Justus-Liebig-University Gießen

Department 02: Economics and Business Studies

Working group of VWL XI for Decision Intelligence

Hon. Prof. Dr. Frauke Schleer van-Gellecom

Business Forecasting

Term Paper

submitted by:

Adrian Michael Fischbach

Friedensstraße/30, 35394 Gießen

Adrian.fischbach@wirtschaft.uni-giessen.de

Master of Science

1st semester

18115586

Supervisor:

Hon Prof. Dr. Frauke Schleer van-Gellecom

Gießen, 30/07/2023

Table of content

1 Introduction 1
2 The ARIMA model, the ARIMA(x) model and the Random Walk Model 2
  2.1 The ARIMA model 2
  2.2 The ARIMA(x) model 3
  2.3 The Random Walk model 4
3 Data 4
4 Estimation 5
  4.1 Challenges to the estimation procedure 5
  4.2 Estimating the ARIMA model 6
  4.3 Estimating the ARIMA(x) model 7
5 Forecasting 9
6 Results and discussion 10
7 Conclusion 12

1 Introduction

Forecasting plays a significant role on a microeconomic as well as on the macroeconomic level. On a microeconomic level, business forecasting allows a company to gauge its position on the market, evaluate its strengths but also potential threats and pitfalls. Business forecasting can alleviate a firm to elaborate a well-founded prospective business agenda for the upcoming year. In that regard, it is neither necessary nor possible to make a 100 percent accurate point forecast for the future, which in fact is not the paramount objective. The more relevant task to tackle is to turn the matter of forecasting, typically metaphorically associated with a "black box", i.e., an enigma with many uncertain variables, to a "grey box", i.e., being able to understand, what variables are responsible for the future of business being opaque and uncertain and being able to use these variables more effective and understanding uncertainty[1]. To corroborate this statement empirically, in case study[2] in collaboration with the Justus Liebig University of Gießen, PWC indicates that 92 % of surveyed CEOs consider financial forecasts as critical, however, worrisomely, 59 % of CEOS lament that the data available is too cryptic and incomprehensive. In fact, usually controllers in companies spend 40 % of their time on forecasting. AI based tools could potentially alleviate this workload, transform it to a lean and less time-consuming workflow. According to the PWC Predictive Excellence Study in December 2020, a leading producer of semiconductors could revise his projected revenue to deviate by merely $200 million by using machine learning regression techniques, compared to forecasts with ordinary measures that deviated between $11.2 and $12 bn. On a macroeconomic level, forecasting can be used to craft austerity measures, to implement fiscal incentives wisely with proper allocation of government spending or to finetune business climate indices like the renowned "Geschäftsklimaindex" published by the IFO institute in Munich. Government officials could use enhanced forecasting techniques to gauge prospective unemployment rates and contemplate on tailored measures to incentives the labor market in a prudent way. Business associates and politicians use these reports as a guide for their prospective agenda and to address crucial policies for the future. The aim of this seminar paper is to examine how the

[1] For more information, see "Von der Black Box zur Grey Box": https://www.pwc.de/de/im-fokus/finance-transformation/von-der-black-box-zur-grey-box-vertrauen-in-ki-basierte-forecasts.pdf
[2] Fore more information, see PWC study :https://www.pwc.de/de/im-fokus/finance-transformation/pwc-predictive-excellence-study.pdf

productivity of the American economy can be predicted, using the ARIMA and ARIMA(x) model. Thus, the paramount objective goal is to assess whether a forest with covariates is more precise than a multivariate model. Therefore, initially, the construction of the models is explained. Subsequently, the empirical data is introduced. Afterwards, the models will be estimated, and the forecast will be conducted. The results from the forecast will be presented and discussed. To conclude, final remarks are given, including incentives for improvement.

2 The ARIMA model, the ARIMA(x) model and the Random Walk Model

2.1 The ARIMA model

An ARIMA process can be considered a compound econometric model, since it is the concatenation of two regression processes, i.e., "AR" and "MA". The explanation for the character "I" will follow shortly. An "AR" process describes the regression of an endogenous variable $\{y_t\}$ on its lagged values:

$$y_t = \beta_0 + \sum_{i=1}^{p} \phi y_{t-1} + \cdots + \phi y_{t-p} + \varepsilon \qquad (2.1.1)$$

where $\varepsilon \overset{i.i.d.}{\sim} (0, \sigma 2)$[3] is an error term (white noise) that captures dynamics in a regression model that can't be explained by the dynamics of the model, where $|\phi| < 1$ and β_0 is a constant (Johnston and DiNardo 2007). Furthermore, "I" stands for integrated and indicates that a time series is made stationary, i.e., has no seasonal trend or stochastic process, i.e., the time series fluctuates around a constant mean, thus is constant, has a constant variance over time, i.e., that the variance is invariant of $\{t\}$. Additionally, the covariance does not depend on the time series but rather on the distance between each observation. Ultimately, "MA" stands for movement average and indicates that an endogenous variable $\{y_t\}$ is regressed on its lagged errors, i.e., the average difference between the predicted value and the actual value in the past (Johnston and DiNardo 2007). Algebraically, this can be expressed as follows:

$y = \sum_{i=1}^{q} \theta e_{t-1} + \cdots + \theta e_{t-q}$, where $|\theta| < 1$ and $\varepsilon \overset{i.i.d.}{\sim} (0,\sigma^2)$. Moreover, an ARIMA process needs to be invertible, i.e., an MA(1) process must be convertible to an AR(∞) process and vice versa. The possibility of inverting a time series is conditioned on the

[3] "i.i.d." stands for independent and identically distributed, which is one of the prerequisites for an efficient statistical estimator.

coefficients $\phi_{\mathrm{t-p}}^{\mathrm{n+p}}$ and additionally the sum of the coefficients have to be less than 1. If this holds, it means that more recent observations have a higher weight in the forecasting process than those from the more distant past. This is because the exponent gradually gets larger and consequently, the coefficient gets taken to a greater power the farer the lagged value lies in the past. Eventually, the ARIMA model can be modelled as follows:

$$y_t = \beta_0 + \sum_{i=1}^{p} \phi y_{t-1} + \cdots + \phi y_{t-p} + \sum_{i=1}^{q} \theta e_{t-1} + \cdots + \theta e_{t-q} + \varepsilon \qquad (2.1.2)$$

Moreover, the following notation needs to be added to the model: ARIMA (p,d,q). The characters in the brackets indicate the rank of the different parameters of the ARIMA model, where p determines how many lagged values of the endogenous variable are most suitable for the model, thus referring to the AR process (Ord et al. 2017). Subsequently, parameter *d* indicates how often the time series has been differenced to obtain a stationary process and thereby is an indicator of the complexity of the time series. Finally, *q* represents of which order the MA-process is, i.e., or how many lagged errors are significant (Ord et al. 2017). The ARIMA model can also be extended to a SARIMA model, where "S" stands for seasonal. To specify a SARIMA model, a second bracket with three parameters, i.e., SARIMA (p,d,q) $(P,D,Q)_m$ needs to be added, where the parameters in the second bracket are analogous to the parameters in the first bracket, the only difference being that the parameters in the bracket, denoted with m, exclusively pertain to seasonal dynamics. For instance, a parameter rank of P=2 indicates that the 24th lag is considered, given that the frequency of the data is monthly. To judge whether a SARIMA model is more useful than an ordinary ARIMA model, one can decompose the time series into its different layers, e.g., white noise, trend, and seasonality. If the seasonal component shows a distinct pattern, choosing the SARIMA model over the ARIMA might be reasonable, which can be determined through criteria in the estimation process of the model.

2.2 The ARIMA(x) model

An ARIMA(x) process is complementary to an ARIMA process in that regard that Eq. (2.1.1) is supplied with an exogenous variable:

$$y_t = \beta_0 + \boldsymbol{x_t} + \sum_{i=1}^{p} \phi y_{t-1} + \cdots + \phi y_{t-p} + \sum_{i=1}^{q} \theta e_{t-1} + \cdots + \theta e_{t-q} + \varepsilon \; (2.2.3)$$

Thus, $\{y_t\}$ and its future values $\{y_{t+1}+\ldots y_{t+T}\}$ also depend on $\{x_t\}$and its future values $\{x_{t+1}+\ldots+x_{t+T}\}$, besides its own lagged values and errors.

2.3 The Random Walk model

The random walk model or naïve model states that $y_t = \phi y_{t-1} + \varepsilon$, where $|\phi| = 1$, i.e., the model has a unit root and therefore follows a stochastic process (Ord et al. 2017). Specific diagnostic tests are needed to distinguish a stationary process from a stochastic process, because it's not always easy to judge at first sight. Firstly, a unit root does not necessarily violate the first prerequisite for a stationary process of having a constant mean over time, but in a subsequent step clearly violates the rule of a constant variance over time. The naïve model can further be extended to include a constant, i.e., $y_t = \beta_0 + \phi y_{t-1} + \varepsilon$. The constant gives the random walk a more deterministic trend, a drift, and therefore is also called the random walk with drift.

3 Data

For the empirical part, the U.S. Retail sales[4] from 1992 – 2022 where chosen. It is used by the Bureau of Economic Analysis as an input for measuring Gross Domestic Product. Additionally, it is used by the Federal Reserve to make more tangible assumptions about economic trends and to deliberate on reasonable monetary policy practices. Ultimately, it is used by the Council of Economic Advisors to make economic policy analysis estimates. Additionally, the U.S. economy is of particular interest, since it tenaciously defends its hegemonial position as the world's strongest economy, which can be derived from several factors. For instance, amongst other factors, it's domestic currency, the U.S. dollar, is the world most widely deployed currency to conduct transactions. Forecasting retail sales was of major interest especially due to its vastly changing and innovative nature and the abundance in exogenous factors that can influence the forecasting accuracy. Furthermore, accurate forecasting in the retail sector can not only potentially lead to substantial cost reduction by forecasting supply chains and needed labor force but also enhance customer satisfaction (Petropolous et al. 2022). Especially when dealing with perishable goods in the retail sector, both, too excessive and too few stocks should be avoided as it results in too much waste or, in the latter case, in supply shortages (Petropolous et al.

[4] For more information on the U.S. Retail Sales sector, see United States Census Bureau: https://www.census.gov/retail/marts/about_the_surveys.html

2022). The frequency for the Retail sales data is monthly. This frequency is preferred over daily data and quarterly data. On the one hand, daily data is the most frequent and thereby most likely produces heteroskedastic residuals. On the other hand, quarterly data, due to the long-time span between observations, lacks in detail (Lütkepohl and Krätzig 2004). For the sake of making the forecasts of the retail sales more understandable, the retail sales from 1992 until 2022 are adjusted for inflation, based on the year 2022 (Ord et al. 2017; Lütkepohl and Krätzig 2004). To adjust for inflation, the modus procedendi is the following: The Consumer Price Index (CPI) of each year was obtained from the Federal Reserve St. Louis database. Afterwards, to adjust each month for inflation, the arithmetical mean of the Consumer Price Index (CPI) in the year 2022 was computed and every monthly CPI was divided by the arithmetical mean of 2022. Hence, for every month of retail sales the price index, relative to the base year is obtained and the retail sales of each month are divided by its respective index. Thereby, real sales figures are obtained. The same modification is applied to potential exogenous variables that are indicated in nominal dollars.

4 Estimation

4.1 Challenges to the estimation procedure

Especially considering the ARIMA(x) model and the deployment of exogenous variables, it is important to choose the most parsimonious model (Ziel and Weron 2018; Box et. al 2008). The intention of the parsimonious model is to be able to explain the influence of exogenous variables with as few regressors as possible and to determine the most efficient trade-off between the goodness of fit and parsimony. For instance, if two models with different amount of regressors contribute the same percentage of explicative variance (R^2), for the sake of simplicity, the simpler model is elected. This is an important aspect in the context of training data and testing data. If the training data is fitted with abundant exogenous variables, it might very precisely predict the current data set, however, if this model is applied to a new data set, it might lose a lot of estimation accuracy, due to the fact the model is not suitable for the new set of data. This predicament is also called overfitting (Box et al. 2008).

4.2 Estimating the ARIMA model

Estimating the ARIMA model means determining the rank of parameters p, d and q. This is crucial in that respect to avoid misspecification of the model. If there are flaws in the specification, e.g., neglecting significant lags, these dynamics are relocated to the error term and will bias further estimations. Suppose for example that the erroneously specified model is the following: $y_t = \beta_0 + \beta_1 x_1 + \beta_{x-1} + \varepsilon_t$. However, due to misspecification, the error term consists of the following parameters:

$\varepsilon_t = \beta_2 x_{t-2} + \eta_t$. Thus, $\varepsilon \overset{i.i.d.}{\sim} (0,\sigma^2)$ does not apply anymore, i.e., $\mathbb{E}[\varepsilon_t | x_1] \neq 0$ (Johnston and DiNardo 2007). Obviously, in any time series model, some existing autocorrelation is inevitable. Ultimately, the model might lack in interpretation credibility. Therefore, the ARIMA model needs to be differenced in an initial step. Differencing ensures that the model is stationary. Stationarity is crucial in that respect if the mean and the variance are not constant, then each sample data taken from the underlying population will exhibit different values. This inhibits the possibility of making statistical inference about the estimators (Ord et al. 2017). Especially in the univariate ARIMA model, where an independent variable is regressed against its own past that is not stationary, the regression is likely to produce a high (R^2). Nonetheless, this high (R^2) results from a spurious regression, meaning that it rather results from both, the endogenous variable and its regressors, moving in the same direction and not due to causal effects. After differencing the time series, the ARIMA model is plotted in the Autocorrelation Function and in the Partial Autocorrelation function. Judged upon how many lags lie outside the confidence bands, i.e., which are significant, the AR(p) order and the MA(p) order will be determined. The Autocorrelation Function determines the rank of the MA(p) process, the Partial Autocorrelation Function respectively the AR(p) process. Judging visually, the lags for the AR(p) model and the MA(p) model are selected and both are separately chosen. The manually chosen models can be complemented by the automatically generated Arima model. This model is generated based on the Hyndman-Khandakar algorithm. The three models are compared based on the Aikaike Information Criterion (AIC) (Ord et al. 2017). The model with the lowest AIC value is chosen. Additionally, based on the best model chosen, the residuals can be plotted on the Autocorrelation Function and the histogram. If the correlogram of the ACF does not display any significant spikes outside the confidence bands and the histogram indicates that the residuals follow a bell-shaped curve, this model can be considered suitable. If there are still significant spikes left although the times series has been differenced, one can conduct diagnostic tests on the residuals, e.g., the Ljung Box test/Portmanteau test, where (H_0) states that all unexplainable dynamics of the model are captured by the white noise term. If the Portmanteu test is not significant, the H_0 can't be rejected, and one can infer that the spikes are significant but do pertain to dynamics captured by the white noise term and don't cause autocorrelation. Furthermore, not rejecting the (H_0) Hypothesis of the Ljung-Box Test is an important prerequisite of conducting an efficient k-fold-Cross validation in the forecasting process (Bergmeir et al. 2017: 14). The estimations are depicted in Table (1). The ARIMA(0,1,2)(0,0,1) and the ARIMA(1,1,0)(2,0,0) model where selected manually upon visual inspection of the ACF and the PACF. The search model was selected by the Hyndman-Khandakar algorithm. Approximation was turned off and the algorithm was allowed to search for very big models, i.e., the constraint for the p, d, P and Q parameters was set ≥ 9. As it turns out, this model appears to have the lowest AIC value and there is no autocorrelation in the residuals.

Table 1 ARIMA model estimation

Model	AIC	Ljung-Box Test
ARIMA (0,1,2)(0,0,1)	-1292	0.0759
ARIMA(1,1,0))(2,0,0)	-1290	0.0004***
stepwise (1,1,2)(0,0,2)	9122	0.00495***
search (2,1,3)(0,0,1)	-1326	0.49

The asterisks indicate at which confidence interval the (H_0) of the Ljung Box test of no autocorrelation of the residuals can be rejected, i.e., (*) = 90%, (**) = 95% and (***) = 99%.

4.3 Estimating the ARIMA(x) model

After having selected the best ARIMA model, the ARIMA(x) model is built on top of the pre-selected ARIMA model. There are multiple ways to determine the most parsimonious model. One can test several models, including multiple regressors and then determining, based on the lowest AIC or BIC value, which model to choose. A common way to eliminate superfluous regressors that might cause a spurious regression due to multicollinearity among the exogenous variables is applying a LASSO[5] regression (Ord et. al 2017). This type of regression penalizes the presence of certain variables that are suspected to cause overfitting by shrinking their

[5] LASSO stands for Least Absolute Shrinkage Selection Operator

coefficients and thereby minimizing their weight on the model (Ord et al. 2017). For this paper, however, for the sake of simplicity, merely one covariate is applied and chosen by its most significant causal impact on the endogenous variable and thereby follows the methodology of Anggraeini et al. (2017: 194). Thus, to find the most suitable exogenous variable, different Vector Autoregressive (VAR) models are built, in which the endogenous variable for every VAR model, is regressed on another exogenous variable. A causality test, including the granger causality test and the instantaneous causality test, is performed on each VAR model to assess, how significantly the exogenous variable granger causes and instantaneously causes the endogenous variable. According to Kirchgässner et. al (2017), granger causality tests whether the forecast error when using the current and past values of both, the exogenous and endogenous variable, is smaller than merely using the current and lagged values of the endogenous variable (Kirchgässner et al. 2017). The instantaneous causality test is an extension of the granger causality test. It additionally imposes the caveat that if the forecast error is smaller when knowing future values of the covariate and the endogenous variable compared to merely knowing future values of the endogenous variable, the covariate significantly instantaneously causes the endogenous variable (Kirchgässner et. al 2017). Thus, in the case of choosing a suitable regressor for the ARIMA(x) model, the criterion for instantaneous causality has higher relevance than the granger causality criterion. This process is iterated for every possible combination of the endogenous variable with all the exogenous variables at hand, until the exogenous variable that most significantly instantaneously causes the endogenous is obtained. This exogenous variable is used to build the ARIMA(x) model. The exogenous variable that most instantaneously caused Retail sales was the Personal Income variable. Other exogenous variables of interest and their causal impact are listed in Table (2). Another idea was to deploy government spending as an exogenous variable. This was of particular interest because, unlike, for instance, Germany, the USA does not have a well-established social welfare system and instead a liberal market economy, where government incentives are transferred to the population more instantaneously. Unfortunately, the respective data was only available on a quarterly basis. To align the data with the Retail Sales data, there would have been the necessity for interpolation which can create major errors and lead to information loss (Lütkepohl and Krätzing 2004: 4). Consequently, it was refrained from using this variable.

Table 2 Causality Test of exogenous variables

Variable	Granger causality	Instantaneous causality
Consumer confidence	0.735	0.687
Consumer Sentiment	0.561	0.765
Expected Inflation	0.0036***	0.004***
Consumer Price Index	0.0005***	0.016***
Personal Consumption Expenditure	0.873	0.757
Personal Income	0.836	0,000***

The asterisks indicate at which confidence interval the (H_0) of no granger causality or no instantaneous causality can be rejected, i.e., (*) = 90%, (**) = 95% and (***) = 99%.

5 Forecasting

To forecast the time series the cross-validation method is used. This methodology gained in popularity in the field of econometric analysis, especially due to the empirical contributions on time series and economic forecasting of the Australian statistician Robin J. Hyndman. According to Hyndman et al. (2017), cross-validation is a method of using available time series data the most parsimonious way without compromising on forecasting accuracy (Hyndman et al. 2017). The data is split into a training and set and a test set. The training set is furthermore partitioned into an arbitrary number (*k*) of sub-samples, called "folds" (Hyndman et al. 2017: 9). This is why this method is also called *k*-fold cross-validation. The pre-estimated model is then trained and evaluated on each fold, thereby through all iteration k-1 folds are used to train the model. The last fold is then used to apply the trained model in the test set. The performance of the model on the testing data is then scrutinized by accuracy metrics such as the Root Mean Squared Standardized Error (RMSSE), the Mean Absolute Percentage Error (MAPE) or Theil's U. Furthermore, a common cross-validation technique is the rolling forecasting technique, i.e., the training set has a rolling origin and with every one-step-ahead forecast, the origin of the training set in moving forward (Hyndman et al. 2017). However, in this forecast, the cross validation with a fixed origin was used with the training set growing with each forecasted observation. Finally, the evaluation of the forecasts is obtained in the following table:

Table 3Evaluation of the forecast

Model	Horizon $_{t+T}$	RMSSE	MAPE	Theil's U
ARIMA	1	0.159	2.99	0.029
ARIMAX	1	0.081	1.17	0.0117
Random Walk	1	0.168	3.6	0.036
ARIMA	2	0.219	4.12	0.0412
ARIMAX	2	0.092	1.35	0.0135
Random Walk	2	0.228	4.82	0.0482
ARIMA	3	0.278	5.23	0.0523
ARIMAX	3	0.106	1.56	0.0156
Random Walk	3	0.286	5.91	0.0591
ARIMA	4	0.333	6.31	0.0631
ARIMAX	4	0.111	1.66	0.0166
Random Walk	4	0.339	6.96	0.0696

6 Results and discussion

From Table (3) the results of the forecast can be derived. A forecasting model is considered as precise when the MAPE is under 5% and the Theil's U coefficient is close to zero (Theil 1971) which can be observed for all models until the forecasting horizon $\{t_{t+2}\}$. The forecasts of the ARIMA model start getting less accurate after $\{t_{t+2}\}$ as its associated MAPE value surpasses the threshold of 5%. However, it seems to be a good model compared to the Random Walk model when comparing all accuracy metrics. In general, it can be inferred from the forecasting results that the ARIMA(x) model indeed seems to be more accurate than the ordinary ARIMA model. A further step could be to perform an Ensemble forecast, e.g., combining the ARIMA and the ARIMA(x) model and taking the arithmetical mean and declaring a new model. This ensemble forecast could be further complemented by an Exponential Smoothing model (ETS) (Hyndman and Athanasopolous 2018). There exist other ordinary accuracy metrics, such as the Mean Absolute Error (MAE) or the Mean Absolute Deviation. These where left out on purpose, because unlike the ordinary accuracy metrics, the MAPE is a scale independent measure, which makes the comparison of economic quantities measured in different units (e.g., absolute quantities, growth rates) more feasible. However, MAPE also has its drawback. Firstly, it is an asymmetric measure, described by the following equation:

$MAPE = \frac{\sum_i^n |\left(\frac{e_i}{y_i}\right) * 100}{n}$ (3.1.1). Given eq. (3.1.1), one observes two components, i.e., $\{e_i\}$ and $\{y_i\}$. A forecasting model cannot keep up with the actual observed values, resulting in positive forecasting errors $\{e_i\}$ and negative forecasting errors $\{y_i\}$, where positive forecast errors, i.e., under forecasts, indicate an upward trend and vice versa. Given Eq. (3.1.1), one can see that the positive forecast errors are being divided by the negative forecast errors. However, a time series does not necessarily have equally significant upward and downward trends. In other words, MAPE penalizes negative errors heavier. The percentage error cannot exceed 100% for forecasts that are too low, whereas there is no upper limit for the forecasts which are too high. Conclusively, MAPE, by construction, favors models that under-forecast over those that over-forecast. Therefore, MAPE is an asymmetric measure (Goodwin and Lawton 1999). Another drawback of MAPE is the composition of the denominator, which is solely consists of $\{n\}$ the number of observations. If the number of observations is zero, MAPE is simply undefined. In case that the number of observations is asymptotically close to zero, MAPE is heavily skewed and can be infinitely high. Symmetrical MAPE, or sMAPE, compensates for the drawbacks that MAPE poses, especially the drawback of favoring models that under-forecast (Goodwin and Lawton 1999). Symmetrical MAPE does that by setting upper and lower boundaries, typically of 200%. Despite having observed that the trained ARIMA(x) in this case study seems to outperform the ordinary ARIMA model, there is still a major drawback to this model. According to Hyndman and Athanasopolous (2017), given the Eq. (2.2.3), unlike in an ordinary multivariate linear regression where the coefficients of the regressors can directly be interpreted as the elasticities of the endogenous variables following a change of the exogenous variable, the interpretation in the ARIMA(x) regression is not that straightforward (Hyndman and Athanasopolous 2017). The elasticities in this model are conditional on the coefficients of the lagged values of the endogenous variable which makes the interpretation less intuitive than in an ordinary multiple linear regression. Furthermore, the ARIMA(x) model is restricted in the sense that only values of the exogenous variable, e.g., observations from $\{x_{t,}\}$ until $\{t_{t+T}\}$ are regressed on the endogenous variable $\{y_t\}$ until $\{y_{t+T}\}$. Thus, only present, and future values are considered. The Autoregressive Distributed Lag Model (ARDL) can be considered as a possible extension of the ARIMA(x) model as it incorporates the features of a multivariate VAR(p) model of order (p) into the ARIMA model. Hence, future observations of the endogenous variable are regressed on values of $\{x_t\}$ until $\{x_{t-p}\}$. One can write the ARDL model in its simplest form:

$$y_t = m + \alpha_1 y_{t-1} + \beta_0 x_1 + \beta_1 x_{t-1} + \alpha_p y_{t-p} + \beta_p x_{t-p} + \varepsilon_t \qquad (7.1.)$$

Applied to this case study, one would not merely predict the future values of the exogenous variable as trivially as was done in the process, but the lags of the time series and the orders of the parameters would be as thoroughly selected as for the endogenous variable. Applying an Autoregressive Distributed Lag Model would have gone beyond the academic scope of this seminar paper, however, in certain economic research fields, for instance, the effect of monetary policy regimes, i.e., where a shock of one exogenous variable does not have an immediate effect on the endogenous variable, it would be more suitable than a simpler ARIMA(x) model. Conclusively, by forecasting an endogenous value that is regressed on an exogenous value, solely observed at the same period, potential crucial impacts of lagged exogenous observations are neglected.

7 Conclusion

As learned throughout this seminar paper, the necessity of business forecasting successively increases in demand. However, there are distinct challenges to the forecasting process and its efficacy to make it genuinely operatable. One major challenge, that becomes evident in Chapter 2 where the model is estimated. Recognizing seasonal patterns and thereby determining the proper lag length is a very challenging aspect, especially in the field of predictive analytics, where only a well estimated model can yield a reliable and accurate forecast. Moreover, providing an accurate long-term forecast can be very challenging. This becomes evident empirically by looking at Table (3), where the RMSSE gradually increases with the forecasting horizon going farther into the future. Besides technical encumbrances that business forecasting entails, one must bear in mind the financial aspect as well. Since an accurate and well fitted forecasting model strongly depends on complete and good quality data, procuring the data can be a considerable cost factor, depending on the specific needs. Analogous to Enterprise Resource Planning systems, such as provided by SAP, companies such as Refinitiv Inc. or Oxford Economics GmbH (ltd.) procure and process economic data and sell them to banks, investment companies and other institutions. These licenses and services oftentimes are very expensive and especially small and midcap companies cannot afford them. Moreover, especially stressed by Lütkepohl and Krätzig (2004), data provided is often processed a priori, e.g., seasonal adjustments are made. For these adjustments many methods exist, and it can be cumbersome to comprehend which exact method was applied (Lütkepohl and Krätzig 2004: 4). Additionally, it is worth arguing if artificial forecasting tools are capable of forecasting unforeseen shocks, e.g., "black swan" events. It might be possible to deploy binary coded dummy variables that are in fact already deployed for measuring the effect of holidays or to capture seasonal patterns. However, computing such variables to improve forecasting remains a challenging task. The paramount aim of artificial, data driven, and sober forecasting is to compensate for the drawbacks of judgmental and emotionally driven forecasting. Nevertheless, in economic environments that are inherently uncertain and volatile, judgmental forecasts might be indispensable, and a prudent blend of both forecasting approaches should be considered. Since the estimation of a forecasting model based on historical data belongs to the field of predictive analytics, a further step would be the prescriptive analytics, where, from a business perspective, a company has the opportunity, based on complex and sophisticated algorithms, optimization models and simulations, e.g., Monte Carlo simulations, to use the privilege of gained insights about future economic environment climate to help suggest various courses of action to be well prepared for future challenges, e.g., by allocating financial resources properly.

References

Anggraeni, W. and K. B. Andri and Sumryanto and F. Mahananto (2017): "The Performance of ARIMAX Model and Vector Autoregressive (VAR) Model in Forecasting Strategic Commodity Price in Indonesia", in Procedia Computer Science, Vol. 124, pp. 189-196.

Bergmeir, C. and R.J. Hyndman and B. Koo (2015): "A Note on the Validity of Cross-Validation for Evaluating Autoregressive Time Series Prediction", in: Computational Statistics and Data Analysis, Vol. 120 (2018), pp. 70-83.

Du Preez, J. and S. F. Witt (2003): „Univariate versus multivariate time series forecasting: an application to international tourism demand", in: International Journal of Forecasting, Vol. 19, No. 3, pp. 435-451.

Goodwin, P. and R. Lawton (1999): "On the asymmetry of the symmetric MAPE", in: International Journal of Forecasting, Vol. 15 (1999), No. 4, pp. 405-408.

Hyndman, R. J. and G. Athanasopoulos (2018). Forecasting: principles and practice. Publisher: OTexts, 2nd edition.

Johnston, J. and J. DiNardo (2007). Econometric Methods. Publisher: McGraw-Hill, 4th edition.

Kirchgässner, G. and J. Wolters (2007). Introduction to Modern Time Series Analysis. Springer.

Lütkepohl, H. and M. Krätzig (2004). Applied Time Series Econometrics. Cambridge University Press.

Ord, K. and R. Fildes and N. Kourentzes (2017). Business Forecasting. Publisher: Wessex Press, Inc., 2nd edition.

Ramos, P. and N. Santos and R. Rebelo (2015): "Performance of state space and ARIMA models for consumer retail sales forecasting", in: Robotics and Computer-Integrated Manufacturing, Vol. 34 (2015), pp. 151-163.

Ziel, F. and R. Weron (2018): "Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks", in: Energy Economics, Vol. 70 (2018), pp. 396-420.