



universität
wien

MAGISTERARBEIT

Titel der Magisterarbeit

“Health Expenditure Forecasting and Technological Change”

Verfasst von

Frank Kronemann, Bakk.rer.soc.oec.

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften
(Mag.rer.soc.oec.)

Wien, 2015

Studienkennzahl lt. Studienblatt:

A 066 913

Studienrichtung lt. Studienblatt:

Magisterstudium Volkswirtschaftslehre

Betreuer

Univ.-Prof. Dipl.-Ing. Dr. Robert Kunst

Health Expenditure Forecasting and Technological Change

Frank Kronemann

Abstract

Forecasting future health expenditure is one important issue that is beyond strategic and planning decision in health policy making. There are certain variables which affect health expenditures. The rising share of older people (*age65*) and the role of technology are seen as key drivers of the continually increasing health expenditures. Therefore this thesis compares two different model approaches with and without exogenous variables and with regard to the forecasting ability. The analyzed countries – Austria and Germany – show similar results with both having a bigger increase in health expenditures caused by the exogenous variable *age65*, while the size of the technology proxy (data for medical devices and pharmaceutical products) is negligible. In contrast, the ARIMA model without an exogenous variable shows a more linear increase for both countries. However, the results of both model estimations predict an increase of per capita health expenditure in Austria by 24.9% in 2020 and by 72.0% in 2030, compared to data from 2012. Health expenditure in Germany will increase by 25.3% and by 76.7%, respectively.

Acknowledgments

Many thanks go to my advisor, Prof. Robert Kunst, for his help and very good guidance over all the progress of work.

Furthermore I would like to express my gratitude to my fellow graduate students at the university and to all my other friends. Moreover I would also like to thank the members of the HealthEcon team for all their advice and help.

A special thanks to my family for supporting me during the whole course of my study. Finally I want to thank Christine for her patience, moral support, and constant encouragement. Words cannot express how grateful I am.

Contents

List of Tables and Figures

1. Introduction	1
2. Materials and Methods	2
2.1. Data	3
2.2. Model	11
2.2.1. Austria – method A	11
2.2.2. Austria – method B	14
2.2.3. Germany – method A	17
2.2.4. Germany – method B	19
3. Forecast	21
3.1. Assumptions	21
3.1.1. Austria.....	21
3.1.2. Germany.....	22
3.2. Results	23
3.2.1. Austria.....	23
3.2.2. Germany.....	25
3.2.3. Evaluating Forecasts	27
4. Conclusion	30
Literature	32
Appendix	35
German Abstract	39
Curriculum Vitae	40

List of Tables and Figures

Table 1: Reform-dummies; Austria and Germany	7
Table 2: Model Variants Austria	13
Table 3: Dickey-Fuller test on first-differenced PCHE	15
Table 4: AIC and BIC for first-differenced PCHE	16
Table 5: Resulting AIC values for combinations of AR and MA (Austria)	16
Table 6: AR(1) model Austria	17
Table 7: Selection of Estimations for Model Method A – Germany	19
Table 8: Resulting AIC values for combinations of AR and MA (Germany)	20
Table 9: MAE and MAPE for pseudo out-of-sample period forecasts (2009 to 2012)	29
Table 10: RMSE and TIE for pseudo out-of-sample period forecasts (2009 to 2012)	29
Table 11: Identification of reforms in Austria regarding health expenditure	35
Table 12: Identification of reforms in Germany regarding health expenditure	36
Table 13: Correlation of Variables (Austria)	37
Figure 1: Classification of factors underlying developments in health care expenditure	2
Figure 2: Public Current Health Expenditures, pc, EUR, PPP2005	4
Figure 3: Smoothed/Weighted deficit and original deficit for Germany (1990 to 2012)	6
Figure 4: New medical products by FDA, 1976-2012 (1976=1)	9
Figure 5: New medical products by FDA in capital stocks (5% depreciation), 1976-2012 (1976=1)	10
Figure 6: Demographical Change, Austria	21
Figure 7: PCHE and its predictions, Austria	24
Figure 8: Forecast Germany method A	25
Figure 9: Forecast Germany method B	26
Figure 10: PCHE and its predictions, Germany	26
Figure 11: Forecast 2010-2012 (sample: 1976-2009), Austria	38
Figure 12: Forecast 2010-2012 (sample: 1992-2009), Germany	38

1. Introduction

Rising Health Care Expenditure and Long-term Care (LTC) have gained more and more priority for political decision makers over the last decades. Health Care Expenditure holds a grand share of public expenditure in all OECD countries and about 12% of GDP in Austria, putting it behind the Netherlands as the country with the second highest health expenditure per capita in 2012, followed by Germany and Denmark (OECD 2014). Predictions of OECD and the Economic Policy Committee (EPC) in the last few years emphasized the rising pressure of future costs in this area. The considerable change in population structure with a high share of older people is occurring due to major increase in life expectancy and a falling birth rate. This phenomenon can not only be observed in Austria but in the entire European Union and constitutes an ambitious challenge for policy makers.

A large part of this thesis' methodology is based on the model of the Ageing Working Group (AWG) – mandated by the EPC – from 2009. Additionally, focus is on improving the accuracy of the prediction using the Willemé/Dumont (2015) method.

Effects on health expenditure are divided into demographical and non-demographical effects. Non-demographical effects however are not fully identified, and a major part of growth in public health expenditure is, among other reasons, technology-declared. Even though this fact is difficult to quantify. Since Newhouse (1992) technology is broadly seen as a main driver regarding expenditures in the medical sector. Given that there is no firm alternative, the most commonly used approach to involve this effect is the total-factor productivity, also called residual approach. The importance of this standard approach is also emphasized by having a lower assumption on income-elasticity than 15 years ago. The higher this elasticity the better demographical and non-demographical/income effects describe the expenditure growth (OECD 2013). Willemé/Dumont (2015) try data based on the number of approved medical devices and pharmaceutical products from the U.S. Federal Drug Admission (FDA) as a technology-proxy and produce surprisingly good results compared to the standard approach. The following thesis will use this approach to estimate a health expenditure forecast for Austria and Germany.

2. Materials and Methods

The empirical literature on health expenditure determinants presents a variety of influencing factors responsible for change in health expenditure. In contrast to the spending on pensions, which are mainly determined by the pension system and demographic factors, health care expenditures depend not only on demographic but on social, economic, political and technological determinants. The literature offers several views on this topic (e.g., Gerdtham and Johnsson 2000, Chernew and Newhouse 2012). Przywara (2010) provides a short overview of factors underlying developments in health care in Figure 1. Specifically the figure divides factors not only into four different groups but also distinguishes between factors affecting the demand or supply side of health care provision.

Figure 1: Classification of factors underlying developments in health care expenditure

	Demographic factors	Health factors	Economic and social factors	Public policy factors
Demand side factors	• Size and structure of the population	• Health status of the population, in particular of elderly cohorts • Death-related costs	• National/ individual income • Income elasticity of demand for health care • Social determinants of health (environment, living conditions) and health-related behaviour • Public expectations and real convergence in living standards	• Health promotion and disease prevention policy
Supply side factors			• Development of new technologies and medical progress • Unit costs in health care sector relative to the other sectors of economy • Resource inputs, both human and capital	• Contribution of public and private budgets to the financing of health care • Insurance schemes • Remuneration schemes in health care sector • Regulation and/or liberalisation of the market for health care services and pharmaceuticals

Source: Przywara (2010)

A major issue is the use of good proxies for the development of new technologies and the medical progress that reflects reality. The following section explains the data that is used in this thesis in order to establish a time series model for Austria and Germany.

2.1. Data

Public current health expenditures (PCHE) in EUR without long-term costs (LTC),¹ and public investments (I) are used as the dependent variable. Furthermore the time series was divided by the Austrian/German population and the according GDP deflator with 2005 prices.

$$PCHE = \frac{(PHE - LTC - I)}{population \times deflator_{2005}}$$

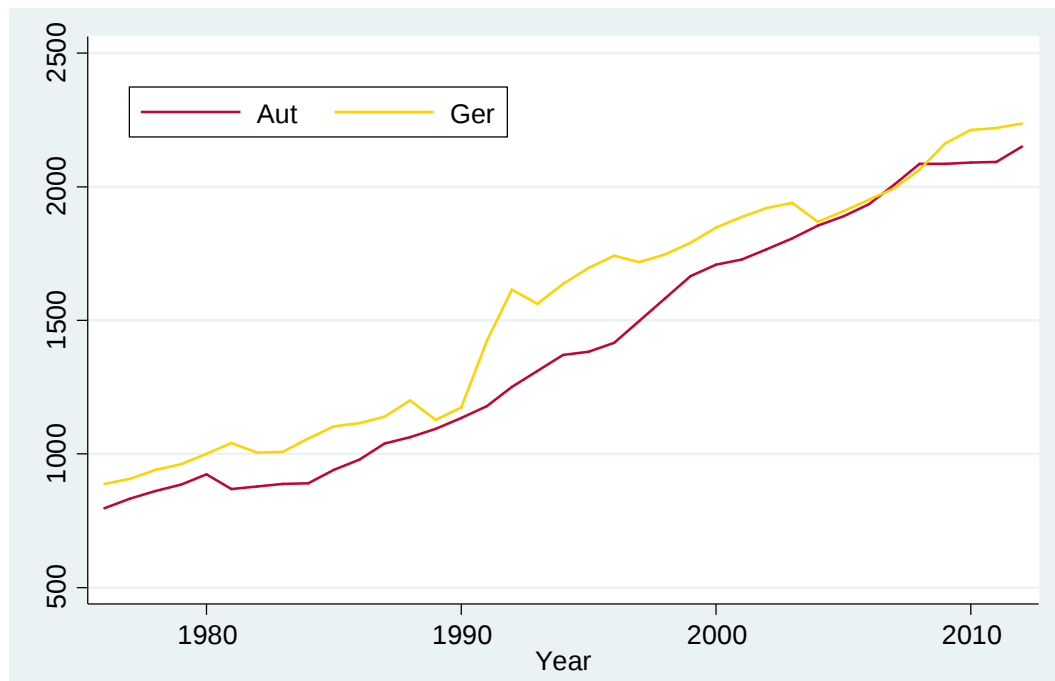
As an alternative it would be possible to use total (current) health expenditures (THE/TCHE) or public health expenditure (PHE). The timeframe for both countries is between 1976 and 2012. The data was taken by OECD Health Statistics and Statistics Austria² and was calculated on the concept of “System of Health Accounts (SHA)”. For the case of Austria there was missing data in the time series for LTC before 1990, which were corrected by rewriting the series from 1990 backwards with the growth-rates of PHE (with LTC) before that point. The German data for investments, long-term costs and overall public health expenditure had breaks in 1991, which were corrected by taking the mean of the adjacent years. The GDP deflator for both countries was routinely calculated by taking the ratio of the nominal and real absolute gross domestic product (GDP) in absolute terms (bn) and multiplying by 100. Furthermore, the population data for each country was used without modification or transformation. Figure 2 represents the two series of further interest.

Demographic Determinants

In order to see the effect of demographical change, *the fraction of the population over 65 years (age65)* is used (in percent of total population). The increase in the proportion of older people is expected to accelerate the growth of health spending. Furthermore it can be assumed

¹ There were two reasons for exclusion: (1) the provision of the time series was not complete and (2) it can be assumed that other determinants for health spending also affect LTC

² OECD Statistics <http://stats.oecd.org/> and Statistics Austria <http://www.statistik.gv.at>

Figure 2: Public Current Health Expenditures, pc, EUR, PPP2005

Source: OECD Health Statistics, Statistics Austria

that life-prolonging medical interventions are to be carried out at advanced ages. The fraction of the population over 65 years therefore can also be partly interpreted as a proxy for medical/technological progress (Riedel et al. 2002). As an alternative it is possible to split the variable *age65* into more intervals (e.g., 65-75y, 76-85y, 85+y). Similar intervals, like a fraction of under 15- or 20-year olds, did not increase the explanatory power of the models and therefore were not used. Variables for life expectancy (*le65*, *le75*) were also checked but not used for the same reason. For Austria the data was taken by Statistics Austria and for Germany by OECD Demographic Statistics.

Social Determinants

As a social indicator, the *unemployment rate*³ (*unemp*) was chosen. Job security enforces health and wellbeing. On the other hand, studies (Paul and Moser 2009, Schröder 2013) show unemployment puts health at risk and increases the danger of premature death. The determinant also – to a certain degree – can be seen as an indicator for social inequalities. For the latter an alternative option would be to choose the *gini-coefficient* or the *rate of academics*. Since these two determinants did not provide more explanatory power than the rate of

³ National definition, in percent

unemployment, they were not used. For Austria data was taken from ArbeitsmarktserviceÖsterreich (AMS)⁴ and Germany's data from OECD Labour Statistics.

Economic Determinants

One difference in the model this thesis uses is the economic determinant. In most cases the *gross domestic product(GDP)* is used as a proxy for the economic development of an economy (Przywara 2010). It has been shown that – in per capita terms – countries with a higher GDP spend more on health (in absolute and relative terms) as countries with a lower GDP (Newhouse 1992, Westerhout 2014). The opposite direction – health status being a driver for GDP – also has been analyzed (Suhrcke et al. 2005), but is not an issue in this thesis. As an alternative, models with *wage income* or the *national deficit of state budget (def)*⁵ were estimated and the latter variable was found to be way more significant than the other two determinants. The national deficit of state budget does not picture the economic development of a country, but in general matches the difference between government spending and income and seems to be a more important factor in health expenditure financing than the country specific GDP. Due to this fact, *def* was used for both cases – Austria and Germany – whereas for the model Germany, the time series was smoothed by a moving average, giving the year $t=0$ a weighting of 0.5 and the previous/following year a weighting of 0.25. It should be noted that this smoothing-filter is two sided and the observation in $t=0$ then depends also on future observations. There may be an impact on causality statements and forecasts.

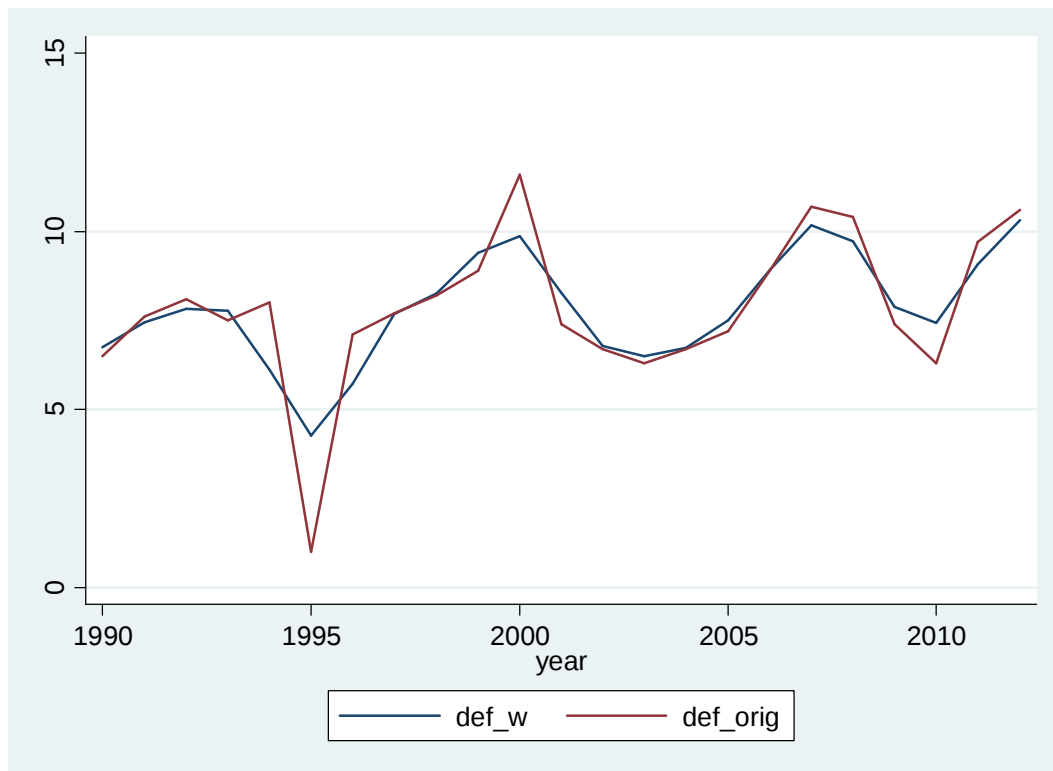
$$\text{def_w}(0) = \frac{1}{4}\text{def}(-1) + \frac{1}{2}\text{def}(0) + \frac{1}{4}\text{def}(1)$$

The smoothing-filter attempts to capture the patterns of the data, but ease data points that are extreme high or low (Figure 3). The data was provided by Eurostat (Austria)⁶ and OECD Statistics (Germany).

⁴ In cooperation with the Main Association of Austrian Social Security Institutions and the Austrian Institute of Economic Research (WIFO)

⁵ Public deficit (Maastricht) in percent of GDP

⁶ Eurostat <http://ec.europa.eu/eurostat>

Figure 3: Smoothed/Weighted deficit and original deficit for Germany (1990 to 2012)

Public Policy Factors

In order to add effects of public policy, legal reforms were investigated.

Table 1 shows a list of reforms that may play a role in health care spending. If a reform had a clear objective (e.g., cost containment) it is also noted. The reforms were added as dummy variables to the models. In fact not all reforms appeared to have explanatory power in the respective models, therefore only reforms in 1989, 1997 and 2005 were added in the case of Austria, and the reforms in 1988, 1993, 2001 and 2007 were added into the German model variants. A full list of identified reforms is available in the appendix (Table 11, Table 12).

Table 1: Reform-dummies; Austria and Germany

	Year	Reform
Austria	1989	Fee for inpatient stays (objective: cost-containment)
	1997	6 th 15a Agreement “Einführung der leistungsorientierten Krankenanstalten-Finanzierung”; increase in prescription charges, introduction of health insurance certificate fee (2005 abolished). “Revenue-Oriented Expenditure Policy” (objective: cost-
	2005	“Reformpool”; Increase in health insurance premiums, Increase in the maximum contribution base, Increase in tobacco tax
Germany	1993	‘Health Care Structure Act of 1993’ came largely into effect on 1.1.1993 (objective: cost-containment and introduction of competitive elements): – Introduction of a risk structure compensation (Risikostukturausgleich) for redistribution of contributions by the health insurances (from 1994); – Abolition of the cost-recovery principle for hospitals – Introduction of legally fixed budgets or spending caps for certain sectors of the health systems Increase of co-payments for drugs (depending on introduction of the drug or the packaging size)
	2001	Between 2001 and 2003 a large number of laws and regulations were adopted (objective: compensate shortcomings of the health insurances). Here are just a few listings: – Reform of risk-structure compensation: Redistributing resources in order to compensate differences between health insurances – ‘Drug Budget Repeal Act’: drug budgets were repealed and replaced by regional agreements. In addition a law limiting the expenditure on drugs was adopted and a one-time-payment of the Association of Research-based Pharmaceutical manufacturer (VFA) was made in order to disclaim on a general price reduction for certain drugs. Introduction of DRG-based payment system
	2007	‘Health reform 2007’: – Health insurance companies increase salary-related rates by 0.5 percentage points – The current limitation of the physicians’ compensation to a fixed total budget is set aside. Instead, the fee will be converted to flat rates per service – Pharmacies have to pay a higher discount (EUR 2.30) per prescription drug than before (EUR 2.00) to the SHI Introduction of a health fund with a uniform contribution rate

Source: Hofmarcher and Rack(2006), Hofmarcher and Quentin (2013), Busse and Blümel (2014)

Determinants of Technological Change

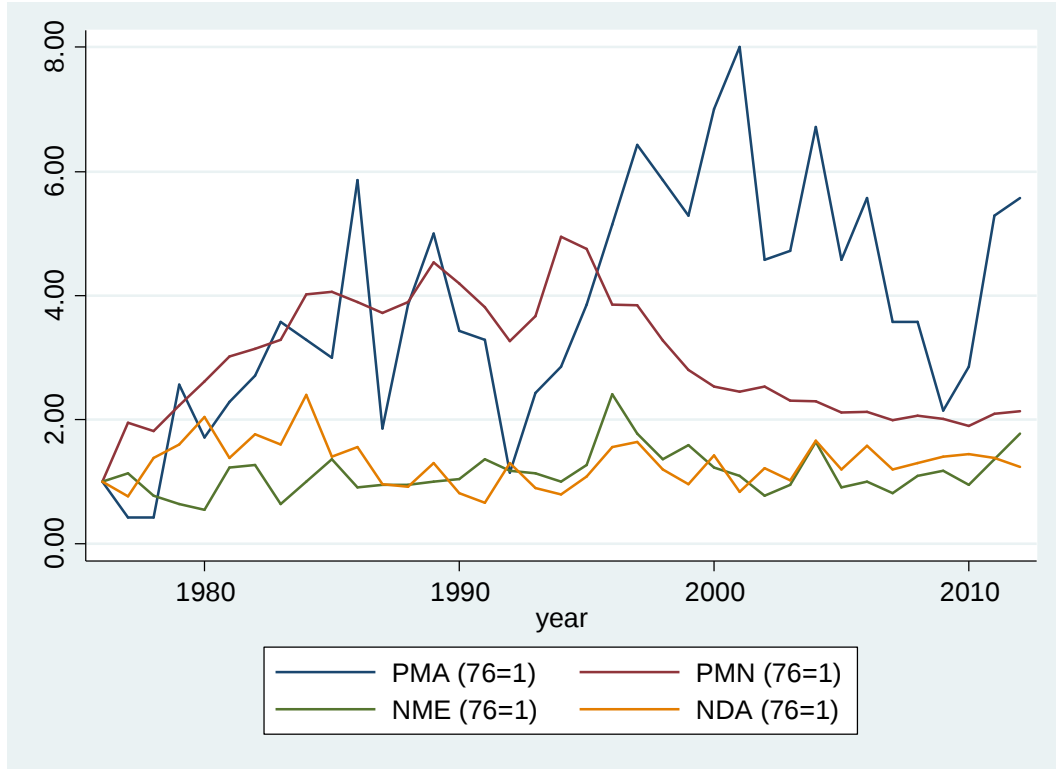
The identification of a determinant for the medical-technological progress is a special challenge. Economists have tried to empirically measure the contribution of technology in different ways. The most common is the total-factor productivity, also called residual approach: if all other variables can be observed and are ‘known’, the contribution of technological progress is captured in the residual (including the errors and all other omitted variables). Alternatives would be to use linear trends, or specific technology indicators, like MRI/CT-scanner density, drug-expenditure, expenditures for physician service or R&D expenditures. Even the percentage of hospitals offering burn care and/or cardiac catheterisation, the percentage of hospitals with organ transplant capabilities and the number of academic health centers were used to estimate the factor of technology. But most of these potential indicators can be explained by other affairs (e.g., the growth in expenditures for physician service can mainly be explained by inflation).

One recent contribution to the literature was the introduction of data by the U.S. Food and Drug Administration (FDA) on the number of approved new medical products as a proxy for medical technological innovation. Willemé and Dumont (2015) use data from four different kind of products, namely, *pre-market approvals (PMA)*, *pre-market (or 510(k)-) notifications (PMN)*, *new drug applications (NDA)* and *new molecular entities (NME)*. The FDA distinguishes between three classes of devices. All ‘Class III’⁷ devices have to stand certain restrictions and require pre-market approval (PMA). Once a device has the PMA-status, it is ready for market introduction and the manufacturer has shown that the device is safe and effective. Medical devices called PMN or 510(k) must be effective and safe as a legal market device, and be registered at the FDA at least 90 days in advance of market-introduction. The devices must be substantially equivalent in safety and effectiveness to a device already on the market. A device that reaches market via the 510(k) process is not considered to be ‘approved’ by the FDA. Nevertheless, it can be sold in the U.S. as a ‘510(k)-cleared’ device. NDAs are new drug applications for drug-licenses in the U.S. market. These applications contain originals only and no generics. New molecular entities or NMEs are a subset of NDAs. An NME is a drug that contains an active moiety that has never been approved by the

⁷ Defined as devices that support or sustain human life or that are of substantial importance in preventing impairment of human health or present a potential, unreasonable risk of illness or injury

FDA or marketed in the U.S.⁸ Figure 4 shows the four medical products in a timeframe from 1976 to 2012. Since the absolute number of each category differs substantially, each one has been indexed to 1 at the initial year. As one can see the progress of technological innovation does not follow a clear trend.

Figure 4: New medical products by FDA, 1976-2012 (1976=1)



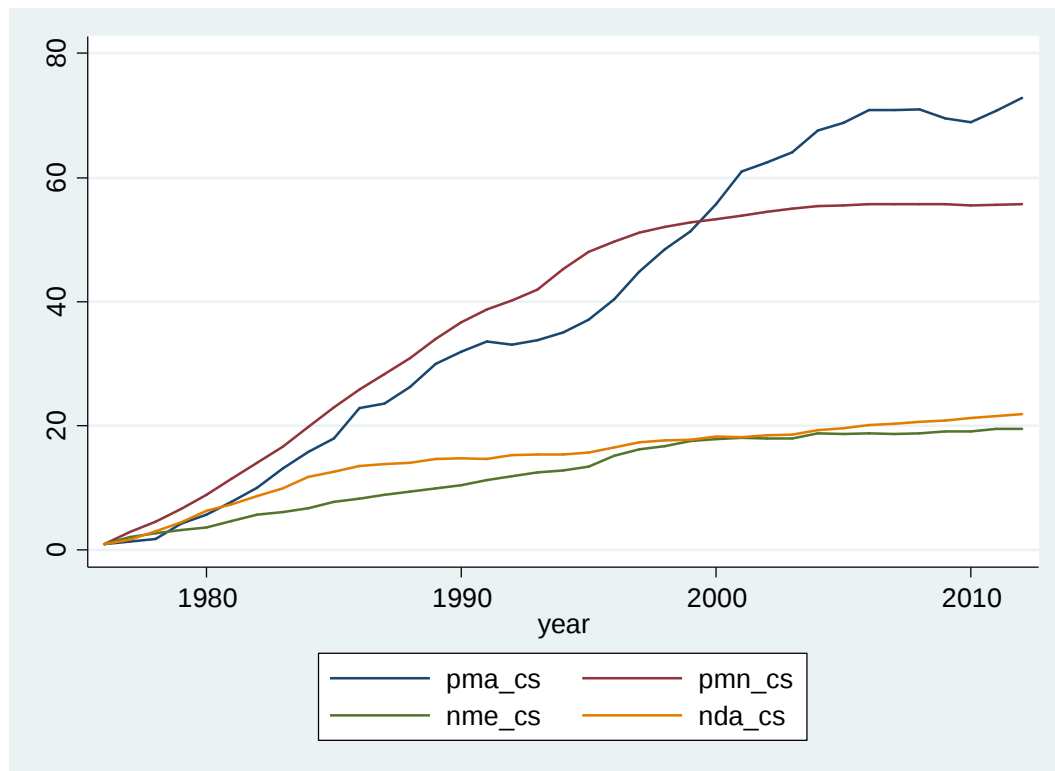
Source: U.S. Food and Drug Administration (2013), illustration acc. to Willemé and Dumont(2015)

Willemé and Dumont (2015) additionally construct capital stocks (CS) of each technology variable using the perpetual inventory method done by Lichtenberg (2005):

$$PMA_{CS}(t) = (1 - \delta) \cdot PMA_{CS}(t - 1) + PMA(t), \text{ with } PMA_{CS}(t = 0) = PMA(t = 0)$$

A certain depreciation rate $\delta \in [0,1]$ determines the dynamic of the capital stock which is based on the capital of the former years. The initial value in this case is the first data point. Figure 5 illustrates the capital stocks of the technology proxies with a depreciation rate of 0.05.

⁸ More detailed information can be found on the U.S. Food and Drug administration website (<http://www.fda.gov/>)

Figure 5: New medical products by FDA in capital stocks (5% depreciation), 1976-2012 (1976=1)

Source: U.S. Food and Drug Administration (2013), own calculation

The idea to use capital stocks with a certain rate of depreciation is motivated by the fact that technology becomes outdated and is replaced by newer technologies. One critique could be the decision making that involves fixing an arbitrary depreciation rate.

In fact for the model of Austria, the method of capital stocks was not used. The untransformed time series with lags between two and four years were used, which can be explained by the fact that the data is based on developments in the U.S. and it may take a while to implement technologies, which are used in U.S. markets, into Austrian markets. Also for Germany there seemed to be an insufficient alternative compared to the untransformed series. Willemé and Dumont (2015) use both methods– transformed and untransformed– while finding significant results in their panel analysis. In this thesis' country specific analysis, no capital-stock models with value for further analysis could be estimated.

Other Determinants

There are further determinants that could affect public current health expenditure, e.g., determinants connected with hospitals or the workforce in the area of healthcare. It was expected to have the following determinants as cost determining factors in a model:

- Physicians per 1,000 inhabitants (*phys*)
- Actual number of beds per hospital (*beds*)⁹
- Days of hospitalization (*beddays*)
- In-patient stays in hospitals (*stays*)

As for the case of Austria, only *phys* seemed to have explanatory power. The data (for Austria and Germany) was taken by OECD health statistics.

2.2. Model

2.2.1. Austria – method A

In order to explain effects that cause health expenditures to rise, *PCHE* (denoted by h_t) was regressed on the covariates listed and described above. The model used was the following:

$$\Delta \log h_t = \alpha + \Delta \log X_t' \beta + \epsilon_t.$$

All variables (except reform-dummies) are in logarithms (log) so as to have normally distributed residuals that result in a valid t-test. Furthermore first differences denoted by the operator Δ (e.g., $\Delta z_t = z_t - z_{t-1}$) were used in order to eliminate possible unit roots. To find out whether or not the data follows a stationary process, an augmented Dickey-Fuller test (DF) was routinely carried out. There are other methods for testing this issue, e.g., the Phillips-Perron test, which had the same result. However, the augmented DF test that is known to perform better in small samples is used in the further analysis (Davidson & MacKinnon 2004).

Having first differences result in stationary data and the application of second differences were not necessary. The vector of k explanatory variables at time t is denoted by X_t' and the vector of coefficients with k elements is denoted by β .

The transformation into logarithms eliminates all values that are negative. Since the entire series on *public deficit* (*def*) is negative, two options arise: taking the logarithm and accepting the loss of information or adding a constant c to each value in order to shift all values into a

⁹ Beds in public hospitals are – among other things – classified as ‘licensed’ or ‘physically available’, reflecting the number of beds that *should* or *could* be in use. ‘Actual’ are those that are *de facto* in use.

positive area. A criticism of the latter method is that some statisticians do not like to add an arbitrary constant to the data. Since all values are in a negative area it seemed legit to take the logarithms of *def*'s absolute values while being aware of the reverse interpretation of the variable. The correlation and partial autocorrelation functions of the data declare that the data is stationary. The modelling of an autoregressive-moving-average model (ARMA) was not realized; instead an AR(1) model was done separately in section 2.2.2. If there was a sign of having autocorrelation in the residuals, the estimator would be consistent and unbiased, but not efficient.

After using *Heteroscedasticity and the Autocorrelation Consistent (HAC-)* estimator to correct for inconstant variance of residuals over time, the four models, as shown in Table 2, were built.

There are various ways to check the stability of a model, which is considered important to having a valid forecast.¹⁰ If one assumes a structural break at a certain point in time, a Chow-Breakpoint test is recommended. Also a lot of visual examinations can be useful in evaluating the stability of a model. A formal statistical test that can be applied is the cumulative sum control chart (CUSUM), which is based on the residuals from the recursive estimates. By applying the CUSUM-test and finding no recursive residual outside the 5%-significance band, each of the four models can be seen as stable.

In order to quantify the severity of multicollinearity the variance inflation factor is estimated, providing a measurement how much the variance of an estimated regression is increased because of correlation between variables. Comparing the Variance-Inflation factor values method and finding all values clearly in a non-suspicious area, the decision is to keep all variables.

¹⁰ Throughout this thesis, the terms 'predicting' and 'forecasting' are used equivalently.

Table 2: Model Variants Austria

Dep. Var: PCHE Austria								
	MV1		MV2		MV3		MV4	
c	0.04	***	0.02	***	0.05	***	0.04	***
age65	1.49	***	1.66	***	1.54	***	1.45	***
def(-2)	-0.01	***	-0.01	***	-0.01	***	-0.01	***
pma(-3)	0.01	*	0.01	**	0.01	***	0.01	***
pma(-4)	0.03	***	0.03	***	0.03	***	0.03	***
pmn(-3)	0.06	***	0.05	***	0.04	(*)	0.05	***
nda(-2)	0.02	***	0.02	***	0.02	***	0.03	***
nme(-3)	0.013	*	0.02	***	0.01	**	0.01	***
unemp	-0.03	*			-0.05	***	-0.03	*
phys	-0.43	*			-0.76	***	-0.81	***
ref_2005	-0.03	***	-0.03	***	-0.02	***	-0.02	***
ref_1997					-0.01	(*)	-0.02	***
ref_1989							0.02	***
Adjusted R ²	0.69		0.67		0.71		0.82	
AIC	-5.44		-5.41		-5.48		-5.94	
BIC	-4.93		-5.00		-4.93		-5.75	

Note: All variables in first differences and logarithms (except reform dummies). Method: Least Squares, HAC standard errors & covariance. Sample (adjusted): 1981 to 2012 (n=32). Significance level: *** p<.01, ** p<.05, * p<.10, (*) p<.15.

It can be noted that the regression is robust in terms of the addition of variables. Change of coefficients is only slight; they remain significant and hold the same sign. In comparing the four model variants, the striking influence of *age65* and *phys* can be noticed. As expected a rising fraction of the older population causes a rise in current public health expenditure. Otherwise an increase of physicians per 1,000 inhabitants will help to reduce current public health expenditures. This can perhaps be explained by the fact of a shortage of physicians in Austria and that an increase in this workforce would help to reduce the payment of physician's long hours and also decrease wait times for special medical treatments, like operations, especially if the wait time is realized in a hospital. One more reason could be that more physicians result in an overall healthier society and therefore a reduction in the amount of severe/costly treatments.

Since *defis* in absolute terms, the sign has to be interpreted the other way around. A lowering of the national deficit in percent of GDP should lower *PCHE* which seems plausible if one assumes a general cut in spending in order to obtain a lower deficit. It is surprising that rising unemployment seems to have reducing effects on the dependent variable. Besides the fact that in MV1 and MV4 *unemp* is not highly significant, there is no intuitive explanation for a

negative sign and one potential critique of the model. Nevertheless, due to the significance of the variable and against intuitive economic understanding, it was decided to keep the variable, because of a lower Akaike information criteria AIC (and higher R^2) in the model MV4.

As it happens the big health reform in 2005 and the reform in 1997 seemed to have had their desired effects of cost-containment.

In contrast to the results of Willemé and Dumont (2015), not only new molecules (NME) and new ‘Class III’ devices (PMA) have a positive sign, but also approved new drugs (NDA) and 510(k)-devices (PMN). As a matter of fact, not only are ‘radically’ innovating products (NME and PMN) experiencing cost-increases, but also ‘incrementally’ innovating products. On the one hand, the cost-increasing effect may be due to treatment expansion as explained by the two authors; on the other hand, this may be caused by special conditions in the Austrian health care system. As the current and special conditions of the Austrian healthcare system are not the topic of this thesis, the latter argument needs further analysis. Another difference besides the sign is the effect of the technology on health expenditure. Willemé and Dumont (2015) are making technology accountable for 43% of growth in total health expenditure. As in the regression above, the effect of technology is much smaller. Furthermore this comparison is not fully appropriate because of (1) the different methodology and (2) the length of time used. In this regression, a time series from 1976 to 2012 was used, compared to the panel estimation from 1980 to 2009, which did not include the latest crisis. Also partial national data (Statistics Austria and OECD) instead of pure OECD data was a resource and may cause differences.

2.2.2. Austria – method B

In method B an autoregressive moving average model, which is in general from the following form, was estimated:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

In order to find out whether or not the variable *PCHE* follows a stationary process, again an augmented Dickey-Fuller test was performed. Since the test cannot reject the null hypothesis the variable to have a unit root, it is probably not stationary but integrated of order 1. Using first differences of the series solved this issue. From an economic point of view, the use of first differences of *PCHE* and in a second step predicting the change in levels of *PCHE* does

not make a difference. Performing the Dickey-Fuller test again on $\Delta PCHE$ leads to the result in Table 3. Also a trend or drift in the data can be excluded.

Table 3: Dickey-Fuller test on first-differenced PCHE

Dickey-Fuller test for unit root		Number of obs = 35	
		Interpolated Dickey-Fuller	
Test	1% Critical	5% Critical	10% Critical
Statistic	Value	Value	Value
Z(t)	-3.823	-3.682	-2.972
			-2.618
MacKinnon approximate p-value for Z(t) = 0.0027			

In a next step an investigation (Table 4) was performed to identify a proper lag-order using the information criteria by Akaike (AIC), Schwarz (SIC, SBIC or BIC) and Hannan-Quinn (HQIC). Alternatively the possibility stands to examine the auto correlation function (ACF) and partial auto correlation function (PACF) for the sample of transformed data. Here the approach via information criteria is preferred. For a small amount of observations (n) or a large number of predictors (k) relative to n , a corrected version of the $AIC = 2k - 2\ln(L)$ –with $\ln$ denoting the natural logarithm and L the maximum likelihood function for the model that minimizes the expected mean squared error –is used (Lindsey & Sheather, 2010). There are again several variants of the corrected AIC (AIC_c), e.g., those done by Hurvich & Tsai (1989) or Hyndman et al. (2008). They are somehow similar as they all add – at least for the cases examined in this thesis – a positive bias-correction to the original AIC, akin to the version of Burnham & Anderson (2004):

$$AIC_c = AIC + \left[\frac{2k(k+1)}{n-k-1} \right]$$

While for the purpose of forecasting, AIC_c or AIC will be the more reliable criterion (compared to BIC or HQIC), solely the AIC will be estimated for comparison. It needs to be stated that the AIC and BIC values are not comparable between the different model variants. Therefore from the two different model methods A and B the best are chosen according to the lowest AIC. The comparison and evaluation between method A and B is then done on the basis of a ‘simulated’ or ‘pseudo’ out-of-sample forecast.

Table 4: AIC and BIC for first-differenced PCHE

Selection-order criteria

Sample: 1982 - 2012

Number of obs

=

31

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-146.834				812.242	9.53765	9.55273	9.58391
1	-143.473	6.7213*	1	0.010	697.607*	9.38535*	9.41551*	9.47787*
2	-143.393	.1606	1	0.689	740.57	9.44469	9.48993	9.58346
3	-143.013	.75906	1	0.384	771.463	9.48472	9.54503	9.66975
4	-143.01	.00581	1	0.939	823.87	9.54905	9.62444	9.78034
5	-141.57	2.8814	1	0.090	802.461	9.52061	9.61109	9.79816

The information criteria are in line to use a lag of 1. Nevertheless, all variants between AR(0 to 5) and MA(0 to 5) were estimated and compared by their information criteria AIC and BIC, so as to find the lowest value in an ARMA(1,0) model, respectively:

Table 5: Resulting AIC values for combinations of AR and MA (Austria)

	MA0	MA1	MA2	MA3	MA4	MA5
AR0	351.45022	348.58406	349.42435	351.42398	351.47696	348.04932
AR1	347.84508	349.84177	351.42423	.	354.63741	350.08868
AR2	349.84003	349.0567	349.94737	354.66985	357.28571	.
AR3	351.49763	353.49704	349.7315	351.95324	353.45496	.
AR4	353.49447	355.14211	352.73973	353.27385	353.12168	350.51489
AR5	352.6953	352.70754	352.06282	356.01248	358.00953	352.51286

An AR(1) model was estimated depending only on *PCHE* itself. The coefficient on the first lag is highly significant and indicates that the past influences the present period in a positive way (Table 6). The AR(1) is stable (or asymptotically stationary) as its characteristic polynomial equation has only roots greater than one in modulus.¹¹

¹¹For the case of AR(1) this property is easily checked, having $|z| = \frac{1}{\phi_1} \approx 2.67 > 1$

Table 6: AR(1) model Austria

ARIMA regression

Sample: 1977 - 2012

Number of obs = 36

Wald chi2(1) = 6.13

Log likelihood = -170.9225

Prob > chi2 = 0.0133

dpche	OIM		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
dpche						
_cons	38.43291	7.301864	5.26	0.000	24.12152	52.7443
ARMA						
ar						
L1.	.3746458	.1513095	2.48	0.013	.0780847	.6712069
/sigma	27.84838	3.282191	8.48	0.000	21.41541	34.28136

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

2.2.3. Germany – method A

In the model method A it would be possible to use data from before the reunification and therefore employ a dummy variable that indicates the presence of such a categorical effect. Nevertheless, after estimating different models with different sample sizes, the models used from 1990 and 1991 onwards performed best regarding explained variability.¹²

The same theoretical model and variable-pool were used as for Austria's model method A. To evaluate the most significant lags of the variables by the Food and Drug Administration, systematically models with all combinations between 1 and 5 lags were estimated separately while having the demographic and economic variable in the model. In a second step the best lags of each variable according to the level of significance (p-value) were hold in the model and step by step combined. Variables with a p-value over 15% were dropped. This was necessary due to the fact of the small sample size and the large number of variables.

Unemployment and variables of the category 'other determinants' had no significant impact and deteriorated the model in almost all estimated combinations, which resulted in a drop of these variables. Of importance is that larger short-term fluctuations occur every 5-8 years, which is in line with changes in the government. Having an eye on this, the reform in 2007

¹²A selection of models is added in the appendix

had a very stabilizing effect in all model variants. Even though this dummy was designed to correct for the categorical effect of the health reform in 2007, it is also possible that other effects, e.g., first effects of the financial crisis of 2007/2008, were covered within. Additionally, the dummy variable `ger_ref_2001(-1)` was implemented, while having the guess that the reform of 2001 showed his impact in 2002. Nevertheless, it was renamed to `dummy_2002`, indicating that the strongest residuals often are not in the place of so-called reforms.

Table 7 shows the chosen model M91_1 among other model estimations.

The variable for the national deficit was not negative for each data point. A switching of the sign and subsequent reverse interpretation as done for Austria was therefore not possible. Also problems may occur by adding for example the highest value in absolute term plus one to every data point. The shifted and positive series may give the possibility to apply a logarithmic calculus. Nevertheless, the interpretation of the variable would not be simple. The applied and more confident solution is to use the alternating series only in terms of first differences.

PMA-devices seem to have a cost increasing effect if they are relatively new. Devices with the age of four years are cost reducing in opposition to devices with the age of one year or two years. There may be several reasons for explanation. There is the possibility of substituting old devices with newer and maybe cost increasing ones in order to improve the provision of health services.

A way around this comes from the possibility of implementing useful devices with evidence in cost reduction. But the process of examining this evidence may take time. Contrary to Willemé/Dumont (2015), pre-market notifications seem to have a cost increasing effect. An intuitive explanation is that PMNs are somehow copies of existing devices, which are shown to be effective and lead to treatment expansion. A NME is a completely new drug, meaning there are no existing treatments or similar drug products. These drugs may be expensive since they are somehow monopolists, but according to the model they induce a fall in health expenditures. A plausible reason is that long-term and chronically-ill patients are a bigger burden to the health system in terms of costs than patients who recover and stop treatment.

Table 7: Selection of Estimationsfor Model Method A – Germany

Dep. Var: PCHE Germany		from 1990 M90_1	from 1990 M90_2	from 1990 M90_3	from 1991 M91_1	from 1991 M91_2
c		0.13 ***	0.20 ***	0.20 ***	0.16 ***	0.21 ***
ger_age65		1.91 ***	2.69 ***	2.01 ***	1.53 ***	1.35 ***
ger_def		-0.11 ***	-0.09 ***	-0.10 ***	-0.03 ***	-0.09 ***
pma(-1)					0.03 ***	
pma(-2)					0.02 ***	
pma(-4)		-0.05 ***	-0.06 ***	-0.06 ***	-0.03 ***	-0.05 ***
pmn(-1)		-0.14 ***				
pmn(-5)					0.14 ***	
nda(-2)			0.04 ***	0.04 ***		0.04 ***
nda(-3)		0.03 ***				
nme(-2)		0.03 ***				
nme(-3)					-0.04 ***	
ger_ref_1988						
ger_ref_1993		-0.14 ***		-0.14 ***	-0.17 ***	-0.16 ***
ger_ref_2001		-0.05 ***				
dummy_2002			-0.02 *		-0.02 ***	
ger_ref_2007		0.05 ***	0.05 ***	0.03 ***	0.05 ***	0.02 ***
REUNION(-1)						
Adjusted R^2		0.86	0.80	0.79	0.94	0.86

Note: Method: Least Squares, HAC standard errors & covariance. All variables are in first differences (except: reform dummies) and logarithms (except:ger_defand reform dummies). Last year of the sample is 2012.

2.2.4. Germany – method B

Germany as it is today was created on the 3rd of October 1990. The reunification of the Democratic Republic of Germany to the then Federal Republic of Germany caused an obvious structural break in the data, which in method A could have been corrected by using a *reunion* dummy-variable. In a model where an auto-regressive (AR) part is used, it seems questionable to use a time-series that consists of two not directly comparable parts. Nevertheless, even if there are possibilities to correct for the break in order to use the whole series – optimal from 1976 onwards –it seems again passable to cut the series at this point. Contrary to variants in the previous method A, the series used for this section's modeland constitutive projection starts directly at 1992.

To find a valid autoregressive moving average model for Germany, the same tests as for Austria were performed, using a modified time-series with data points after 1991. A Dickey-

Fuller test without linear trend cannot reject the null hypothesis that the data has a unit root; hence it is not stationary and instable. Therefore first differences of the health expenditure data were used and a further DF-test shows the data to be stationary (with p-value < 0.001), concluding the data to be integrated of order one. The DF-test with linear trend also cannot reject the null hypothesis for non-differencing data. Respectively, the test with differencing data yields a p-value under 0.0001. The series was modified and a new series was created by taking the trend variable (starting from 1992) into account. This was done in three steps. First, the data of the series was used and the differences of each $PCHE_t$ and $PCHE_{t-1}$ were calculated. Second, the mean of all differences was taken and third the new values for the new series were calculated: $PCHE_t - \mu(PCHE) \times trend\ variable$. This was done for each period. The trend variable increases by 1 over time, starting with a value of 0 for the first period. After successfully performing tests for (trend-)stationarity on the log-diff data, the autocorrelation and partial autocorrelation functions were plotted to determine an appropriate lag order of the model. Since both functions did not show a clear and identifying pattern, in a next step AIC and BIC values were computed for all combinations of MA(0 to 5) and AR(0 to 5).

Table 8: Resulting AIC values for combinations of AR and MA (Germany)

	MA0	MA1	MA2	MA3	MA4	MA5
AR0	-83.875408	-82.156154	-81.411431	-82.806996	-82.892875	-81.429469
AR1	-82.05565	-80.552761	-83.640967	-81.10542	-81.193798	-79.884987
AR2	-80.953393	-84.399746	-84.089678	-79.26926	-78.626967	-78.972575
AR3	-79.928299	-82.737689	-77.930684	.	.	.
AR4	-77.996058	-80.96537	-79.255562	.	.	.
AR5	-78.49005	-78.582027	-76.663668	.	.	.

According to both information criteria, the ARMA(2,1) model yields the best result, presenting the AIC values in Table 8. The second best result according to the AIC is ARMA(2,2). Since there are only 20 observations, a model with a low lag order of 1 was also estimated. The ARMA(1,0) model was then preferred over, e.g., an ARMA(1,1) model, for the same reasons as before.

3. Forecast

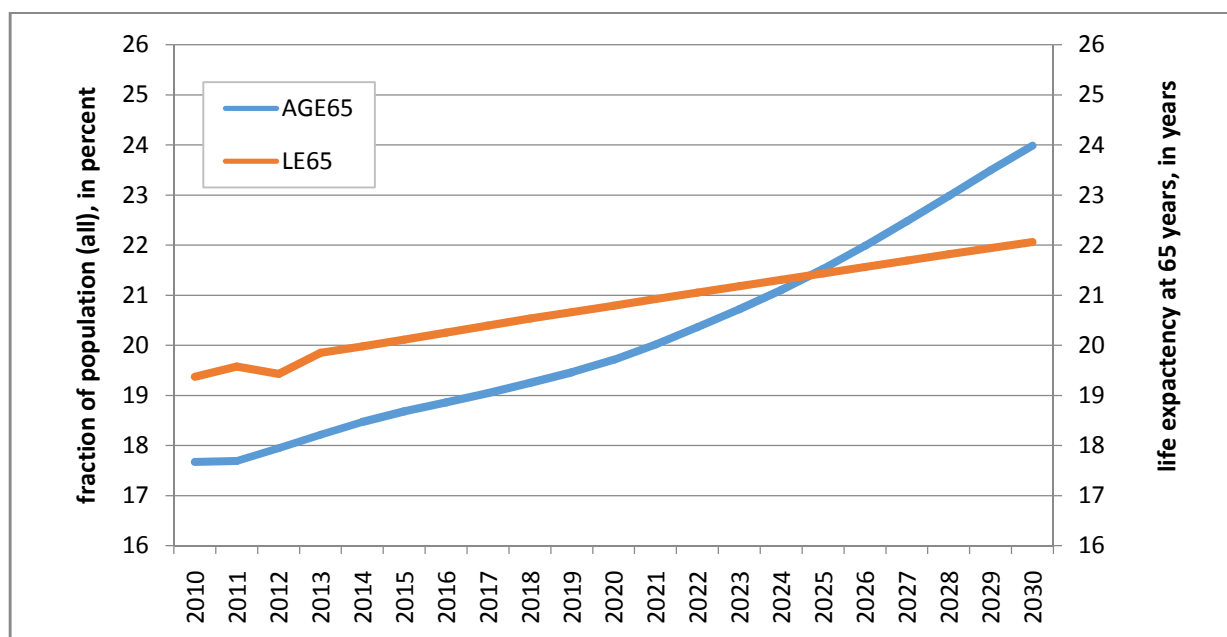
The forecast on current public health expenditure without investments and costs for long-term care is based on the models as presented in the previous chapter. For both countries the original sample (till 2012) will be extended by the forecast sample (2013 to 2030).

3.1. Assumptions

3.1.1. Austria

The fraction of the population that is older than 65 years develops on the basis of Statistics Austria's main scenario (2013) as seen in Figure 6.

Figure 6: Demographical Change, Austria



Note: AGE65: Share of population that is 65 years or older; LE65: life-expectancy at the age of 65

Source: Statistics Austria (2013)

The rate of unemployment (national definition) in 2013 was 7.6%. Based on the WIFO-forecast it will drop from 7.9% in 2014 to 7.7% in 2018. One assumption of the further forecast states that the rate of unemployment will remain constant from 2018 till 2030.

The data for medical devices (PMA and PMN) and new drugs (NME and NDA) was continued from 2014 until 2030 based on the average growth rates from 1976 to 2013.

As a legal requirement of the Stability and Growth Pact by the European Economic and Monetary Union, the national deficit is not allowed to exceed the limit of -3.0%. Same as before, the data was continued from 2014 until 2030 on the basis of the average growth rates from 1976 to 2013 without violating this requirement.

3.1.2. Germany

The share of population older than 65 years develops according to variant 1 of the 13th coordinated population projection (Federal Statistical Office 2015). The variant 1 of the projection, which is named *continuity at weaker immigration*, takes 2013 as the base year. This variant assumes the average number of children per women to be constant at 1.4, while the average reproductive age of women is slightly increasing from 30.7 years to 31.8 years in 2028 and constant afterwards. The life expectancy at birth and the life expectancy at the age of 65 is moderately increasing from on average 80.3 years to 86.8 years and 19.1 years to 23.5 years in 2060, respectively. The net annual immigration falls stepwise to 100,000 in 2021 and stays constant at this level. As a result, the share of population older than 65 years increases from 20.9% (2013) to 27.5% (2030).

As for Austria the number of medical devices, new drugs and the national deficit – for the latter taking the requirement of the -3.0% frontier into account - were continued on the basis of the average growth rates of the previous years.

3.2. Results

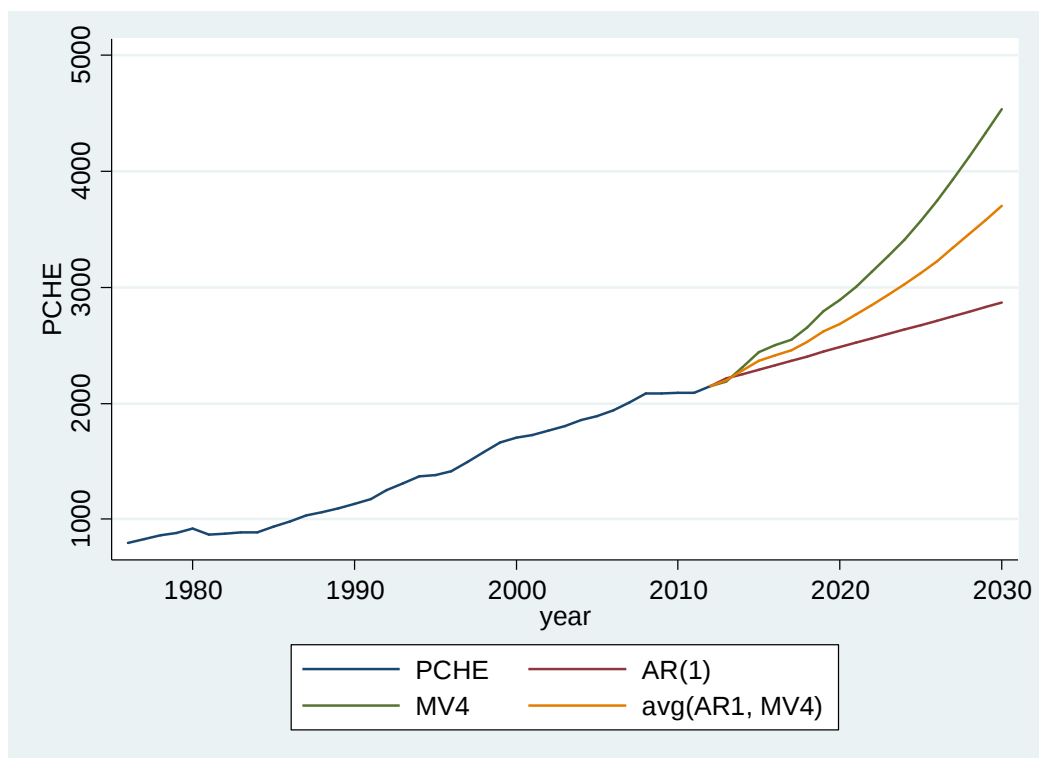
3.2.1. Austria

As the previous assumptions have been made, model A forecast results with increases of current health expenditure per capita of EUR 2892.03 in 2020 and EUR 4538.08 in 2030 (Figure 7: green line MV4).

Since in general all estimated models are wrong to a certain degree it seems to be the reasonable that the acceptance of ‘thick modeling’ – the simple method of combining models and give equal weights – has increased in the last decade. The combination of forecasts was pioneered by Bates and Granger (1969) and since then has been refined in many fields of application, becoming subtler. Wenzel (2001) for example proposes a hits-and-miss methodology in order to weigh different forecasts. If a model ‘hits’, meaning either gets the right direction or has a smaller error than competing models, it is weighted by a higher factor. Often simple averages can be understood as benchmarks to more complex weighting schemes and hold on surprisingly well, legitimating giving method A and B a weighting of 0.5.

The mean standard error in the AR(1)-model for the predicted values is 7.30 and, in contrast to the forecast from the model MV4, stays the same. The forecast is also much more linear than in the variant MV4. Comparing the predicted values and their significant difference, it appears valid to use the average of both predictions (Figure 7). The steep rise of expenditure in the variant MV4 is prone to the big influence of the variable that determines the age of the population. The model AR(1) only indirectly takes a rising population effect into account since there will be a connection between age and health expenditure, but the main reason for a further rise in health expenditure is the continuous rise of health expenditure in the past.

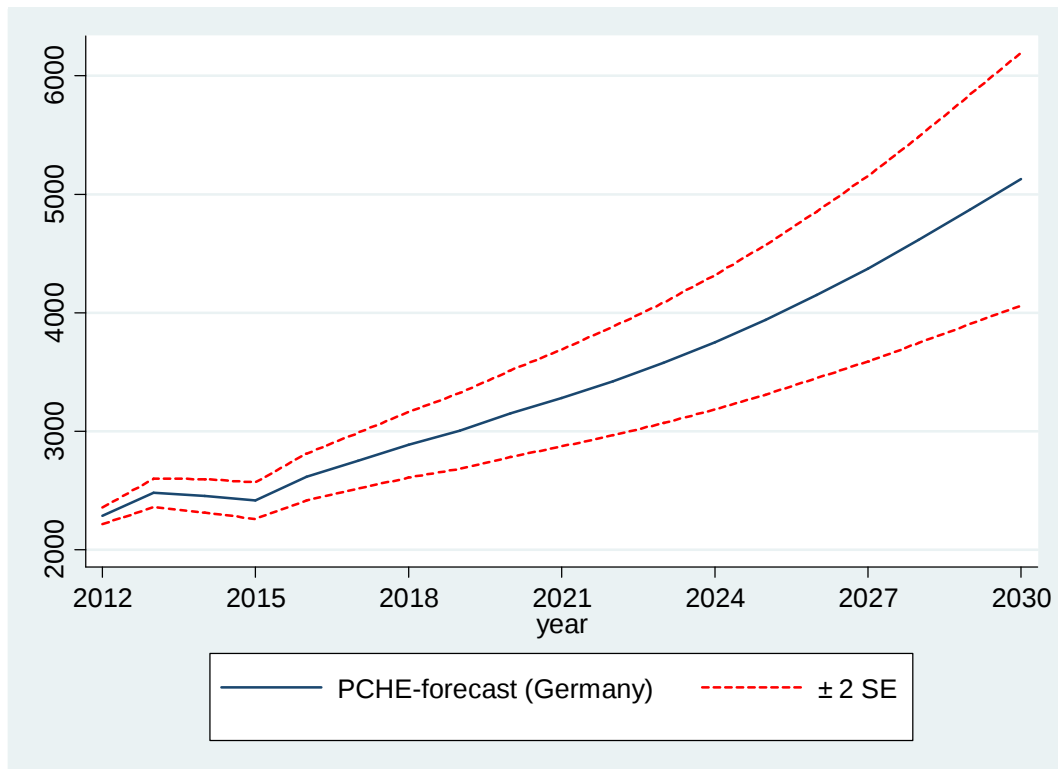
Considering both models equally, the current health expenditure for Austria will increase to EUR 2687.75 in 2020 and EUR 3702.93 per capita in 2030, which is a 24.9% or 72.0% rise, respectively, compared to the expenditures in 2012.

Figure 7: PCHE and its predictions, Austria

3.2.2. Germany

In Figure 8 there are two small edges observable at the beginning of the forecast for Germany. Afterwards a constant rise of health expenditure can be seen. Naturally the longer the period of the forecast, the more inaccurate it becomes. The red dotted lines denote an interval of standard errors.

Figure 8: Forecast Germany method A



According to the previous chapter, the model ARMA(2,1) was selected and projected in a dynamic way. Also the model AR(1) was projected and was very similar to the prediction of the ARMA(2,1) model. More interesting is the shape of the ARMA(2,2) forecast. At the beginning of the out-of-sample forecast a short-term fall in health expenditure can be observed, which is kind of similar in shape – not in value – to the model calculated in method A. For comparison, the models ARMA(2,1) and ARMA(2,2) are presented in Figure 9.

Figure 9: Forecast Germany method B

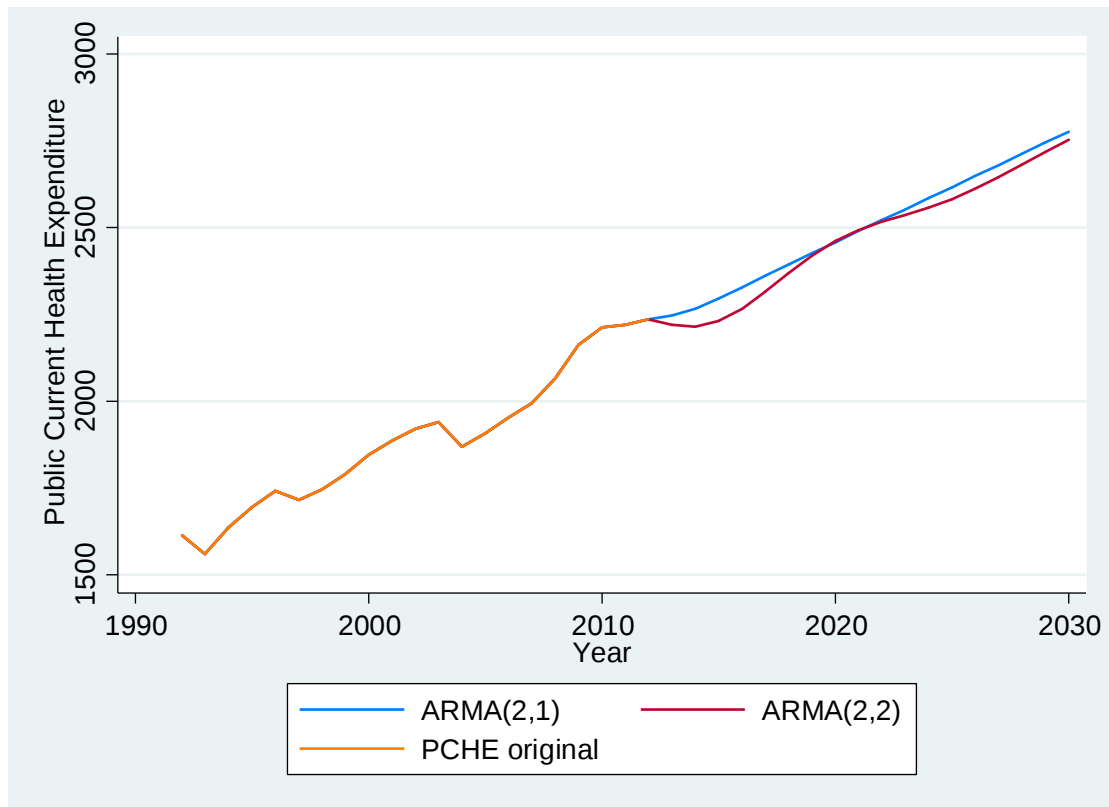
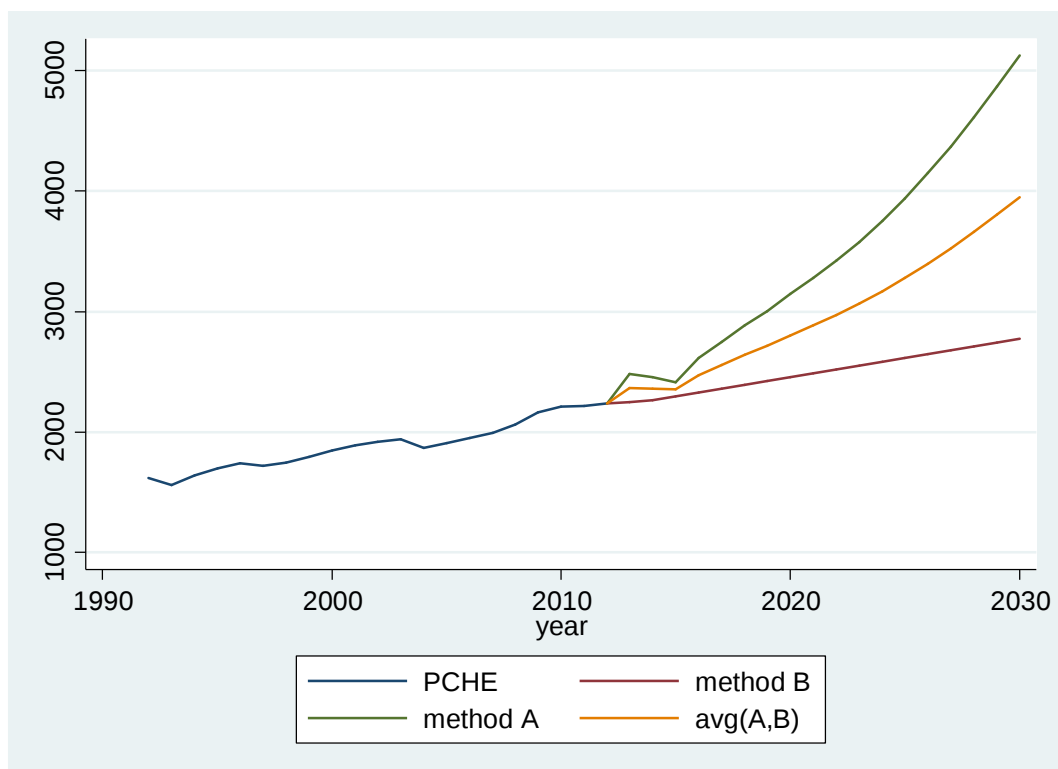


Figure 10: PCHE and its predictions, Germany



Similar to the results for Austria, the model from method A forecasts much higher health expenditure per capita than method B. This is also due to the strong influence of an ageing population predicted by official agencies. In Figure 10 the difference between the two models becomes visible, resulting in an increase of health expenditures to an amount of EUR 2803.92 in 2020 and EUR 3951.91 in 2030, taking both models with equal weight into account. Compared to 2012, this is an increase of 25.3% and 76.7%, respectively.

3.2.3. Evaluating Forecasts

There are several ways to evaluate forecasts. Fair (1986) and Pena et al. (2001) state the most common evaluations are based on the mean errors (ME), mean absolute errors (MAE) or root mean squared errors (RMSE), defined as follows:

$$ME = \frac{1}{n-m} \sum_{t=m+1}^n (y_t - \hat{y}_t)$$

$$MAE = \frac{1}{n-m} \sum_{t=m+1}^n |y_t - \hat{y}_t|$$

$$RMSE = \sqrt{\frac{1}{n-m} \sum_{t=m+1}^n (y_t - \hat{y}_t)^2}$$

Nevertheless, the RMSE seems to be the more valuable criterion since it is more sensitive than the other two measures. Through the squaring, a disproportionate weight is given to very large errors.

However, there is no “good” or “worse” criterion. Often other measures like the mean percentage error (MPE) and the mean absolute percentage errors (MAPE) are also used, having the advantage of being scale-independent when considering multiple forecasting points. Otherwise, they have the disadvantage of being extreme when any y_t is close to zero and thereby distort the value of the actual forecast model (Tofallis 2014). They also can become unfeasible if the data points of y_t are shifted in the direction of $\hat{y}_t$ data points.

$$MPE = \frac{1}{n-m} \sum_{t=m+1}^n \left(\frac{y_t - \hat{y}_t}{y_t} \right) \times 100$$

$$MAPE = \frac{1}{n-m} \sum_{t=m+1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100$$

Furthermore, Theil's coefficient of inequality (TIE) is used in the evaluation of forecasts (e.g., Huang et al. 2014). This measure is bounded between 0 and 1, with 0 as the case of a perfect forecast and 1 indicating a negative relationship or one of the variables being identical to zero.

$$Theil's\ IE = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n y_t^2} + \sqrt{\frac{1}{n} \sum_{t=1}^n \hat{y}_t^2}} \times 100$$

Since it is not always clear which performance measurement is the best to use, it seems justified to compare multiple measurements. In Table 9 and Table 10 some of these measurements are presented. Method B refers to the chosen models for Austria and Germany, ARMA(1,0) and ARMA(2,1) respectively, while the sample has been split into two parts: the construction sample and the validation/hold-out sample. The first part is used to fit the model and predict a following period, while the second part gives the advantage of evaluating the forecast result.

There are no guidelines on how to split the data. However, since the actual number of observations is low, the split should also be marginal in order to prevent a relevant loss in efficiency while the model is fitted only to part of the original time-series. Therefrom the years 2010 to 2012 were cut in order to have a “pseudo” or simulated” out-of-sample comparison (appendix: Figure 11, Figure 12). A true out-of-sample analysis would be to estimate the model based on data up to and including today (here: 2012), construct a forecast of tomorrow's value Y_{t+1} , wait until tomorrow, record the forecast error and re-estimate the model before making a new forecast of Y_{t+2} . Since this procedure will take its time, often “pseudo” or “simulated” out-of-sample analysis are performed. These analysis mimic the procedure described, using some date $T_0 < T$, rather than today's date T , as starting point. The tables evaluate one-step ahead forecasts and multi-step forecasts. In one-step forecasts all models were newly estimated and evaluated. The results show that the overall forecasting

performance of method A and B evaluated via RMSE, MAE, MAPE and Theils' Inequality Coefficient is quite similar on both one-step and multi-step forecasts.

Overall, it can be observed that the models from method A in general perform better according to the measurements. In one-step-forecasts, method A forecasts more accurately than method B, regardless of the forecast error measurement considered. The disparity for Austria in MAE and RMSE values ranges from 16% to 3%, respectively. The improvements for Germany are higher and reach from 28% (MAE) to 19% (MAPE).

Table 9: MAE and MAPE for pseudo out-of-sample period forecasts (2009 to 2012)

Country	Method	One-step forecasts		Step-ahead of multi-step forecasts					
				1		2		3	
		MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
Aut	A	21.84	1.11	27.59	1.28	49.07	2.30	26.01	1.22
	B	25.93	1.22	35.45	1.65	13.06	0.62	21.22	1.48
Ger	A	34.00	1.70	49.70	2.22	65.11	2.92	40.20	1.81
	B	46.98	2.11	45.87	2.07	53.02	2.38	69.87	3.14

Note: Mape in %

Table 10: RMSE and TIE for pseudo out-of-sample period forecasts (2009 to 2012)

Country	Method	One-step forecasts		Step-ahead of multi-step forecasts					
				1		2		3	
		RMSE	TIE	RMSE	TIE	RMSE	TIE	RMSE	TIE
Aut	A	23.48	0.32	27.59	0.65	51.89	1.24	30.93	0.74
	B	24.11	0.35	33.32	0.83	17.34	0.41	33.95	0.80
Ger	A	37.78	0.49	36.24	0.80	72.09	1.13	27.59	0.36
	B	47.00	0.62	45.87	1.04	53.50	0.86	71.59	0.92

4. Conclusion

Compared to other countries in Europe, Austria and Germany had the second and third highest health expenditure per capita in 2012 (OECD 2014), despite most countries facing the same challenges for financing health expenditures. The direction of health expenditure is given by the past and this growth is expected to accelerate due to an increasing share of the population being made up of older individuals. The review of empirical literature on health expenditure determinants has illustrated that age is a crucial factor, but also other determinants like technology have to be involved.

This thesis tries to evaluate current public health expenditures without long-term costs, while applying two different methods. One method (A) uses exogenous variables in order to explain health expenditures, and the other method (B) tries to project the future only based on past experiences of itself. Method A therefore not only uses standard variables, but also a technology proxy as recently appeared in the literature. It was determined that for both countries—Austria and Germany—the technology proxy did not have a big impact on health expenditures. The share of people over 65, on the other hand, had a large significant impact, which in the future could result in an exponential increase of costs. For the case of Austria also the fraction of physicians seems to be a relevant dimension in containing costs, as they help to reduce costs. Both countries have already reacted with reforms that aim to contain costs, though not all had their desired effect from a statistical point of view. Method B in contrast forecasts a more comfortable scenario, as health expenditure will rise linearly. However, the results of both estimations predict an increase of per capita health expenditure in Austria by 24.9% to EUR 2687.75 in 2020 and by 72.0% to EUR 3702.93 in 2030, compared to data from 2012. Health expenditure in Germany will increase by 25.3% to EUR 2803.92 in 2020 and by 76.7% to EUR 3951.91, compared to the last available data in 2012.

While some economists may argue that the gross domestic product needs to be present even in a model using national deficit, this indeed could be a relevant critique. Furthermore, even if literature is united in seeing the ageing of populations as a major challenge for future governments, the effect could be too large in both methods of this thesis. A third point that could not be avoided while investigating countries in Europe is the use of data from the United States Food and Drug Administration. Unfortunately, data on new medicines approved by the European Medicines Agency is only available for a very limited timeframe. In fact the

European Database on Medical Devices (Eudamed) can only be accessed by regulatory authorities.

Still, the methods to achieve the objective of this thesis – to predict health expenditures – were successfully applied and it will be interesting to compare results with actual prospective data. While out-of-sample forecasts often lose their precision the longer they stand, the direction of the results may be a beneficial indication to decision makers in the health system.

Literature

- Bates, J., Granger, C.W.J. (1969): The combination of forecasts. *Operational Research Quarterly*(20), 451-468.
- Burnham, K.P. (2004): Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2), 261-304. doi: 10.1177/0049124104268644
- Busse, R., Blümel, M. (2014): Germany: Health system review. *Health Systems in Transition*, 2014, 16(2):1-296.
- Chernew, M.E., Newhouse, J.P. (2012): Health care spending growth. In *Health Economics. Handbooks in Economics*: North-Holland.
- Czypionka, T., Riedel, M., Röhring, G., Leutgeb, J. (2011): Zukunft der Gesundheitsausgaben und Gesundheitsfinanzierung in Österreich II: Prognose der öffentlichen Gesundheitsausgaben in Österreich und Methodenvergleich mit Ageing Report 2012. Project report. Institute for Advanced Studies, Vienna.
- Davidson, R., MacKinnon, J.G. (2004): *Econometric Theory and Methods*. New York: Oxford University Press.
- Economic Policy Committee (2009): 2009 Ageing Report: Economic and budgetary projections for the EU-27 Member States (2008-2060). *European Economy* 2|2009, Brussels.
- Economic Policy Committee (2011): The 2012 Ageing Report: Underlying Assumptions and Projection Methodologies. *European Economy* 4|2011, Brussels.
- Fair, R.C. (1986): *Handbook of Econometrics*, Vol. III. Z. Griliches and M.D. Intriligator (eds), Elsevier Science Publishers BV.
- Federal Statistical Office (2015): Germany's Population by 2060 – Results of the 13th Coordinated Population Projection. Statistisches Bundesamt, Wiesbaden.
- Gerdtham, U.G., Johnsson, B. (2000): International comparisons of health expenditure: Theory, data and econometric analysis. In *Handbook of Health Economics*. Elsevier, 11-53.
- Hair, J. F., Anderson, R., Tatham, R.L., Black, W.C. (2006): *Multivariate Data Analysis*. Upper Saddle River, NJ.
- Hofmarcher, M., Rack, M. (2006): Austria: Health system review. *Gesundheitssysteme im Wandel*, 2006; 8(37):1-247.

- Hofmarcher, M., Quentin, W. (2013): Austria: Health system review. *Health Systems in Transition*, 2013; 15(7):1-291.
- Huang, T., Fildes, R., Soopramanien, D. (2014): The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research*, 237(2), 738-748. doi: 10.1016/j.ejor.2014.02.022
- Hurvich, C.M., Tsai, C.L. (1989): Regression and time series model selection in small samples. *Biometrika*, 76(2), 297-307. doi: 10.1093/biomet/76.2.297
- Hyndman, R., Koehler, A., Ord, K., Snyder, R. (2008): *Forecasting with Exponential Smoothing: The State Space Approach*. Springer, Berlin Heidelberg.
- Kutner, M.H., Nachtsheim, C.J., Neter, J. (2004): *Applied Linear Regression Models* (4th ed.). McGraw-Hill Irwin.
- Lichtenberg, F.R. (2005): Pharmaceutical knowledge-capital accumulation and longevity. In: *Measuring Capital in the New Economy*. NBER.
- Lindsay, C., Sheather, S. (2010): Variable selection in linear regression. *Stata Journal* 4:650-669.
- Newhouse, J. P. (1992): Medical Care Costs: How Much Welfare Loss? *Journal of Economic Perspectives*, 6(3), 3-21. doi: 10.1257/jep.6.3.3
- Organisation for Economic Co-operation and Development (OECD) (2013): Public spending on health and long-term care: a new set of projections. OECD Economic Policy Papers No. 06. OECD, Paris.
- Organisation for Economic Co-operation and Development (OECD) (2014): *Health at a Glance: Europe 2014*. OECD, Paris.
- Paul, K. I., Moser, K. (2009): Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, 74(3), 264-282. doi: 10.1016/j.jvb.2009.01.001
- Pena, B., Tiao, G.C., Tsay, R.S. (2011): *A Course in Time Series Analysis*. John Wiley & Sons, New York.
- Przywara, B. (2010): Projecting future health care expenditure at European level: drivers, methodology and main results. *European Economy, Economic papers* 417. European Commission, Brussels.
- Riedel, M., Hofmarcher M., Buchegger, R., Brunner, J. (2002): *Nachfragemodell Gesundheitswesen. Projektbericht Teil 2*. Institute for Advanced Studies, Vienna.

- Schroder, M. (2013): Jobless now, sick later? Investigating the long-term consequences of involuntary job loss on health. *Adv Life Course Res*, 18(1), 5-15. doi: 10.1016/j.alcr.2012.08.001
- Spanos, A., McGuirk, A. (2002): The problem of near-multicollinearity revisited: erratic vs systematic volatility. *Journal of Econometrics*, 108(2), 365-393. doi: 10.1016/s0304-4076(01)00144-0
- Suhrcke, M., McKee, M., Stuckler, D., Sauto Arce, R., Tsoлова, S., Mortensen, J. (2006): The contribution of health to the economy in the European Union. *Public Health*, 120(11), 994-1001. doi: 10.1016/j.puhe.2006.08.011
- Tofallis, C. (2014): A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society*, 66(8), 1352-1362. doi: 10.1057/jors.2014.103
- Wenzel, T. (2001): Hits-and-misses for the evaluation and combination of forecasts. *Journal of Applied Statistics*, 28(6), 759-773. doi: 10.1080/02664760120059282
- Westerhout, E. (2014): 14. Population ageing and health care expenditure growth. *International Handbook on Ageing and Public Policy*, 178.
- Willeme, P., Dumont, M. (2015): Machines that Go 'Ping': Medical Technology and Health Expenditures in OECD Countries. *Health Econ*, 24(8), 1027-1041. doi: 10.1002/hec.3089

Appendix

Table 11: Identification of reforms in Austria regarding health expenditure

Year	Reform
1978	1 st 15a Agreement – „Revenue-Oriented Expenditure Policy“ (objective: cost-containment)
1989	Fee for inpatient stays (objective: cost-containment)
1992	Contribution increases
1993	Federal Care Allowance (Bundespflegegeld); Money transfer from the federal budget
1997	6 th 15a Agreement „Einführung der leistungsorientierten Krankenanstalten-Finanzierung“; increase in prescription charges, introduction of health insurance certificate fee (2005 abolished). „Revenue-Oriented Expenditure Policy“ (objective: cost-containment)
2000	Refund of sales growth in pharmacies, margin reductions for wholesalers, renewed increase in prescription charges
2001	Ambulance fee (abolished 2003)
2002	Revision compensation fund (Ausgleichsfond) (objective: cost-containment)
2003	Unification of contribution rates; Increase contribution rates for pensioners
2005	„Reformpool“; Increase in health insurance premiums, Increase in the maximum contribution base, Increase in tobacco tax
2006	– PharmMed: Less stringent restrictions for licensing pharmacies to encourage competition – Service-fee for e-card instead of health-certificate fee (Krankenscheingebühr)
2008	– Cap on prescription-fees; Individuals that reach 2% expenditure on fees of annual income are exempt from paying for fees for the rest of the year. – „Framework Contract for pharmaceuticals“ to slow increase in medication costs – Linking tax expenditure on hospitals to general taxation → Additional funding of EUR 100 Mio per year for hospitals: “Financial Equalization Act 2008” in agreement with 15a.
2010	Minimum income agreement: Recipients (former: social benefit receiver) obtain access (e-card)
2010-2014	Structural Health Fund; funded with EUR 260 Mio of general tax revenue (Expenditure of funds on the achievement of certain objectives) (objective: Reduce level of indebtedness)
2011-2015	Follow-Up Contract (2008 Framework): Pharma-Industry must refund EUR 82 Mio of their benefits till 2015 to the Federal Social Security Institutions (FSSI)
2011/12	LTC-Fund; “Long-Term Care Allowance Reform Act” – covers increases in cost by federal states & local authorities (EUR 685 Mio: 2/3 federal government, 1/3 federal states)
2013	Annual payment of EUR 6 Mio from Pharma-Industry to FSSI

Source: Hofmarcher and Rack (2006), Hofmarcher and Quentin (2013)

Table 12: Identification of reforms in Germany regarding health expenditure

Year	Reform
1988/89	The ‘Health Care Reform Act of 1989’ came into effect on 1.1.1989: – Adjustment of co-payments for dentures depending on regular dental examinations – Fixed amounts for drugs and medical devices – Public committees to regulate shared medical equipment in the outpatient and inpatient care
1993	‘Health Care Structure Act of 1993’ came largely into effect on 1.1.1993 (objective: cost-containment and introduction of competitive elements): – Introduction of a risk structure compensation (Risikostrukturausgleich) for redistribution of contributions by the health insurances (from 1994); – Abolition of the cost-recovery principle for hospitals – Introduction of legally fixed budgets or spending caps for certain sectors of the health systems – Increase of co-payments for drugs (depending on introduction of the drug or the packaging size)
1997	‘Health Insurance Contribution Relief Act’ come into effect on 1.1.1997 (objective: cost-containment) – Excluding operational dental services and oral surgery services from the service catalog for those born after 1978 (abolished 1998) – Reduction of all contribution rates by 0.4% – Reduction of rehabilitative services – Increased co-payments for drugs and rehabilitative measures (1999 and 2000 partially reduced) – Removal of Health Promotion Services (partially reintroduced in 2000)
2001	Between 2001 and 2003 a large number of laws and regulations have been adopted (objective: compensate shortcomings of the health insurances): – Reform of risk-structure compensation: Redistributing resources in order to compensate differences between health insurances – ‘Drug Budget Repeal Act’: drug budgets were repealed and replaced by regional agreements. In addition a law limiting the expenditure on drugs was adopted and a once-payment of the Association of Research-based Pharmaceutical manufacturer (VFA) was made in order to disclaim on a general price reduction for certain drugs. – Introduction of DRG-based payment system
2003	‘Statutory health insurance (SHI) Modernization Act’: – Certain services are excluded from the SHI catalog – Co-payments are adjusted

-
- Introduction of consultation fee for medical and dental visits
 - Increase of manufacturer discounts for cash at patented drugs
-

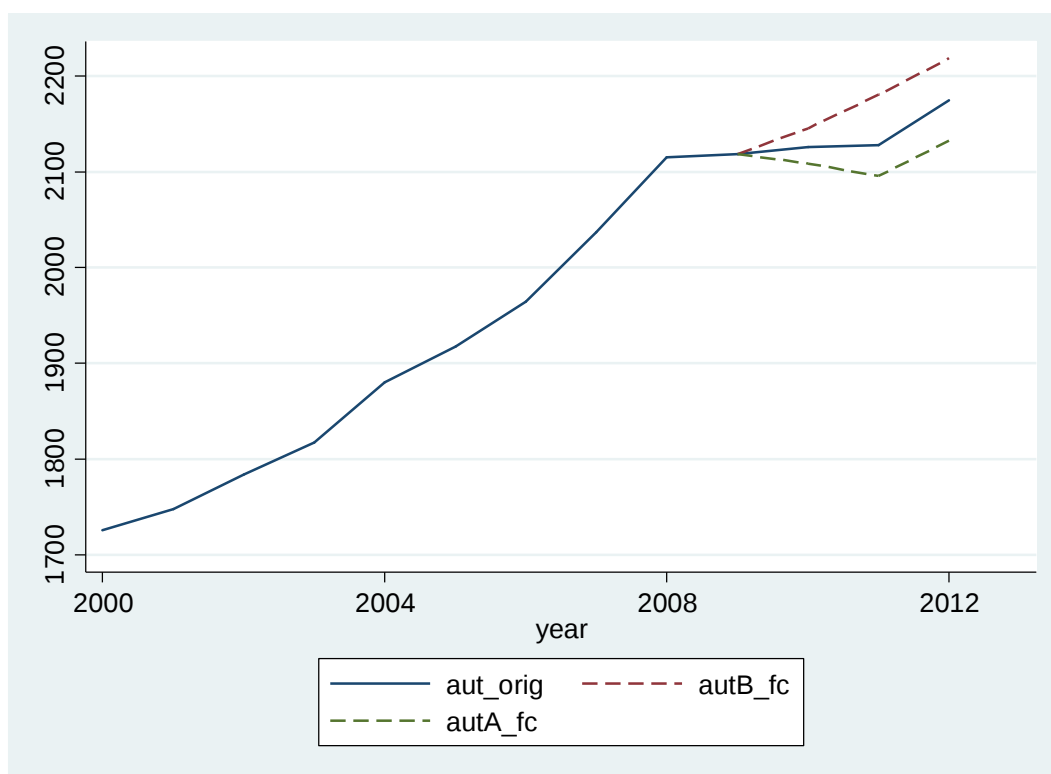
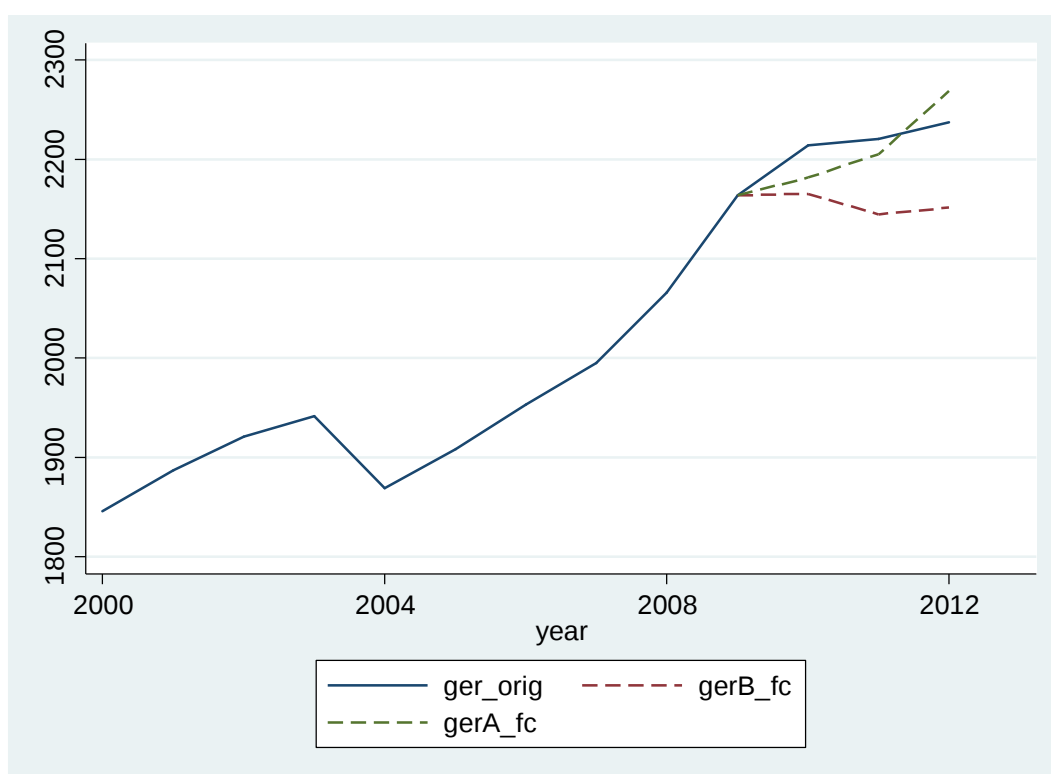
2007 ‘Health reform 2007’:

- Health insurances increase salary-related rates by 0.5 percentage points
 - The current limitation of the physicians’ compensation to a fixed total budget is set aside. Instead, the fee will be converted to flat rates per service
 - Pharmacies have to pay a higher discount (EUR 2,30) per prescription drug than before (EUR 2,00) to the SHI
 - Introduction of a health fund with a uniform contribution rate
-

Source: Busse and Blümel (2014)

Table 13: Correlation of Variables (Austria)

	PCHE	DEF	PMA	PMN	NDA	NME	UNEMP	PHYS
PCHE	1,00							
DEF	-0,33	1,00						
PMA	0,51	-0,30	1,00					
PMN	-0,42	0,43	0,08	1,00				
NDA	-0,10	0,00	0,08	-0,05	1,00			
NME	0,18	0,12	0,37	0,23	-0,02	1,00		
UNEMP	0,77	-0,01	0,61	0,15	-0,19	0,41	1,00	
PHYS	0,99	-0,27	0,54	-0,30	-0,11	0,24	0,85	1,00

Figure 11: Forecast 2010-2012 (sample: 1976-2009), Austria**Figure 12: Forecast 2010-2012 (sample: 1992-2009), Germany**

German Abstract

Die Prognose von Gesundheitsausgaben ist ein wichtiger Indikator für strategische und planerische Entscheidungen im gesundheitspolitischen Bereich. Verschiedene Faktoren, wie der Anteil an älteren Personen (*age65*) und die Wirkung von technologischem Fortschritt werden dabei als zentrale Gründe für kontinuierlich steigende Gesundheitsausgaben gesehen. Aus diesem Grund vergleicht die vorliegende Arbeit zwei verschiedene Modelansätze – mit und ohne exogene Variablen – auf ihre Prognosefähigkeit. Die analysierten Länder – Deutschland und Österreich – zeigen dabei ähnliche Ergebnisse. Modelvariante A verursacht auf Grund der exogenen Altersvariable *age65* einen größeren Anstieg an Gesundheitsausgaben, während die Effekte des Technology-Proxy (Daten für medizinische Geräte und pharmazeutische Produkte) vernachlässigbar bleiben. Im Gegensatz dazu zeigen die ARIMA Modelle ohne exogene Variable in Variante B einen vergleichsweise linearen Anstieg für beide Länder. Auf Grund der Ergebnisse beider Varianten wird eine Erhöhung der pro-Kopf Gesundheitsausgaben in Österreich von 24.9% bzw. 72.0% für den Zeitraum 2012 bis 2020 bzw. 2030 vorausgesagt. In Deutschland werden für denselben Zeitraum die Gesundheitsausgaben um 25.3% bzw. 76.7% ansteigen.

Curriculum Vitae

Frank Kronemann

WORK EXPERIENCE

03/2013 – 08/2015 **Research Assistant at the Department HealthEcon**

IHS Institute for Advanced Studies
Stumpergasse 56, 1060 Vienna (Austria)

10/2011 – 02/2013 **Telemarketer**

Sodexo Motivation Solutions Austria GmbH
Heiligenstädter Lände 27c, 1190 Vienna (Austria)

05/2005 – 04/2015 **Supporting the management**

TwoSeasons – Christian Rauchecker GmbH
Hermann Bahr Promenade 2, 5020 Salzburg (Austria)

08/2011 – 09/2011 **Intern at the Department Internal Auditing**

Verbund-AG
Am Hof 6a, 1010 Vienna (Austria)

EDUCATION

09/2012 – 08/2015 **Magister rer.soc.oec. (MSc equivalent) in Economics**

University of Vienna, Vienna (Austria)
Diploma-thesis: *Health Expenditure Forecasting and Technological Change* in econometrics
Supervisor: Univ.-Prof. Dipl.-Ing. Dr. Robert Kunst

09/2008 – 06/2012 **Bakk.rer.soc.oec. (BSc equivalent) in Economics**

University of Vienna, Vienna (Austria)
1st thesis: *Shared Space* in law and economics
Supervisor: Ao. Univ.-Prof. Dr. Wolfgang Weigel
2nd thesis “Evidence-based Medicine” in health and pharma economics
Supervisor: Dr. Jörg Mahlich

2006-2008 **Studies in German Philology and Engineering**
University Salzburg

LANGUAGE SKILLS

German	mother tongue
English	fluent
Other	Latinum, Graecum

COMPUTER SKILLS

Good command of standard scientific software (Stata, MS Office, Eviews, EndNote,...)

PUBLICATIONS AND PROJECT REPORTS

Czypionka T, Röhring G, Kronemann F (2015): Gesundheit und Pflege. In Schnabl A, Pretenthaler F (eds.): [Economic effects of public services of the Roman Catholic Church] Ökonomische Effekte der öffentlichen Leistungen der römisch katholischen Kirche. Institute for Advanced Studies and Johanneum Research, Vienna.

Czypionka T, Kronemann F (2015): [International comparison of hospital payment systems] Bezahlungssysteme von Spitälern im internationalen Vergleich. Published in Health System Watch, HSW II/2015.

Czypionka T, Röhring G, Kronemann F, Reiss M (2015): [Hospital financing: Where do resources come from?] Schwerpunkt Spitalsfinanzierung: Woher kommt das Geld? Published in Health System Watch, HSW I/2015.

Czypionka T, Kraus M, Kronemann F (2015): [Payment systems in primary health care] Bezahlungssysteme in der Primärversorgung. Study on behalf of the Austria Federation of Social Security Institutions. Institute for Advanced Studies, Vienna.

Czypionka T, Kraus M, Kronemann F (2014): [Provision of service in primary health care] Leistungserbringung in der Primärversorgung. Study on behalf of the Austrian Federation of Social Security Institutions. Institute for Advanced Studies, Vienna.

Czypionka T, Berger M, Kronemann F, Reiss M, Sigl C (2014): [Quality management in the outpatient sector] Qualitätsmanagement im ambulanten Bereich - Teil 2. Published in Health System Watch, HSW I-B/2014.

Riedel M, Czypionka T, Kronemann F (2013): [Telemedicine in the Austrian health system] Telemedizin im österreichischen Gesundheitssystem. Study on behalf of T-Mobile. Institute for Advanced Studies, Vienna.