

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

Faralli Zully, Spörri Marc

May 18, 2018

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 3 |
| 2 | Data Description and Data Analysis | 3 |
| 3 | Transformation of the Data into a Stationary Time Series | 4 |
| 3.1 | Method 1: Trend Elimination by fitting polynomial models | 4 |
| 3.2 | Method 2: Trend elimination by differencing | 5 |
| 3.2.1 | Modelling the estimated noise sequences | 5 |
| 3.2.2 | Testing the noise sequences | 6 |
| 4 | Fitting and Testing the Model | 7 |
| 4.1 | Order and model selection | 7 |
| 4.2 | Diagnostics of the model | 9 |
| 5 | Prediction of future values | 10 |
| 6 | Discussion | 12 |
| 7 | Conclusion | 12 |
| A | Tables | 15 |
| B | Simulations | 15 |
| C | Cross-validation of the lag2-differenced series | 15 |

1 Introduction

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 and Data Analysis

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.

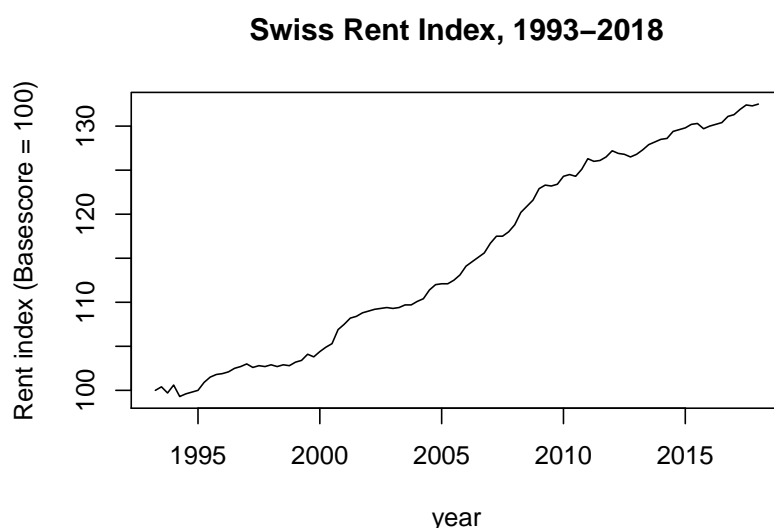


Figure 2.1: Swiss rent index, years 1993 to 2018

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).

The first step in any analysis of a time series is to identify possible discontinuities or a sudden change of level in the series (Brockwell and Davis, 2002, p. 23). As we can see from figure 2.1 there seems to be a evidently 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 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. 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 [**!!!TODO: (zitieren 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 (Box and Cox, 1964). 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).

3 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 test 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))

Furthermore we test the estimated noise sequences by examining some simple tests whether or not the obtained residuals are values of independent and identically distributed random variables (as mentioned above: that should not necessarily occur, otherwise we would be confronted with too much randomness in our stationary series and our prediction work would be simply done by estimating mean of the white-noise, which is zero for all predictions X_{n+1} , X_{n+2} , X_{n+3} etc. in case of a mean-centered stationary series).

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 4.

3.1 Method 1: Trend Elimination by fitting polynomial models

Since the time series 2.1 indicates a probable polynomial trend, especially a linear one, we are starting by fitting different polynomial trends by ordinary least squares estimation (Hastie et al., 2009, p. 11).

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 at first glance the time series figure 2.1 looks more like a linear than an exponential trend, we are going to test for a logarithmic trend either.

The regression models by ordinary least squares estimation (Hastie et al., 2009, p. 11) showed highly significant trend coefficients for all for polynomial models and no significant seasonal effects.¹

So all of the polynomial models would explain our data extremely well, which is indeed not surprising given the fact that the trend is consistently positively monotonic and therefore highly consistent with other positively monotonic models as our 4 polynomial trend models are indeed (further details on the estimated trend and seasonal coefficients as well as their significance can be found in figure A.1 in the tables appendix).²

Nevertheless, even though there is evidently a highly significant polynomial trend in the Swiss rent index series, that doesn't mean necessarily that its residual time series are stationary. Unfortunately that happened as we can see from the four noise sequences in figure 3.1. The series wander up and down for quite some periods, meaning that the basic properties of stationary processes, time-independent expectation and constant variance over time are not fulfilled here (Brockwell and Davis, 2002, p. 49). Since we can already be sure from the residual plots that trend elimination by fitting polynomial models doesn't help to obtain a stationary series, we will not show further diagnostics on these and continue by trying to eliminate the trend by differencing.³

3.2 Method 2: 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 hopefully get a stationary series (Brockwell and Davis, 2002, p. 35). Again we will show first the residual plots. Afterwards testing will be done by looking at the sample ACF and PACF as well as the Ljung-Box-Test for independence of the residuals and the Augmented Dickie-Fuller-Test for stationarity.

3.2.1 Modelling the estimated noise sequences

As we can see in the figure 3.2 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 variance.

However, the lag1-differenced series show mostly patterns of a white-noise which is not what we intend since in case of a white-noise too much randomness is induced and the 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).

¹ Since the data are collected quarterly we used a linear regression model with d=4 dummy predictors to check for significant seasonal coefficients, meaning a dummy for every quarter.

² Interpretation of p-values of regression model coefficients are to be handled with care in the case of time series since these models are assuming independence of the observations which hardly ever holds true in a time series except white-noise. One has to keep in mind that in case of time series we don't want to draw inference on the regression coefficients, our purpose for the regression coefficients is a different one: by looking at their p-value we get an idea whether or not there is a polynomial trend to eliminate in order to get a stationary time series

³ Since the residual plots show strong time-depending covariances, by consequence the plots of the sample auto-correlation function (ACF) and partial auto-correlation function (PACF) of the fitted polynomial models showed likewise a strong auto-regressive process of order 1 (decreasing values in the ACF-plot and 1 spike in the PACF-plot with a lot of significant auto-correlations outside the 95-percent-confidence bounds of $\pm 1.96/\sqrt{n}$). This can be found in the appendix

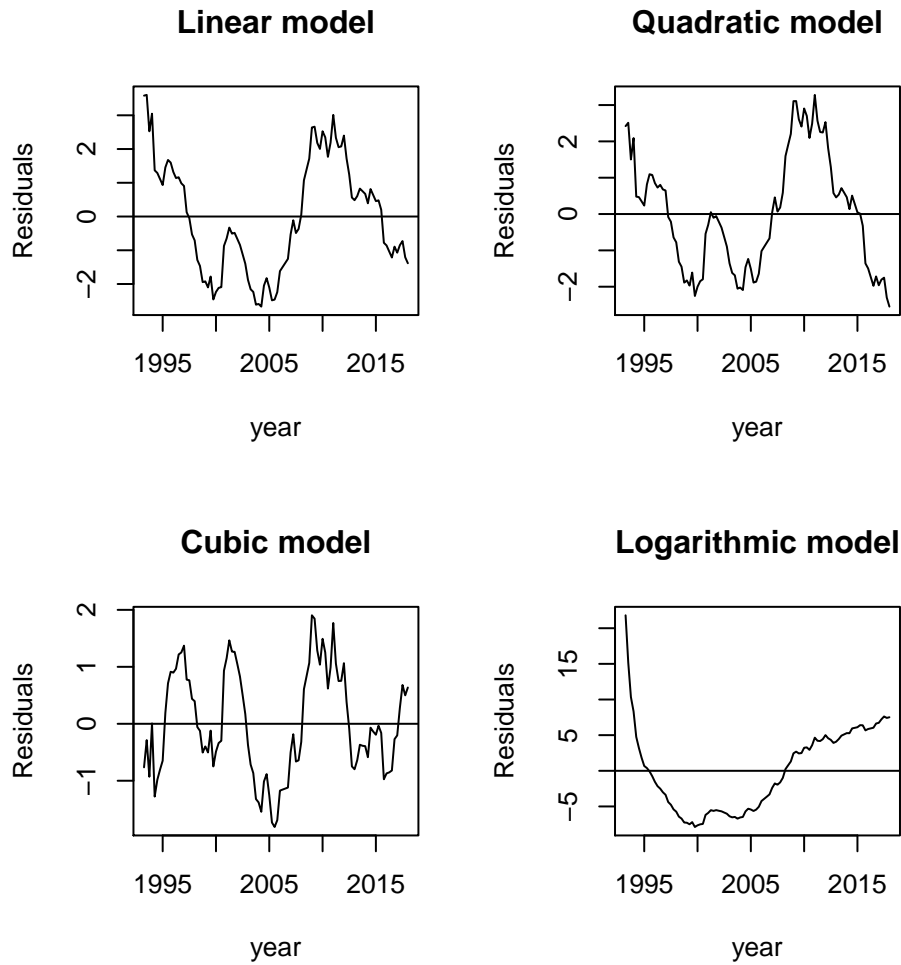


Figure 3.1: Residuals of the fitted linear, quadratic, cubic and logarithmic model

3.2.2 Testing the noise sequences

We are going to test the estimated trends first visually, by looking on the sample ACF and the sample PACF, afterwards 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 the Augmented Dickie-Fuller-Test for stationarity (Dickey and Fuller, 1979).

As we can see in figure 3.3) the white-noise of the lag1-differenced series is confirmed by the sample ACF where all lags up to 40 fall within the bounds of $\pm 1.96/\sqrt{n}$ (Brockwell and Davis, 2002, p. 39). This underlines that the lag1-differenced series is too versatile and contains too much randomness in order to do good predictions.

In contrast, the sample ACF of the lag2-differenced model doesn't indicate a white-noise: there is at least a highly significant at lag1 and the following three lags are touching the $\pm 1.96/\sqrt{n}$ - 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.

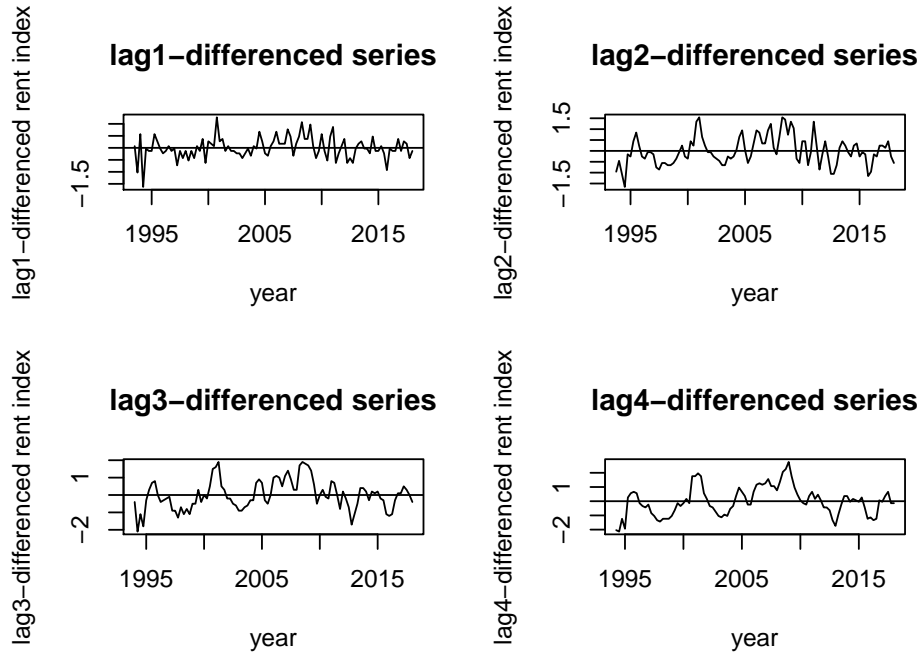


Figure 3.2: Residuals of the differenced series

| |
|--|
| • lag1-differentiation • lag2-differentiation • lag3-differentiation • lag4-differentiation • |
| • p-value Ljung-Box-Test (H_0 : White-noise) • 0.35 • 0.00 • 0.00 • 0.00 |
| • p-value Augmented Dickie-Fuller-Test (H_0 : Non-Stationarity) • 0.01 • 0.05 • 0.11 • 0.42 |

From the tests overview in section 3.2.2 we can see for the lag1-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 lag1-differenced series.

In contrast again, the Ljung-Box-Test for the lag2-differenced series rejects the H_0 of independence which is a pleasant result giving us the sureness that there is no white-noise in the lag2-differenced series.

The Augmented Dickie-Fuller-Test suggest for the lag1- as well as for the lag2-differenced trend elimination a stationary time series (p-values < 0.05).

That is a very good result and we can finally conclude to stick with the lag2-differentiation in order to transform our data, since the lag1-differenced series is characterized by too much randomness and white-noise and the lag3- and lag4-differenced series show too large covariances (time-dependence/non-stationarity, indicated by a p-value of the ADF-Test by 0.11 and 0.42, respectively).

Finally we center our transformed data by subtracting the mean so that they are prepared fit and test ARMA-models.

4 Fitting and Testing the Model

4.1 Order and model selection

From the sample ACF and PACF (figure 3.3) of our lag2-differenced series we can see in the ACF-plot a highly significant spike at the first auto-correlation $\hat{\rho}_1$ and the following bars dying out quickly, lying inside 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

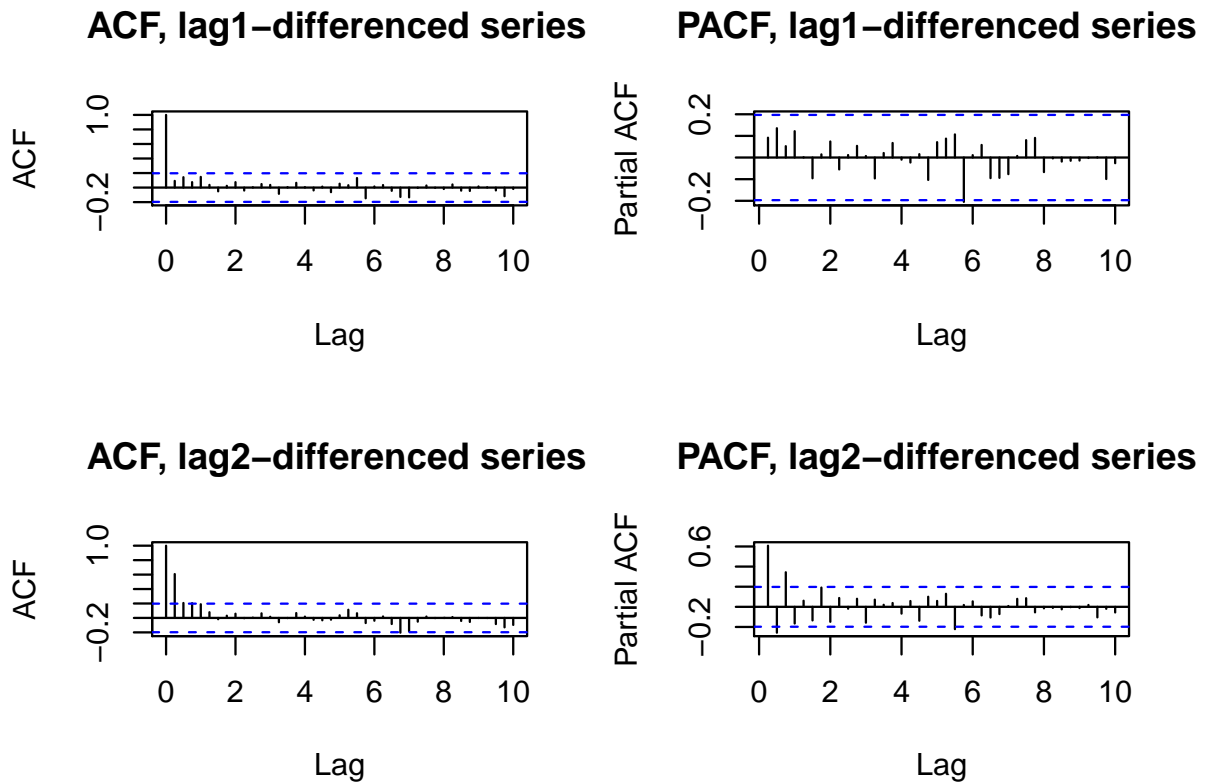


Figure 3.3: sample auto-correlation and partial auto-correlation function of the lag1- and lag2-differenced series

| | q=0 | q=1 | q=2 | q=3 | q=4 | q=5 |
|-----|----------|----------|----------|----------|----------|----------|
| p=0 | 195.0016 | 121.1726 | 121.8342 | 122.9241 | 124.4193 | 124.1333 |
| p=1 | 150.5512 | 121.5382 | 121.4741 | 123.3266 | 125.1512 | 125.9862 |
| p=2 | 145.2901 | 122.3800 | 123.3443 | 125.4699 | 127.3046 | 127.4469 |
| p=3 | 131.5927 | 123.8479 | 125.2583 | 126.7079 | 125.1262 | 125.9619 |
| p=4 | 130.5384 | 124.3852 | 126.3848 | 124.5604 | 125.9526 | 127.9747 |
| p=5 | 132.3732 | 126.3838 | 128.2622 | 126.2500 | 127.8924 | 129.9223 |

Figure 4.1: AICC Matrix

ACF-plot let us suggest an MA(1)-model could fit our data 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 since the bars die out so quickly. Is this the case, then an exponential decrease (in absolute values) in both ACF and PACF plot would rather suggest an ARMA-model.

Hence, according to the ACF and PACF plots it looks primarily like an MA(1)-process, but we cannot reject for sure that it could be an ARMA-process as well, therefore AICC⁴ we compare as well AICC-values for ARMA-models with different orders of p and q⁵ in order to select our model.

From the AICC-Matrix (figure 4.1) we can see that an MA(1)-model gives us the lowest AICC-value which means this model is supposed to perform best. That confirms our guess from the sample ACF/PACF plots. However, we are also going to fit an ARMA(1,1) and an ARMA(1,2) model, because both of them showed as well very low AICC-values (see AICC-Matrix in figure).⁶

⁴ The lower the AICC-value the better the model's performance (Sakamoto et al., 1986).

⁵ p indicates the order of the auto-regressive part and q the order of the moving-average part. Its coefficient are generally noted as ϕ (AR-part) and θ (MA-part).

⁶ Estimations by the `autofit()`-function of the "itsmr" package of R ([**!!!TODO: !!!**]) suggest as well an MA(1)-

We now estimate coefficients ((Brockwell and Davis, 2002, p. 83)) for these three model candidates by the `arima()`-function of the R-package "stats" [**!!!TODO:**].

$$\text{MA(1)-model: } X_t = Z_t + \theta_1 Z_{t-1} \text{ where } Z_t \sim WN(0, \sigma^2)$$

with the estimated coefficient and the corresponding standard error (in brackets):

$$\hat{\theta}_1 = 1.0000(651) \qquad \hat{\sigma}^2 = 0.1847$$

$$\text{ARMA(1,1)-model: } X_t - \phi_1 X_{t-1} = Z_t + \theta_1 Z_{t-1} \qquad \text{where } Z_t \sim WN(0, \sigma^2)$$

with the estimated coefficients and the corresponding standard errors (in brackets)

$$\hat{\phi}_1 = 0.1336(1071) \qquad \hat{\theta}_1 = 0.9786(784) \qquad \hat{\sigma}^2 = 0.1838$$

$$\text{ARMA(1,2)-model: } X_t - \phi_1 X_{t-1} = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} \qquad \text{where } Z_t \sim WN(0, \sigma^2)$$

with the estimated coefficients and the corresponding standard errors (in brackets):

$$\hat{\phi}_1 = 0.7974(2055) \qquad \hat{\theta}_1 = 0.3092(2528) \qquad \hat{\theta}_2 = -0.6685(2583) \qquad \hat{\sigma}^2 = 0.1793$$

First we can see that ARMA(1,1)-model is not supposed to be a good model. Its estimated $\hat{\phi}_1 = 0.13$ is not significant, since its standard error of 0.11 is nearly as high as the coefficient. In contrast, an ARMA(1,2)-model looks quite promising. First, its AR-term would become significant, taking a value of $\hat{\phi}_1 = 0.79$ and second, we can derive evidence for the estimates of the two MA-terms by considering an additional MA-term. An ARMA(1,3)'s $\hat{\theta}_1 = 0.35$ and $\hat{\theta}_2 = -0.60$ would not differ a lot from the ARMA(1,2)'s $\hat{\theta}_1 = 0.31$ and $\hat{\theta}_2 = -0.67$. Furthermore an additional MA-term $\hat{\theta}_3 = 0.05$ is not significant since it would be close to 0 with a large standard error, indicating that the explanatory power by adding a third $\hat{\theta}$ -term would be rather poor, indicating that an ARMA(1,3) is overfitted model. Therefore emphasizing that a moving average of order 2 would be a suitable choice. 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.

The MA(1) is also performing very well, however with a value of 1.00 $\hat{\theta}_1$ implies a unit root on the moving-average polynomial. But that is only a minor issue in comparison to unit roots on the autoregressive polynomial (Brockwell and Davis, 2002, p. 42–83) which is not the case here. Therefore we want to stick with the MA(1) and ARMA(1,2) as our candidate models.

4.2 Diagnostics of the model

To check the validity of our candidate 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 $\hat{\rho}_h$ should be inside the bounds $\pm 1.96/\sqrt{n}$.

Third we use the Ljung-Box Test (Ljung and Box, 1978) to validate or not that the residuals are white-noise by showing p-values for each lag [**!!!TODO: cite page 5 Lecture notes 11**].

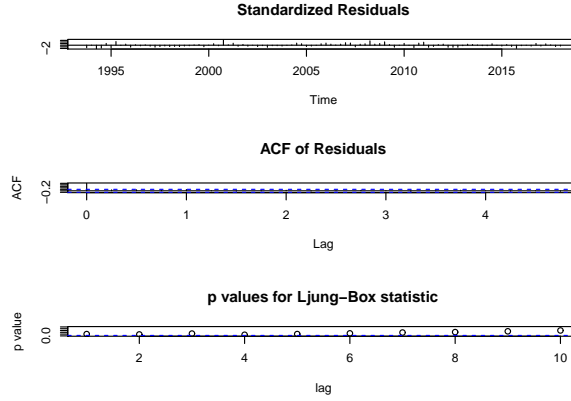


Figure 4.2: MA(1)-Model: Standardized Residuals, Auto-correlation function of the Residuals and the p-values for the Ljung-Box-Test at each lag

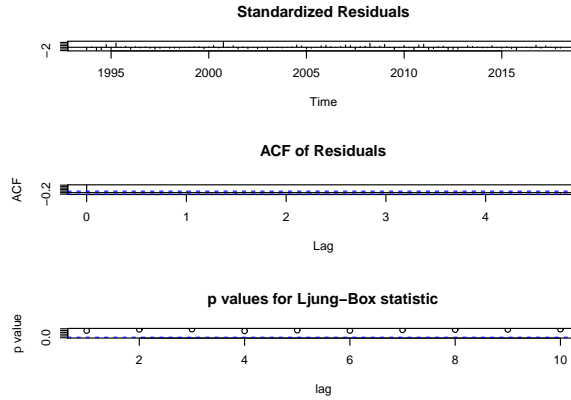


Figure 4.3: ARMA(1,2)-Model: Standardized Residuals, Auto-correlation function of the Residuals and the p-values for the Ljung-Box-Test at each lag

We are visualizing these diagnostics by using the `tsdiag()` function from the "stats" R-Package [**!!!TODO:**]. The following figure 4.2 and figure 4.3 show these three diagnostics for both of our candidate models, the MA(1) and the ARMA(1,2). We can see that the standardized residuals are supposed to be white-noise in both 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 $\pm 1.96/\sqrt{n}$. 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 H_0 that these residuals are a white-noise. So both models MA(1) and ARMA(1,2) would be valid and could be used for forecasting.

5 Prediction of future values

In this section we want to predict future values on the base of our fitted MA(1)- and ARMA(1,2)-models. We will do that by first forecasting our transformed (lag2-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.

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 also suggested by the `autofit()`-function showing very low AICC-values. This can be seen as a succesful robustness check in terms of order selection of p and q .

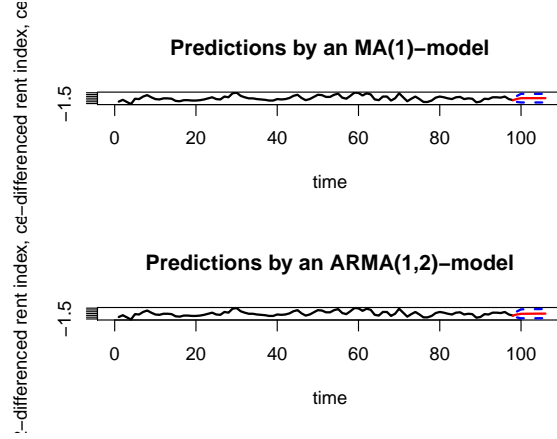


Figure 5.1: Prediction and Confidence Bands for the lag2-differenced Series, MA(1) and ARMA(1,2)

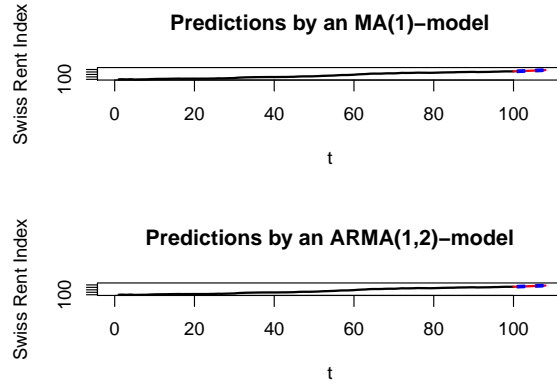


Figure 5.2: Prediction and Confidence Bands for the initial series, MA(1) and ARMA(1,2)

From figure 5.1 we can see the prediction of a process of adaptation to the zero-mean for the next 8 quarters. Both models show a prediction tendency to the zero-mean which is nothing surprising for a mean-centered stationary time series. However, the ARMA(1,2)-model shows a smoother and slower adaptation process to the zero-mean, representing a more reasonable forecast than the MA(1). In comparison the MA(1)-model would only predict for X_{n+1} a value different from the zero-mean. For all the following X_{n+2}, X_{n+3} etc. MA(1) it would abruptly predict zero-mean. Hence, the MA(1)-model's prediction power is not that high and beside the prediction of X_{n+1} , wouldn't even differ from predictions based on a pure white noise sequence. The quite large 95-percent confidence bands are not that surprising in a strongly differenced series as we have here. Differenced series can contain quite large amplitudes as we can.⁷ However, it would be desirable to reduce the variance and by consequence the confidence bands by increasing the number of observations.

Having a look at the predictions for the initial time series⁸ figure 5.2 we can intuitively see that

⁷ The indication of confidence bands is weakly justified here since the series' residuals are successfully tested for normal distribution by the Shapiro-Wilk-Normality Test. [!!!TODO: missing sentence] showed a p-value of 0.12 in case of MA(1) and 0.07 in case of ARMA(1,2), meaning that the hypothesis H_0 of normal distribution can slightly not be rejected (Shapiro and Wilk, 1965). However, a quantile-quantile-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.

⁸The prediction of the initial series bases on retransformation of the lag2-differenced series into the initial series.

the predictions in both models MA(1) and ARMA(1,2) are reasonable, indicating a strongly monotonic trend in the future. However, due to the smoother and slower adaptation process to the expectation of the mean we definitely want to stick with the ARMA(1,2)-model as our favourite model.

6 Discussion

[!!!TODO: Discuss your results, the strengths and the limitations of your own analysis.] The strength of the project lies in the forecast of the ARMA(1,2)-model which gives us a reasonable monotonically increasing prediction which smoothly adapts to the expected mean within the upcoming few quarters.

The difficulty of this project was the problem to get the data transformed into a stationary timeseries in a way that the residuals don't show too much randomness (no white-noise) on the one hand, and on the other hand that we get a few significant auto-correlations $\hat{\rho}_h$ in order to do good predictions based upon past observations.

Since the initial time series was strongly monotonically increasing and therefore likewise strongly positively auto-correlated, achieving a stationary series became a hard task: first, trend elimination by polynomial trend fitting was not possible because of the heavy auto-covariances, second lag1-differencing turned out to be not applicable neither and causing the opposite problem, meaning too much white noise.

Lag2-differencing finally did the job. But that is not a totally satisfying transformation since the series was definitely not seasonal as proven in the section 3, and lag2-differentiation normally is used to get rid of a seasonal component with $d=2$.

Furthermore the MA(1)-model turned out to show a minor issue due to a unit root in the moving-average polynomial (estimated $\hat{\theta}_1 = 1.00$) which could possibly indicate an overdifferencing according to (Brockwell and Davis, 2002, p.194). When a root of the process is close to the unit circle the asymptotic distribution of the maximum likelihood can be quite inadequate (Davidson, 1981). However, other authors have shown that an estimation by maximum likelihood differ only slightly from estimations by other methods as (Davis and Dunsmuir, 1996) have shown and that $\hat{\theta}_1$ close to 1 is not that big of a deal and it is certainly from minor importance in contrast to a unit root on the autoregressive polynomial (Plosser and Schwert, 1977).

To conclude, the choice of the correct data transformation was the dilemma between a lag1-differentiation which would have lead to an undesirable white-noise sequence and a lag4-differentiation which was not anymore stationary (p-value of augmented Dickie-Fuller-Test of 0.42 (editor's comment: alternative hypothesis is stationarity (Dickey and Fuller, 1979))). In between those two extreme options only the lag2-differentiation was stationary on the one hand and allowed for better prediction power thanks to at least one significant autocorrelation.

Furthermore, our model of choice, the ARMA(1,2) process, does in fact not show a unit root on the moving average polynomial (which would not be that severe as we evoked above), but we have to note that its $\hat{\theta}_1$ -coefficient lacks of reliability since its standard error was not that small. The relatively small n of 100 observations was one reason for the relatively high standard error. Hence, for further predictions on the rent index a larger number of observations would be desirable.

7 Conclusion

Both of the finally suggested candidate models, the MA(1)- and the ARMA(1,2)-model turned out to be good models in order to predict our data. Both of them passed the diagnostics of independent residuals very well and both are predicting a reasonable, strongly monotonic trend in the future. However, due to the smoother and slower adaptation process to the expected of the mean we decided to stick with the ARMA(1,2)-model.

This smoothly adapting positive trend is more realistic and goes confirm not only with the long-lasting and consistently increasing trend in the past, but as well with hypothesis about the Swiss rent market development where expert predict a slight slow-down of the growth in rents but a still remaining positive (but slower increasing) trend [**!!!TODO: zitiere nzz oder so**].

Further analysis could be done on the data analysis by treating the two segments before and after the subprime crises (before and after the year 2009) in different ways, since they showed 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.

Further improving work could be done, first of all on the transformation into a stationary time series. Normally, if one uses a lag2-differentiation this is suggesting a seasonal component. This was clearly not the case in our timeseries as tested in section 3. Therefore the transformation process into a stationary time series is not completely satisfying and could be certainly improved by finding other transformation methods, for instance to check as well for the Brockwell-Davis-Method (Brockwell and Davis, 2002).

Second, given the fact that the quite promising MA(1)-model (in terms of the ACF and PACF patterns as well as the lowest AICC) didn't turn out be an ideal choice because of the unit root detected in the moving average polynomial. Even though it is not a severe issue, it could be worth to test drawing inference in another way, meaning that the estimation of the θ -coefficient shouldn't rely only on the maximum likelihood function, but to try alternative estimations which are less vulnerable to the asymptotic distribution of the maximum likelihood, for instance trying a locally best invariant unbiased (LBIU) approach Davis and Song (2011), and afterwards comparing the two different estimates. That could be worth a try, especially if we take into account that a simulated MA(1)-process with $\theta_1 = 1.00$ (see figure B.1 in the [**!!!TODO: apeendix**]) is indicating the model PACF to decrease much slower than the sample PACF, suggesting that a $\theta_1 = 1.00$ is maybe overestimated and alternative estimation methods are justified. A simulated model with $\theta_1 = 0.5$ for instance would look more like the sample ACF/PACF (see figure B.2 in the [**!!!TODO: apeendix**]).

References

- Box, G. and Cox, D. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society*, B 26:211–252.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, 2nd Edition. Springer, New York, NY.
- 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.

- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction, 2nd Edition*. Springer, New York, NY.
- 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.
- Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (for complete samples). *Biometrika*, 52(3/4):591–611.

A Tables

```

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 A.1: summary statistics of the fitted cubic model

B Simulations

C Cross-validation of the lag2-differenced series

Cross-validation with a test data set considering only the last 35 of the 100 observations (see ?? above) confirms furthermore the non-stationary character of the lag2-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)` 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 aswell the data of the first period (1993 to 2009) lag2-differenced series show independent observations as well (see figure C.1), 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.

```

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 A.2: summary statistics of the fitted seasonal model

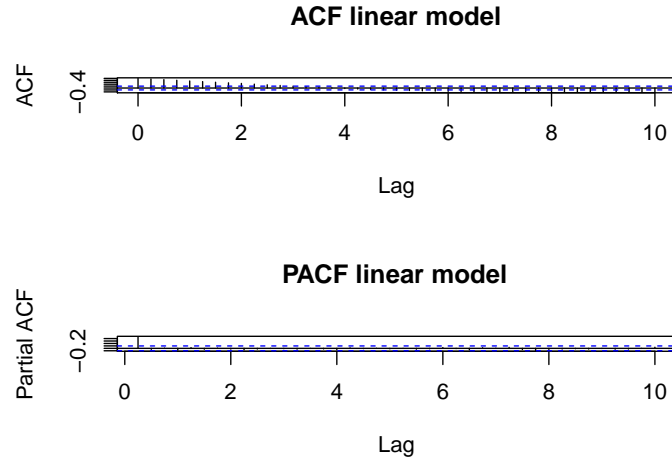


Figure A.3: Sample Autocorrelation and Partial autocorrelation function of the linear model

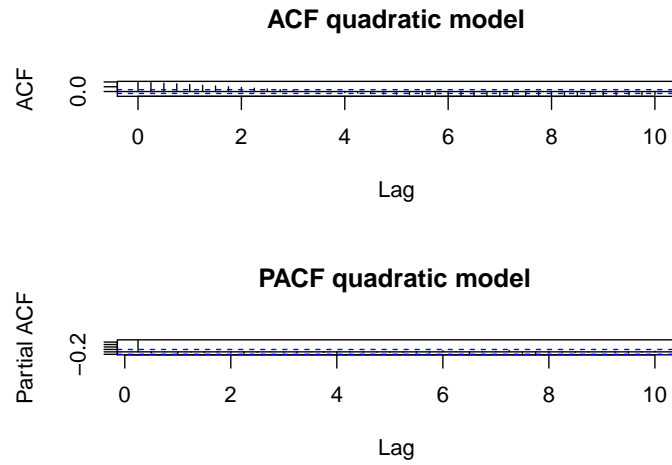


Figure A.4: Sample Autocorrelation and Partial autocorrelation function of the quadratic model

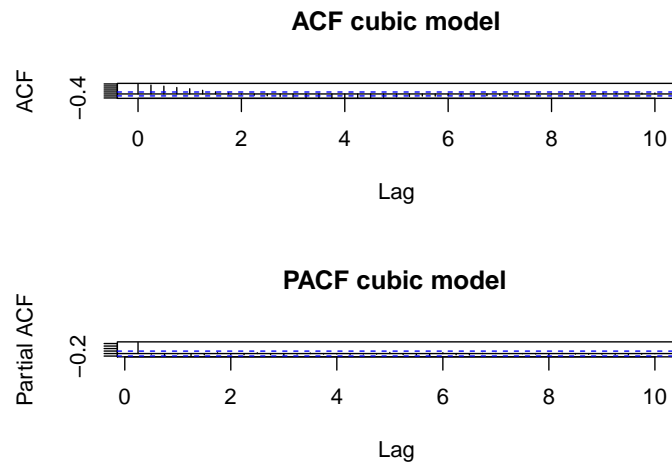


Figure A.5: Sample Autocorrelation and Partial autocorrelation function of the cubic model

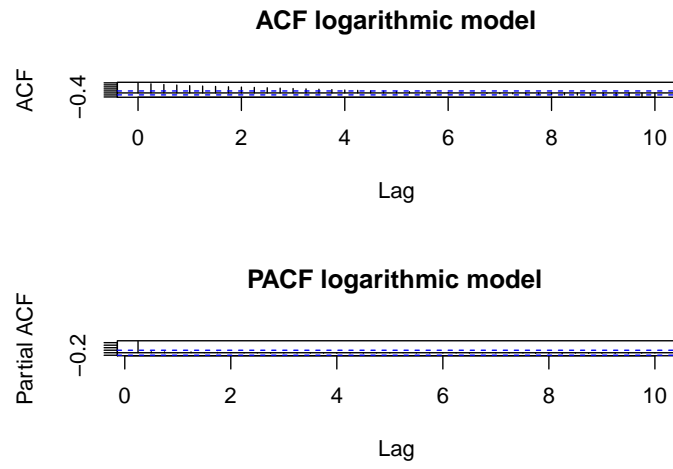


Figure A.6: Sample Autocorrelation and Partial autocorrelation function of the logarithmic model

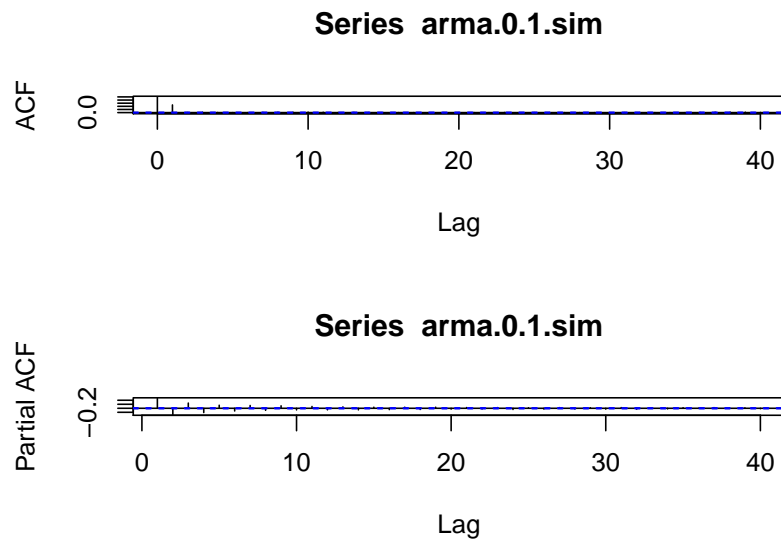


Figure B.1: model ACF and PACF of a simulated MA(1)-process with $\hat{\theta}_1 = 1.00$

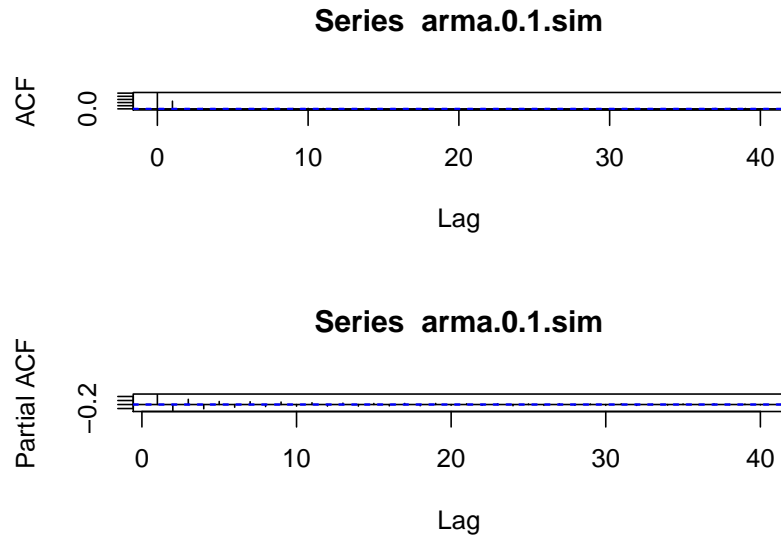


Figure B.2: model ACF and PACF of a simulated MA(1)-process with $\hat{\theta}_1 = 0.5$

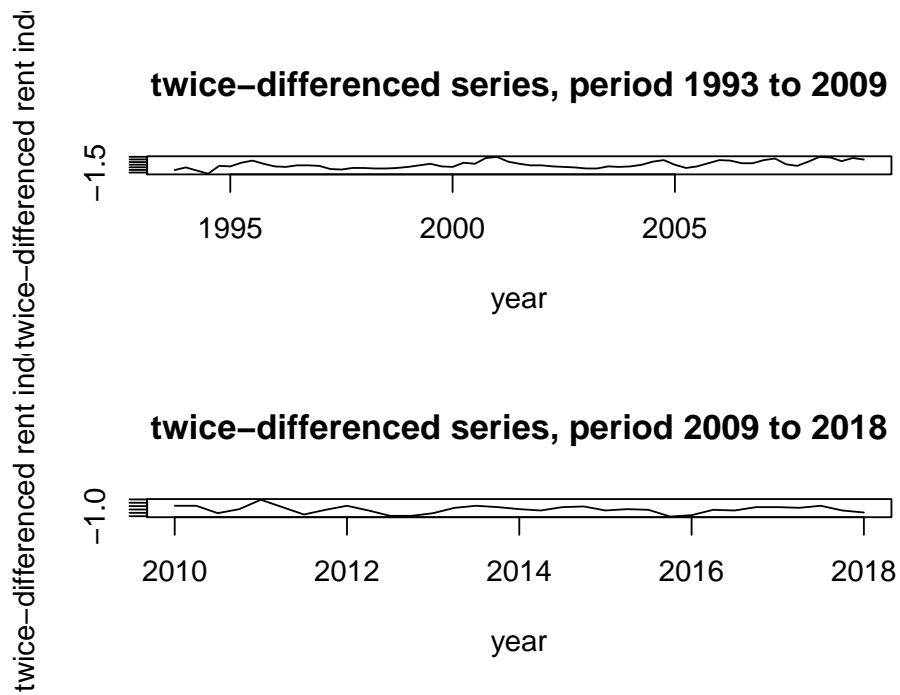


Figure C.1: mean-centered lag2-differenced series, periods from 1993 to 2009 and 2009 to 2018