 Fitting and Testing the Model

4.1 preliminary estimation and order selection

From the sample ACF and PACF (figure 3.16) (figure 3.17) of our centered twice-differenced model we can see in the ACF-plot a highly significant spike at lag 1 and the following bars dying out quickly, lying inside the 95percent-confidence – bounds (with bars at lag 2 and 3 slightly touching the 95percent-confidence-bounds). The combined view of the PACF-plot with its exponentially decreasing patterns (in absolute values) and the 1 spike in the ACF-plot let us suggest an MA(1)-model could fit our data quite well. However, since we are operating with only 100 observations, one can not exclude that there could be underlying exponentially decreasing patterns in the ACF plot as well. Exponential decrease (in absolute values) in both ACF and PACF plot would rather suggest an ARMA-model then. Since we're following a conservative approach we are going to test for MA(1) as well as for ARMA-model. In opposition to exclusive AR(q)- or exclusive MA(p)-models where we can derive the order of p or q from the number of spikes in the ACF and PACF plot, the same holds not true to guess the orders of p and q for ARMA-models. Therefore in case of ARMA-models we have to choose the order of p and q in another way, by comparing models with different p and q and decide which performs best in terms of the corrected Akaike Information Criterion (AICC). The following matrix (figure ??) shows the AICC for different orders of p and q. The one with the lowest value performs best (Sakamoto et al., 1986). We can see that an MA(1)-model gives us the lowest AICC-value which is very well confirmed by the ACF/PACF plots. However, since the sample ACF and PACF, as mentioned above, could possibly both decrease exponentially (in absolute values), we are going fit as well an ARMA(1,1) and ARMA(1,2) model, since both of them show very low AICC-values, too (see AICC-Matrix in figure). Estimation of the model candidates We estimate our three model candidates by the arima function . This function implies a rough estimation by (hannan rihannen or yule walker) and afterwards as optimization method bfgs (Brockwell, P. J. and Davis, R. A. (1996) Introduction to Time Series and Forecasting. Springer, New York. Sections 3.3 and 8.3.). aswell reference the bfgs.

4.2 Comparing the sample ACF/PACF and model ACF/PACF

If we simulate several models ($n=10000$) with our estimated coefficients we can see that the simulated model ACF and PACF of the MA(1) suits our sample ACF and PACF best. The one big spike at lag1 in the sample ACF is very well modeled. However the PACF is decreasing too slowly, indicating that the estimated theta with a value of 0.99 is probably overestimated. (It would fit the sample PACF better when we reduce the evidently overestimated MA to 0.4, for instance).

BrockwellDavis p. 193 a root near 1 of the autoregressive polynomial suggests that the data should be differenced before fitting an ARMA 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 requires an initial value which we obtained by the Hannan-Rinnanen-Alg. The BFGS-algorithm is a local optimization algorithm, meaning that it does not find the global 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 the BFGS-algorithm a good initial value to get more secure that it will find the global minimum and not only a local 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.
- Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root,. *Journal of the American Statistical Association*, 74:427–431.
- Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178.
- 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.
- Sakamoto, Y., Ishiguro, M., and G., K. (1986). *Akaike Information Criterion Statistics*. D. Reidel Publishing Company.