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

Faralli Zully, Spörri Marc

May 12, 2018

Contents

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.

In this project we analyze the time series of the quarterly collected Swiss rent index from 1993 to 2018. The main goal of the project is to draw inference from the rent index time series and to predict its future values and a future trend. In order to do that we do first a transformation of the data in a mean-centered stationary time series by eliminating the trend and possible seasonal components. Afterwards we are going to model our transformed, stationary data by fitting and testing a set of autoregressive moving average processes (ARMA) in order to select the best candidate model for our transformed, stationary data (Brockwell and Davis, 2002, p. 82–110). The selected model will be used to describe and interpret our data as well as to predict future values for the Swiss rent index.

2 Data Description

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

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

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 diagnostig 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(Brockwell and Davis, 2002, p. 23). As we can see from figure 4.1 there seems to be a clear positive monotonic trend in the time series of the Swiss rent index. However, the increase seems to be higher in the 1990ies. It may be advisable to analyse the series by breaking it into two segments (Brockwell and Davis, 2002, p. 23) . For that purpose we have a look at the segmented plots 1993 to 2009 and 2009 to 2018 (see figure **??**).
We can see a slightly slower increase in the period from 2009 to 2018 indicating that during the years 1993 to 2009 the growth in rents was faster than in the years after the year 2009. This fact can be well explained by the big real estate depression in the early 90ies in Switzerland, therefore the time series starting at 1993 is consequently starting from a relatively low initial level and, given the low initial level and the better conjunctural perspectives, the growth was faster during the 1990ies (irgendwas zitieren 90er jahre depression), whereas from the year 2009 on the growth in rents slowed down, which can be very well explained by the US subprime crises triggered in the year 2008 and followed by a long-lasting global recession (zitiere irgendwas).
However, they are apparently no outliers, sudden changes or major discontinuities and there is evidently no reason to do further data adjustments or variance-stabilizing transformations like Box-Cox-transformations, for instance boxcox64. Therefore we start our work in the next section by transforming our data into a stationary time series (Brockwell and Davis, 2002, p. 45–82).

4 Transformation of the Data into a Stationary Time Series

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 obtain a 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 neither, since there would no further modelling to be done except to estimate their mean-expectation (which would be zero in the centered transformed series) and variance. The objective is to obtain 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 stationarity by tests like the Augmented Dickie-Fuller-Test for the null hypothesis of a unit root (with the alternative hypothesis of stationarity, by consequence (Dickey and Fuller, 1979)) as well as 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 by examining some simple tests whether the obtained residuals are values of independent and identically distributed random variables (as mentioned above: that should not occur, otherwise we have to much randomness and our prediction work would be simply done by estimating mean and variance of the white-noise).
In a final step we check them visually by plotting the autocorrelation-function as well as the partial auto-correlation 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, on the other hand (Brockwell and Davis, 2002, pp. 83–110). Once we have found a good way to transform our data into a stationary series we will fitting and testing a set of candidate models for our transformed data in section 5.

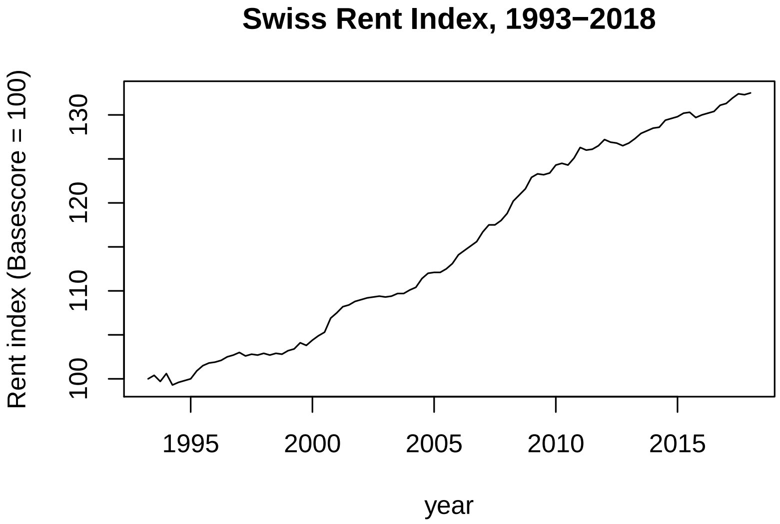


Figure 4.1: Swiss rent index, years 1993 to 2018

4.1 Testing Seasonal effects

As we have seen in figure 4.1 a trend was obviously visible, however we cannot be that sure if it doesn't exist a seasonal component neither. Therefore we check first whether there exist (or not) possible seasonal components, afterwards we want get rid of the trend.
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 4.3 none of the estimated seasonal coefficients are significant. We can be sure now that they are no seasonal impacts on our data, hence what remains to do is to eliminate the trend. (Our twice-differenced, non-seasonal decomposition model looks like the following (Brockwell and Davis, 2002, p. 24):

$$X_t = m_t + Y_t, \text{t} = 1,\ldots,\text{n}, where EY_t = 0 and m_t is the trend component and Y_t is the white - noise \tag{1}$$

)

4.2 Method 1: 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 4.1 indicates a probable polynomial trend, especially a linear one, we are starting by fitting different polynomial trends by ordinary least squares estimation (**?**).

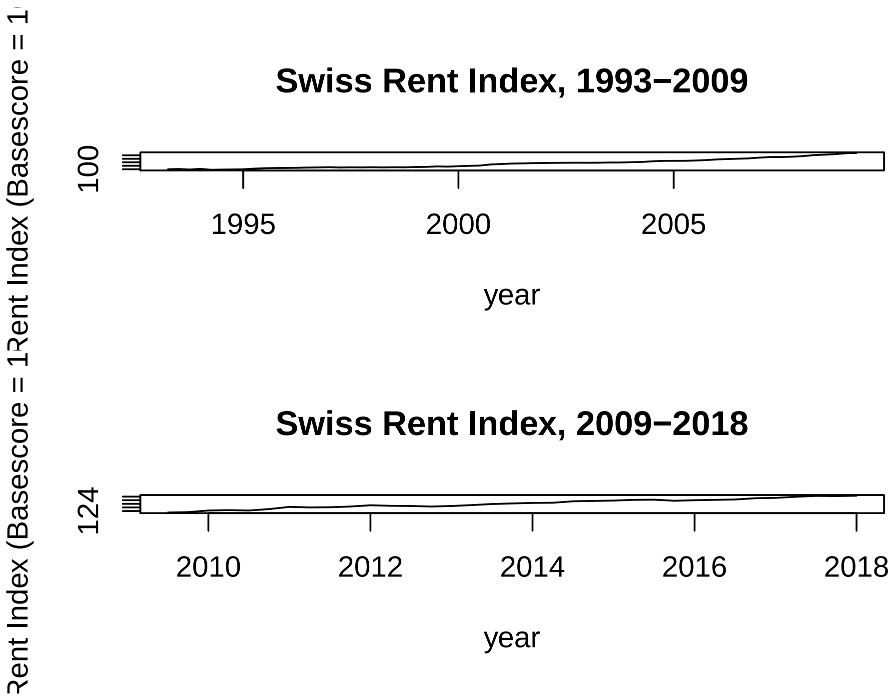


Figure 4.2: Swiss Rent index, periods from 1993 to 2009 and from 2009 to 2018

```
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 4.3: summary statistics of the fitted seasonal model

4.2.1 Modelling the estimated noise sequence

We are now modelling a noise sequence (residual time series) by estimating and subtracting a linear, a quadratic, a cubic 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 the time series 4.1 looks more like a linear than an exponential trend at first glance, we are going to test for a logarithmic trend either.

The coefficients in all 4 models are highly significant and with an coefficient of determination (the adjusted R^2) of around 98 to 99 percent [1], the linear, the quadratic and the cubic model would explain our data extremely well (which is not surprising given the fact that the trend is consistently positive monotonic and therefore highly consistent with other positively mono-

[1] This means that the polynomials can predict upto 98 to 99 percent of the variance of the variance in the dependent variable (rent index

```
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 4.4: summary statistics of the fitted cubic model

tonic models as our 4 polynomial trend models are). The logarithmic model showed a bit less explanatory power with an R^2-Value of 0.72, even though its coefficient is highly significant as well. [2]

Unfortunately the four residual plots for the polynomial models are evidently not indicating a good stationary noise sequence, as the series wanders up and down for quite some periods (see figure 4.5). Time-independent expectation and variance are basic properties of stationary processes (Brockwell and Davis, 2002, p. 49). This is not the case here, as we can see large covariances and residuals which are evidently depending on time. However, we want to test the four estimated trends, anyhow.

4.2.2 Testing the noise sequences of the subtracted polynomial trends

We are going to test the estimated trends first visually, by looking on the sample autocorrelation-function (ACF) and the sample partial autocorrelation-function (PACF), second we are applying three diagnostic tests for stationarity and independence of the observations, respectively: The Ljung-Box-Test (Ljung and Box, 1978) for independence as well as for stationarity the Augmented Dickie-Fuller-Test (Dickey and Fuller, 1979) and the KPSS-Test (Kwiatkowski et al., 1992).

From the ACF-plot and the PACF-plot (cite bd02 zu acf pacf)) in (figure **??** exemplified on the cubic model because it was the best-performing polynomial model in terms of the adjusted R^2, we can see that there are a lot of significant lags outside the 95-percent-confidence bounds of $+/-1.96/sqrtn$. These bars don't die out quickly and this large amount of significant sample auto-correlations Rho_hat are indicating time-dependence and non-stationarity (Brockwell and Davis, 2002, p. 21). [3]
Let's now have a look at the results of the three tests in figure **??**: 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 not independent (ist das so??? über-

[2] We have to be careful with the interpretation of the p-values since this regression models assume independence of the observations (which hardly never holds true in a time series except white-noise). we don't want to do inference, our purpose of the summary statistics of fitted linear models is a different one: we want to get an idea whether there is a polynomial trend to eliminate in the time series in order to get the series stationary

[3] Similarly strong auto-correlation patterns hold true for the ACF and PACF plots of our other polynomial trend models (linear, quadratic and logarithmic) which will not be shown here.

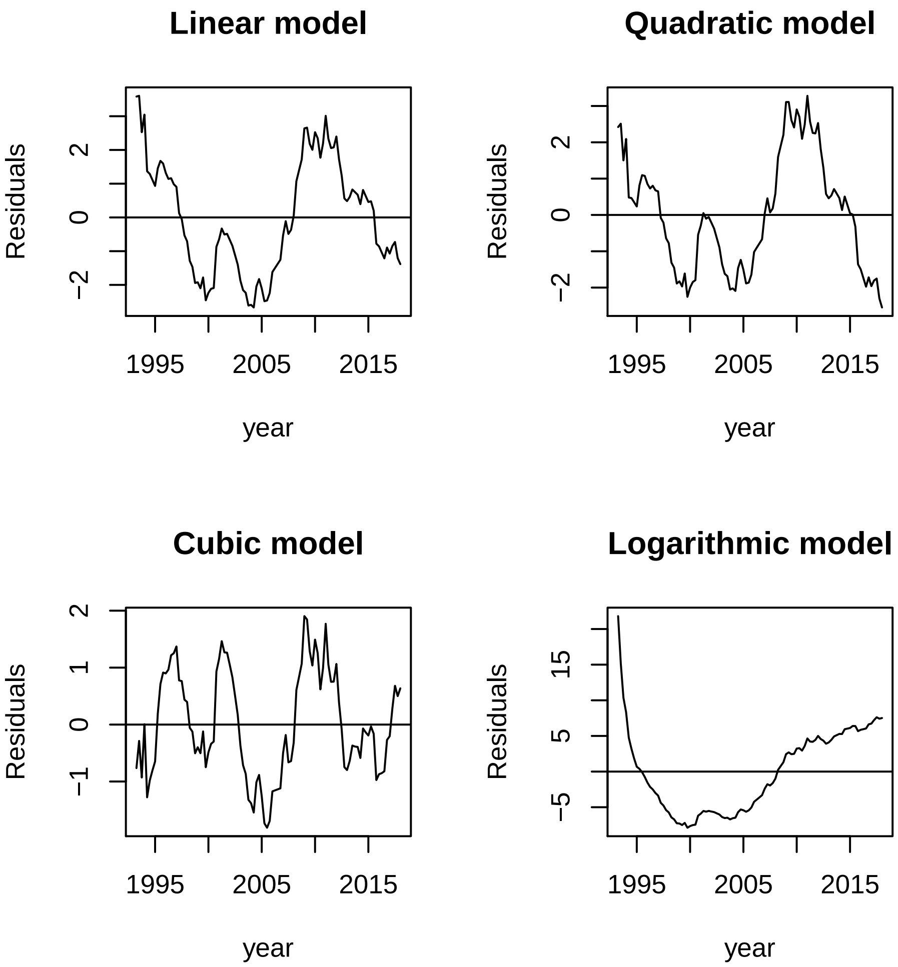


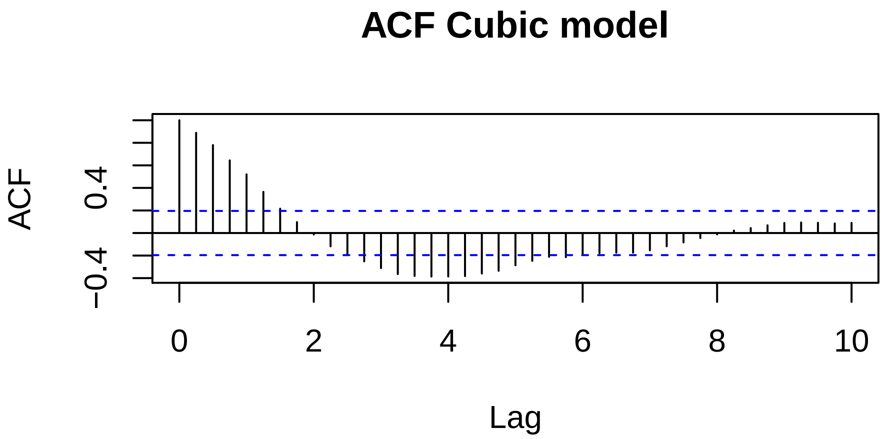
Figure 4.5: Residuals of the fitted linear, quadratic, cubic and logarithmic model

prüfen). We can see that the Ljung-Box Hypothesis H_0 of independence has to be rejected for all four polynomial models. In fact, that is not bad in general, since we want to avoid a white-noise-series, but given the fact of the large covariances seen in the residual plots small p-values for the Ljung-Box-Test emphasizes in these four cases the assumption that time-independance is not guaranteed.
Furthermore the Augmented Dickie-Fuller test[4] and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test provide as well strong evidence against stationarity for all 4 polynomial cases.

4.3 Trend elimination by differencing

Since the fitted polynomial models didn't help to get a stationary time series we will try to eliminate the trend by differencing. We will differentiate the data at different lags (1, 2, 3 and 4) in order to generate a noise sequence and consequently get a stationary series. (Brockwell

[4] We should take in consideration that the Dickie-Fuller-Test is mostly for testing autoregressive series for the presence of a unit root of the autoregressive polynomial in order to decide whether or not a time series should be differenced (Brockwell and Davis, 2002, p. 194)) (Dickey and Fuller, 1979)



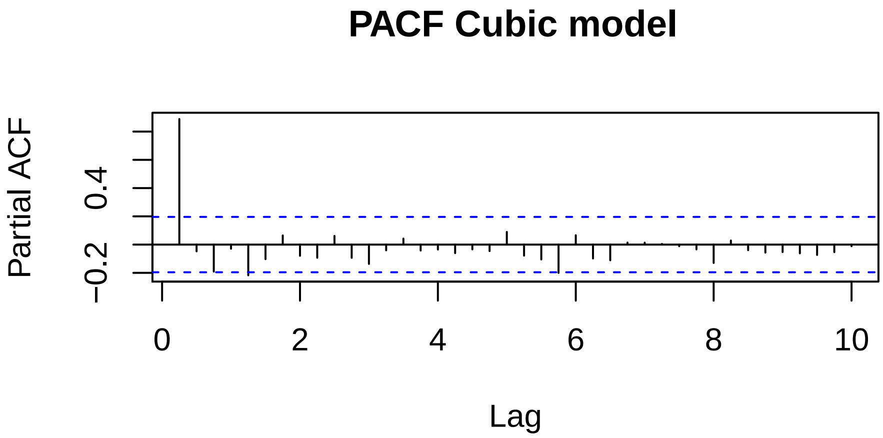


Figure 4.6: Sample Autocorrelation and Partial autocorrelation function of the Cubic model

and Davis, 2002, p. 35)). Again we will show first the residual plots and testing them afterwards by the sample ACF and PACF as well as the Ljung-Box-, Augmented Dickie-Fuller- and the KPSS-Test.

4.3.1 Modelling the estimated noise sequence

As we can see in the figure 4.7 the first two models look already quite promising in contrast to the residual time series above by polynomial trend elimination, showing here obviously time-independent noise sequences with presumably constant expectation and variances.
However, the once-differenced series shows mostly patterns of a white-noise which is not what we intend since in case of a white-noise and too much randomness prediction quality becomes rather poor (Expectation of all the predictions X_{t+1}, X_{t+2}, X_{t+3} etc. would become zero in case of a mean-centered series).

4.3.2 Testing the noise sequences of the differenced series

As we can see in figure 4.8) the white-noise of the once-differenced series is confirmed by the sample ACF where all lags up to 40 fall within the bounds of $+/-1.96/sqrtn$ (Brockwell and

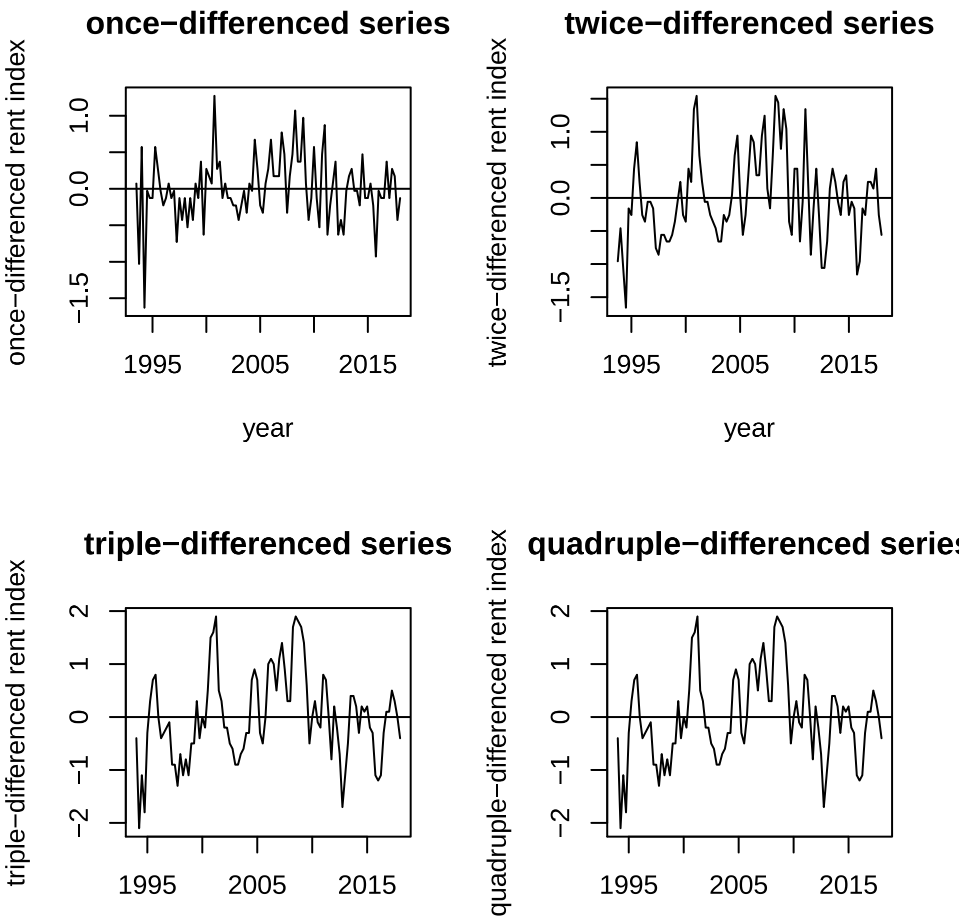


Figure 4.7: mean-centered, differenced series (1x, 2x, 3x, 4x)

Davis, 2002, p. 39). This underlines that the once-differenced series is to versatile and contains too much randomness in order to do good predictions.

In contrast, the sample ACF of the twice-differenced model doesn't indicate a white-noise: there is at least a highly significant correlation at lag 1 and the following three lags are touching the $+/-1.96 sqrt/n$ - 95percent-confidence-bounds, followed by no more significant bars which die out quickly. This is a good sign, meaning that on one hand we have at least correlation at lag 1 which is needed to do prediction, on the other hand the covariances seem obviously not too large to be considered time-dependant and therefore we can assume a stationary series here.[5]

From the tests overview in figure **??** we can see for the once-differenced series a p-value of 0.35 for the Ljung-Box-Test. That was to be expected, meaning that we cannot reject the Hypothesis H_0 that there is independence between the observations (Ljung and Box, 1978) and therefore confirming the white-noise structure in the once-differenced series.

In contrast again, the Ljung-Box-Test for the twice-differenced series rejects the H_0 of inde-

[5] Triple- and quadruple-differencing is not appropriate neither. Whereas once-differencing is a white-noise, the sample ACF and PACF of the triple- and quadruple-differencing showed significant bars periodically appearing outside the $+/-1.96/sqrtn$ - bounds which doesn't allow reliable modelling. This plots will not be shown here.

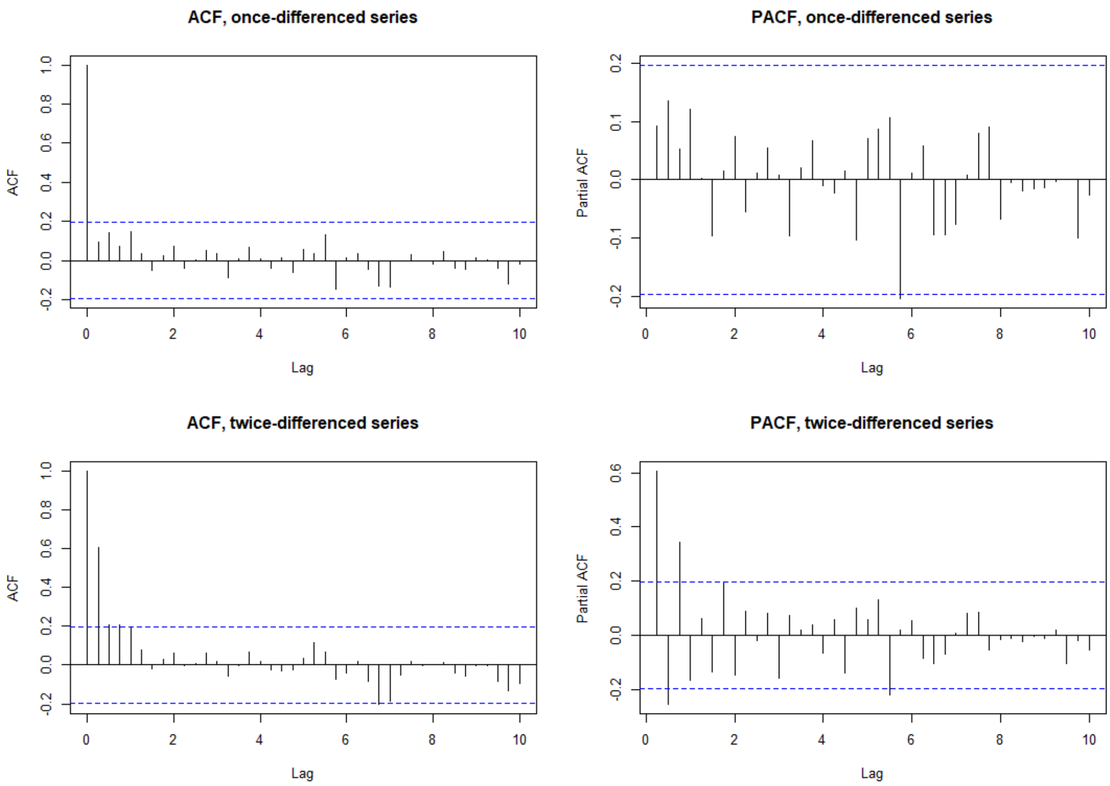


Figure 4.8: sample auto-correlation and partial auto-correlation function of the once- and twice-differenced series

pendence which is a pleasant result giving us the sureness that there is no white-noise in the twice-differenced series.
The Augmented-Dickie-Fuller- and the KPSS-Test suggest for the once- as well as for the twice-differenced trend elimination a stationary time series. That is a very good result and we can finally conclude to stick with the twice-differentiation in order to transform our data, since the once-differenced series is characterized by too much randomness and white-noise and the triple- and quadruple-differenced series show too large covariances (time-dependence/non-stationarity). We now center our transformed data by subtracting the mean so that they are prepared fit and test ARMA-models.

4.4 Cross-validating the twice-differenced series

5 Fitting and Testing the Model

5.1 Order and model selection

From the sample ACF and PACF (figure 4.8) of our twice-differenced series 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 an 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. This is as well confirmed by the ACF/PACF plots as evoked above. However, since the sample ACF and PACF, could possibly as well interpretated both as decreasing exponentially (in absolute values), we are going to fit as well an ARMA(1,1) and an ARMA(1,2) model, especially because both of them show a very low AICC-values, too (see AICC-Matrix in figure). In the footnote: Other estimation by the autofit()-function of the itsmr R-Packages (cite itsmr) suggests as well an MA(1)-model as best model with a very similar estimate for θ_1 and very close AICC. Furthermore ARMA(1,1) and ARMA(1,2) are suggested by the autofit()-function showing very low AICC-values as well. That can be seen as a robustness check in terms of selecting the order of p and q, end footnote.

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.

5.2 Evidence from simulation: Comparing the sample ACF/PACF and the 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 modellized. However, the PACF is decreasing too slowly, indicating that the estimated θ_1 with a value of 0.99 is probably overestimated. A root of the moving-average polynomial close to 1 indicates that the data were probably overdifferenced (Brockwell and Davis, 2002, p. 193).

Further evidence for an ARMA(1,2)-model as best candidate model is provided by taking into account an additional MA-term. We can see then that in an ARMA(1,3) $\hat{\theta_1}$ and $\hat{\theta_2}$ don't differ a lot from the $\hat{\theta_1}$ and $\hat{\theta_2}$ in the ARMA(1,2)-model. Furthermore $\hat{\theta_3}$ is not significant since it's close to 0 with a large standard error, indicating that the gain by adding a third $\hat{\theta}$-coefficient would be rather poor, hence, an moving average of order 2 is sufficient enough. But we have to note that $\hat{\theta_1}$ and $\hat{\theta_2}$ are rather small compared to their standard errors. $\hat{\theta_1}$ is only slightly larger than its standard deviation. $\hat{\theta_2}$ is 2.5 times larger. However, since we pointed out a unit root issue in the moving average polynomial for the MA(1)- and ARMA(1,1)- models we want to stick with the ARMA(1,2) as our best candidate model, anyhow. Our final model equation is therefore:

$$X_t - \hat{\phi_1} X_{t-1} = Z_t + \hat{\theta}_1 X_{t-1} + \hat{\theta}_2 X_{t-2} \tag{2}$$

5.3 Diagnostics of the models

In order to check the validity of our ARMA models we apply the following diagnostics: First, we plot the rescaled residuals and verify that there is not some obvious dependence, trend or else. We should see a white-noise if the model is valid. Second, we have a look at the ACF of

the rescaled residuals. Since we are supposed to have a white-noise approximately 95percent of the sample-autocorrelations rhohhat should be inside the bounds $+/-1.96/sqrtn$. Third we use the Ljung-Box Test (Ljung and Box, 1978) to validate that the residuals can be considered as a white-noise by showing p-values for all possible lags (cite page 5 Lecture notes 11).
We are visualizing this by using the tsdiag() function from the "stats" R-Package (cite R Core Team). The following figure xy shows these three diagnostics for each of our model candidates MA(1), ARMA(1,1) and ARMA(1,2). We can see that the standardized residuals are supposed to be white-noise in all three cases. In addition, from the ACF of the rescaled residuals we can see that not a single bar (out of 20) lies outside the bounds bounds $+/-1.96/sqrtn$. And from the Ljung-Box-Tests we can see that the p-value is larger than 5 percent for each lag for each model, meaning that we cannot reject the Hypothesis H0 that these residuals are a white-noise. So the residuals for all three models seem to be indeed a white-noise, therefore all three models would be valid (good ACF, large p-values on Ljung-Box-Tests) and could be used for forecasting.

5.4 simulation?

However, we want to forecast on the base of our ARMA(1,2) since this model has proved to serve our case best for different reasons evoked above and especially, it doesn't consist of polynomials on or close to the unit circle.

6 Prediction of future values

In this section we want to predict future values on the base of our fitted ARMA(1,2)-model. We will do that by first forecasting on our transformed (twice-differenced) series. Afterwards we will do a retransformation into the initial time series and predict its future values. In addition, we will give 95 percent confidence bands in case of normally distributed residuals.

6.1 Forecasting the transformed series

From figure 6.1 we can see the prediction of a process of adaptation to the zero-mean for the next few quarters. This is in fact nothing surprising for a mean-centered stationary time series and the smooth positive adaptation process seems to be a reasonable forecast with quite some prediction quality in comparison to a white-noise which would predict directly a zero for the tplus1 observation. Likewise the quite large 95-percent confidence bands are not that surprising in a strongly differenced series as we have here. Differenced series like that can contain quite some randomness and result by consequence in large amplitudes as we can see. Footnote: note that the indication of confidence bands is weakly justified here since the series' residuals are successfully tested for normal distribution: the Shapiro-Wilk Test showed a p-value of 0.07, meaning that the hypothesis H0 of normal distribution can knapp not be rejected (cite shapiro wilk). However, a qq-plot showed patterns of normal distribution for the close-to-mean observations but slightly heavier tails compared to a true normal distribution. Therefore interpretations of confidence bands are to be handled with care, end footnote. However, it would be quite desirable to reduce the variance and by consequence the confidence bands by increasing drastically the number of observations.

6.2 Forecasting the initial series

Now that the stationary time series is predicted we need to obtain predictions for the initial time series. For this purpose we have to consider how we removed the trend and, by consequence,

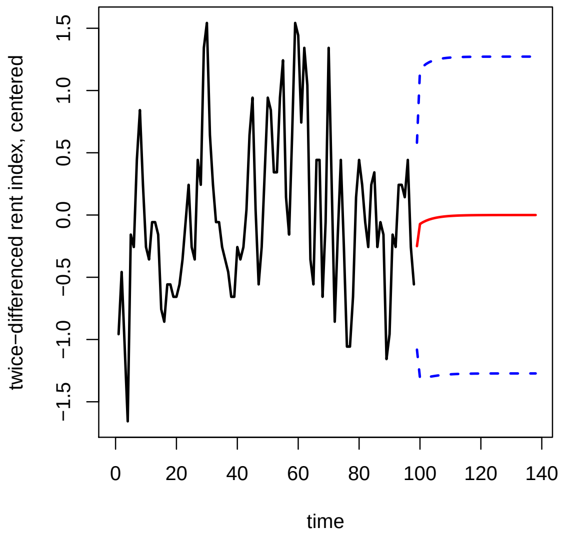


Figure 6.1: Prediction and Confidence Bands for the twice-differenced Series

retransform the twice-differenced series into the initial series.

This smooth positively predicted trend goes confirm not only by the long-lasting and consistently increasing trend in the past, but as well with theory about the rent market where expert predict a slight slow-down of the growth in rents but a positive trend still remaining (but less increasing) (zitiere nzz oder so).

7 Discussion

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

The strength of this process was the success to get the data stationary in a way that the transformed residuals don't show too much randomness (white-noise) on the one hand, and on the other hand that the data anyhow show a bit of correlation in order to allow predictions based on past observations. That was not an easy task, since the initial time series looked mostly monotonically increasing and therefore strongly positively auto-correlated, but differencing the data twice did the job. On the same time, the twice-differenced-transformation turned out to be a weakness for modelling since there exist the danger of overdifferencing (Brockwell and Davis, 2002, p.194), indicated especially by the unit root in the moving average polynomials in case of an MA(1)-model and in the ARMA(1,1) whose estimated theta1 coefficients were on the unit circle or very close to the unit circle. The problem here lies in the maximum likelihood estimation which possesses properties which chan pose problems for estimation(Davidson, 1981). When a root of the process is close to the unit circle the asymptotic distribution of the maximum likelihood can be quite inadequate. However, the estimation by maximum likelihood differ only slightly from other estimations by other methods as (Davis and Dunsmuir, 1996) have shown. Therefore our "problematic" estimate of theta close to 1 can still be considered quite accurate. Furthermore (Plosser and Schwert, 1977) underline that a root close to the unit circle is not that big of a deal as long as the perturbances of overdifferencing are understandable.

To conclude, the choice of the correct data transformation was the dilemma between once-differentiation which leads to a white-noise series which doesn't allow good prediction, and the twice-differentiation which allows better prediction but suffers possible over-differentiation.. However, our model of choice, the ARMA(1,2) process, doesn't show unit root problems any more, but its theta coefficients lack of reliability since its standard errors were not that small. In the context of large variance the relatively small n of 100 observations played an important role here. Hence, for further predictions on the rent index more data would be desirable.

8 Conclusion

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

Further improving work could be done on the transformation into a stationary time series. Given the fact that the quite promising MA(1)-model (in terms of the ACF and PACF patterns aswell as the AICC) didn't turn out be an ideal choice because of the unit root detected in the moving average polynomial and the consequential danger that excessive use of the difference transformation could induce a non-invertible moving average process (Brockwell and Davis, 2002, p.194) (Plosser and Schwert, 1977), one could try a third transformation option: meaning neither fitting a polynomial trend, neither differentiating, but try the Brockwell-Davis-Method (cite b d filtering method) for instance. Another possibility would be to draw inference in another way, meaning that the estimation of the theta coefficient shouldn't rely only on the maximum likelihood function, but to try alternative estimations like a locally best invariant unbiased (LBIU) approach, for instance Davis and Song (2011).
Last but not least, furthermore improving work could be done during the transformation step by treating the two periods before the subprime crises (year 2009) and after in a different way, since they show slightly different increasing patterns. One possibility could be, for instance, to try first to logarithmize the stronger increasing data from 1993 to 2009 and afterwards eager for a linear trend elimination or differencing transformation on the total of the data.

References

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer New York, New York.

Davidson, J. (1981). Problems with the estimation of moving average processes. *Journal of Econometrics*, 16:295–310.

Davis, R. and Dunsmuir, W. (1996). Maximum likelihood estimation for ma(1) processes with a root on or near the unit circle. *Econometric Theory*, 12:1–29.

Davis, R. and Song, L. (2011). Unit roots in moving averages beyond first order. *The Annals of Statistics*, 39(6):3062–3091.

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.

Plosser, C. and Schwert, G. (1977). Estimation of a non-invertible moving average process: The case of overdifferencing. *Journal of Econometrics*, 6:199–224.

Sakamoto, Y., Ishiguro, M., and G., K. (1986). *Akaike Information Criterion Statistics*. D. Reidel Publishing Company.