



Universiteit  
Leiden  
The Netherlands

## Case study: labour market reactions to GDP developments

Xinzhu, L.

### Citation

Xinzhu, L. (2017). *Case study: labour market reactions to GDP developments*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master thesis in the Leiden University Student Repository](#)

Downloaded from: <https://hdl.handle.net/1887/3596285>

**Note:** To cite this publication please use the final published version (if applicable).

MATHEMATICAL INSTITUTE

MASTER THESIS

STATISTICAL SCIENCE FOR THE LIFE AND BEHAVIOURAL SCIENCES

---

# Case study: labour market reactions to GDP developments

---

Author:  
Liu Xinzhu

First Supervisor:  
Dr. Shota Gugushvili  
Leiden University

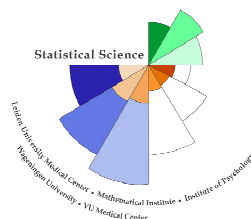
Second Supervisor:  
Prof. dr. Aad van der Vaart  
Leiden University

Third Supervisor:  
Dr. Yvonne Adema  
Centraal Planbureau

January 2017



Universiteit  
Leiden  
The Netherlands



Centraal Planbureau

## Acknowledgements

I wish to express my sincere gratitude to my supervisor Dr. Yvonne Adema for giving me the opportunity to work six months at the CPB. During this precious experience, Yvonne and two other colleagues Rob Luginbuhl and Adam Elbourne provided excellent academic support.

I dedicate the success of my research to my thesis supervisor Dr. Shota Gugushvili. Shota provided a strong guidance throughout my research, and innovated ideas I am not enough of an experienced time series analyst to do myself. Most useful feedback was given by Prof. Aad van der Vaart during the last stage of this work, and I would like to acknowledge Aad for making the task of writing a thesis such an enjoyable experience.

Special thanks to my friends Wouter van Loon, Dora van Oort, and Dick Oosthuizen who expanded my horizon of knowledge. I am especially grateful towards Wouter for his inspirations for programming.

Finally I would like to thank my parents for their support during my years of education. Their constant encouragements have brought me through the hardest time and made me feel confident to pursue my goals.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Aims of this thesis . . . . .	3
1.2	Structure of this thesis . . . . .	4
<b>2</b>	<b>Data description</b>	<b>5</b>
2.1	Data description . . . . .	5
<b>3</b>	<b>Stationarity and stationarity tests</b>	<b>7</b>
3.1	Tests . . . . .	8
3.1.1	Augmented Dickey-Fuller (ADF) test . . . . .	9
3.1.2	Elliott, Rothenberg & Stock Unit Root (ERS) test . . . . .	10
3.2	Sensitivity to lags . . . . .	10
3.2.1	Automatic lags . . . . .	10
3.2.2	OLS estimation . . . . .	10
3.2.3	Pacf and lags . . . . .	11
3.3	Residuals diagnostics: complete versus reduced data . . . . .	12
3.4	Summary . . . . .	13
<b>4</b>	<b>Vector autoregression</b>	<b>15</b>
4.1	VAR lags . . . . .	15
4.2	VAR for the complete versus reduced data . . . . .	16
4.3	VAR stability . . . . .	16
4.4	VAR Residual diagnostics . . . . .	17
4.5	Impulse response analysis . . . . .	21
4.5.1	Orthogonalization by Choleski decomposition . . . . .	21
4.6	Summary . . . . .	23
<b>5</b>	<b>Structural vector autoregression</b>	<b>25</b>
5.1	Parameter restriction and identification . . . . .	25
5.2	Overidentification tests and parameter estimation . . . . .	26
5.2.1	Tests for the complete data . . . . .	27
5.2.2	Tests for the reduced data . . . . .	27
5.3	Structural impulse response analysis . . . . .	29
5.4	Summary . . . . .	31
<b>6</b>	<b>Identification of shocks via a change in volatility</b>	<b>32</b>
6.1	Identifying SVAR via changes in volatility . . . . .	32
6.2	Implementation of the approach . . . . .	33
6.3	Simulation test . . . . .	34
6.4	Parameter estimates . . . . .	39
6.5	Summary . . . . .	39
<b>7</b>	<b>Methods comparison</b>	<b>40</b>
7.1	Predictions: VAR and SVAR . . . . .	40
7.2	Predictions: identification of shocks via a change in volatility . . . . .	41
7.3	Comparison of the predictions through simulation . . . . .	42
<b>8</b>	<b>Discussion</b>	<b>44</b>

<b>9</b>	<b>Appendix</b>	<b>46</b>
	Appendix A Data and descriptive	46
	Appendix B VAR estimation method	48
	Appendix C Estimation of (over-identifying) structural parameters	49
	Appendix D Non-linear estimation of Lutkepohl's parameters	49
	Appendix E Results of simulation test	51
	Appendix F In-sample forecasts	52
	Appendix G Source of R code used in this thesis	54
	G.1 R code Chapter 2 . . . . .	54
	G.2 R code Chapter 3 . . . . .	58
	G.3 R code Chapter 4 . . . . .	62
	G.4 R code Chapter 5 . . . . .	64
	G.5 R code Chapter 6 . . . . .	66
	G.6 R code Chapter 7 . . . . .	73
<b>10</b>	<b>References</b>	<b>77</b>

## Abstract

Vector autoregression (VAR) and structural vector autoregression (SVAR) are economic models that capture the linear interdependency between time series. VAR is purely mechanical in its construction of independent structural residuals, whereas SVAR for the same purpose employs economic content. Another model is built based on VAR and SVAR using the inconsistent variance covariance structures of the residuals. This helps enhance the advantages of these two models while avoiding their limitations.

We study the reaction of the employment rate to the GDP growth rate in the Netherlands. The reaction has a lagged effect, which is modelled by VAR, SVAR, and the built model. Due to the Great Recession, the effect size becomes asymmetrical, which is better operated by the built model than VAR and SVAR. The effectiveness of the built model is perceived in terms of its fitted values.

By fitting the models to data without the recessionary period, the results of VAR and SVAR are consistent with the CPB findings. The employment rate decreases in reaction to shocks to the GDP rate and fluctuates back to equilibrium after three lagged periods. The fluctuation may be caused by the confounding factors that might have imposed opposite forces than the variable reactions.

# 1 Introduction

Economists study the relationship of various labour market variables with GDP, such as employment, unemployment, labour supply etc. During the Great Recession in the late 2000s, an economic decline was observed in world markets [Davis, 2009]. The Centraal Planbureau (CPB) as the economic policy analysis institute for the Netherlands has done abundant research to monitor this recession’s impact on the Dutch labour market. The CPB usually addresses the changes in labour market variables by growth rates, which make economic interpretations accessible. In the CPB’s publication *The Dutch labour market during the Great Recession* [van den Berge *et al.*, 2014], the employment loss due to the Great Recession is shown to be substantial. From 2008 till 2013, the employment loss was approximately 222 thousand full-time equivalents (Figure 1.1), corresponding to an employment growth rate of  $-2.5\%$ .

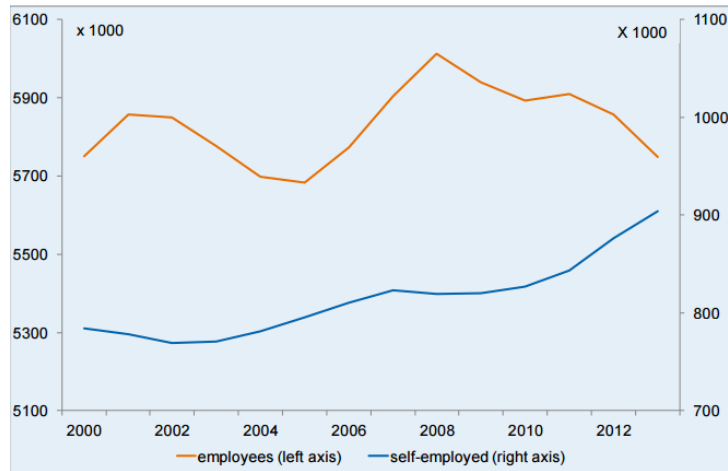


Figure 1.1: Employment of employees and self-employed in full-time equivalents.

In the CPB publication *Roads to Recovery* [Lanser *et al.* 2014], the decline of the Dutch employment rate during the Great Recession was shown to differ from other countries. In the first phase of the recession in 2008 and 2009, the decline was mild and slower than expected whereas most other countries experienced a sharp decline. During the other countries’ recovery phase starting 2011, the Dutch employment continued to decline. The slower decline in the first phase, thus the lag between GDP and employment, was said to be cushioned by labour hoarding <sup>1</sup>, temporary workers and self-employment <sup>2</sup>, and real wages <sup>3</sup>. The continuous decline in the second phase could be explained by the “double dip” <sup>4</sup> in the Netherlands.

There are many economic theories that link GDP and employment. Changes in GDP measure the economic state of a country, and affect the behavior of firms on how many people they want to hire, thus employment. Changes in aggregate employment follow changes in GDP: firms hire workers if the economic situation improves and displace workers if the economic situation deteriorates. However this relationship is not as simplistic as it seems. There may e.g. be a lag in the reaction of aggregate employment to the business cycle due to labour hoarding after a recession or due to recovery of productivity after the recession is over. Moreover, the reaction of firms may depend on the labour market situation (whether or not the labour market was tight

<sup>1</sup>Labour hoarding implies that firms refrain from lay-offs of redundant workers during a recession [CPB, 2014].

<sup>2</sup>These two elements play an opposite role to employment in its reaction to the crisis [CPB, 2014].

<sup>3</sup>Lowered labour demand leads to a trade-off between wage and employment losses CPB, 2014.

<sup>4</sup>Two recessions occurred in the Netherlands in 2008 and 2012.

before the recession) and on their financial situation (labour hoarding might not be feasible if firms do not have the financial means to do so). The question is how the relation between GDP and labour market variables can be modelled and estimated empirically, especially for the Dutch market where the GDP-employment relationship is peculiar.

Okun's law [Okun, 1962] is a classical econometric model that measures the empirical relationship between GDP and unemployment. Several models were developed based on Okun's law, one of which for instance models the relationship between employment and GDP growth rates [Kitov & Kitov, 2011]. However Okun's law is under some criticisms of lacking long-term effect between the two variables [Stock and Watson, 2012]. Additionally several authors have found instability in Okun's law over time [for example, Knotek, 2007].

Akkemik (2007) and Chinn et al. (2014) proposed an error correction model (ECM) that accounts for a long-term relationship between GDP and employment. However the ECM requires the existence of cointegration relations between the variables, that is, a linear combination of the variables must have a long run equilibrium. This is not applicable in the case of labour market growth rates, which usually do not form such a cointegrated trend.

The vector autoregression (VAR) model and the structural vector autoregression (SVAR) model may be good macroeconomics models for the current purpose [see, for example, Raoufinia, 2016; Gosselin Lalonde, 2002]. On the one hand the changes in employment lag behind changes in GDP <sup>5</sup>, on the other hand GDP is also a measure of the total production of the economy which is determined by employment. Therefore the two variables are correlated. One may expect that the lagged values of GDP have an effect on employment rate and perhaps its contemporaneous effects. Likewise, employment rate has impact on GDP and it is hard to conclude about their causality. VAR and SVAR models have been widely used by the CPB to measure the interrelatedness between labour market variables [see, for example, Elbourne and Teulings, 2011]. VAR models are used because VARs are simple to estimate [Elliott and Timmermann, 2008] and a large class of models can be well approximated by a VAR model [Fernandez-Villaverde et al. (2007)]. On top of these advantages, VAR models are developed to incorporate non-statistical economic information [Pfaff, 2008], that is the SVAR models.

## 1.1 Aims of this thesis

In this thesis the relationship between the real GDP growth rate and employment rate is estimated. There are two goals of this project. One is to examine how employment reacts to changes in GDP in the Netherlands by answering:

- How strong is the reaction of employment to changes in GDP;
- How long is the time lag for this reaction of employment to GDP;

Another goal of this project is to produce a model, built on existing models, which will improve the analysis of the size of the reaction and the lag in asymmetrical good and bad economic times. This model will be used to answer:

- How well it produces fitted values and predictions;
- If its fitted values and predictions outperform those of the existing models.

---

<sup>5</sup>Describes the situation where the leading variable GDP is cross-correlated with the values of the lagged behind employment at earlier times.



## 1.2 Structure of this thesis

This thesis is organized as follows. In Chapter 2 a detailed description of the data is provided. Chapter 3 consists of the theoretical background of several unit root tests that are used to test the stationarity of the series. Stationarity tests are necessary because stationarity or otherwise can strongly influence the behavior of the series e.g. persistence of economic shocks will be infinite for nonstationary series. Chapter 4 introduces the vector autoregression model and its relevant tests that are used to handle the doubts about series stationarity. The Chapter ends with a focus on the impulse response function and Cholesky decomposition. In Chapter 5 the structural vector autoregression model is introduced, followed by overidentification tests and parameter restrictions. In Chapter 6, an alternative model for SVAR that compensates the traditional VAR model is described. The model setup is described in steps. Simulation tests are included in this chapter to test the reliability and validity of the implementation. Chapter 7 assesses the prediction values using the two methods: identification of shocks using SVAR and a change in residual volatility. The thesis ends in Chapter 8 with a discussion that answers the research questions by relating the impulse responses to the CPB findings. It also reviews the advantages of the built model over the classical models. The statistical analyses are performed in the R-software environment. The R code for each chapter and some additional information are provided in the appendices.

To sum up, the main body of this thesis consists of VAR, SVAR, and a model built on the basis of them. The comparisons between the three methods are made throughout this thesis mainly by means of their residuals, model fit, and the estimated parameters. The predictions made by these models are used as final comparisons.

## 2 Data description

Economic Co-operation and Development (OECD) and Centraal Bureau voor de Statistiek (CBS) are the major sources of data used in CPB publications. The aggregated economic data from the CBS are less accessible than the OECD data mainly because of strict security rules of the CBS. The economic data on OECD are updated every six months. For this thesis, the latest available data on OECD published on 30 October 2015 is used. Once there is an update in data, the old data can be replaced, and the same models fit on it.

### 2.1 Data description

Data collected from Economic Projections OECD Economic Outlook No 98 <sup>6</sup> is used in this thesis. The data consists of 22 variables <sup>7</sup>, we selected the two variables **Gross Domestic Product Volume Market prices** (nominal GDP) <sup>8</sup> and **Total Employment** <sup>9</sup> for the Netherlands. These two series encompass the years from 1960 to 2017. Since data was published in 2015, years 2016 and 2017 were economic projections and were excluded from the analyses. The function `ts` in package `stats` can be used for visual representations for the two processes, which are provided in Figure 2.1. A numerical summary of the data is provided in Appendix <sup>10</sup>.

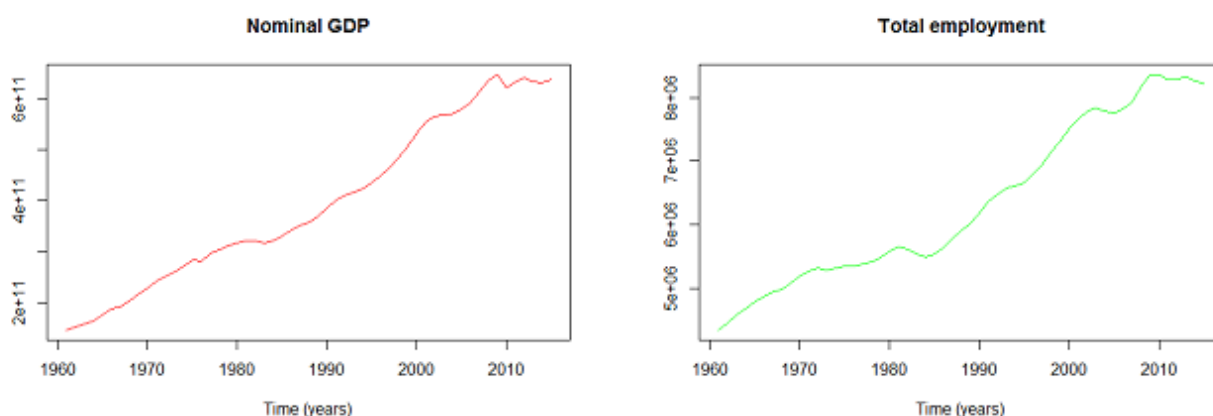


Figure 2.1: Time series for nominal GDP and total employment

As seen in these two plots, nominal GDP continued to increase from 1961 until year 2008, when it had a sharp decrease caused by the Great Recession. At the same time the total employment in the Netherlands decreased only mildly. In the phase when multiple other countries who witnessed recovery around 2011, the decline of the Dutch total employment continued. The data for these two variables are consistent with the CPB publications introduced in the previous chapter.

---

<sup>6</sup><http://stats.oecd.org/>

<sup>7</sup>See G.1.

<sup>8</sup>GDP evaluated at current market prices, GDP being the monetary value of all the finished goods and services produced within a country's borders in a specific time period.

<sup>9</sup>Total employment is defined as the sum of civilian employment and members of the armed forces.

<sup>10</sup>A, Table A.1.

GDP volume market prices	Total employment
Min. :1.467e+11	Min. :4347823
1st Qu.:2.795e+11	1st Qu.:5334329
Median :3.631e+11	Median :5956745
Mean :3.999e+11	Mean :6369475
3rd Qu.:5.668e+11	3rd Qu.:7766083
Max. :6.868e+11	Max. :8505093

Table 2.1: Data summary

### 3 Stationarity and stationarity tests

A time series is a sequence of numerical data points in successive order, which can be expressed as  $(X_1, \dots, X_n)$ . A stationary time series is one whose statistical properties such as mean and variance do not change over time. Stationarity is often a prerequisite for time series analysis. If  $(X_1, \dots, X_n)$  were a draw from unknown distribution, conclusions about the series distribution, as well as tasks such as understanding the system and making predictions, would be impossible [Van der Vaart, 2013].

One of the methods to transform a non-stationary series to a stationary series is through differencing adjacent terms in the series. A time series may have stationary differences when itself is nonstationary. Differencing removes a linear deterministic trend in a non-stationary series [Van der Vaart, 2013]. First-order differencing can also transform a non-stationary series with a stochastic trend to a stationary series. For instance, the time series  $\{x_t\}$  is a random walk if

$$x_t = x_{t-1} + w_t \quad (1)$$

where  $\{w_t\}$  is stationary, for example a white noise series - a stationary series with mean zero and zero autocorrelations. First-order differencing yields:

$$\nabla x_t = x_t - x_{t-1} = w_t \quad (2)$$

The first-order difference of  $\{x_t\}$  produces the stationary series  $\{w_t\}$ .

In Figure 2.1, the two series for the nominal GDP and total employment are non-stationary: they both exhibit an increasing trend. First order differencing was applied to both series, and divided by the value at the former time point to generate new variables **real GDP growth rate**<sup>11</sup> and **employment growth rate**<sup>12</sup>. For convenience these two growth rates will be abbreviated as **GDP** and **employment** for the rest of this thesis unless specified otherwise. The current analysis used data from 1960 till 2015. The differenced data contains GDP and employment from 1961 till 2015. The two series are plotted in Figure 3.1 using function **ts** in package **stats**. Because of the small sample size, the converted two variables are documented in Appendix<sup>13</sup> to clearly show their trend numerically.

GDP	Employment
Min. :-0.03676	Min. :-0.014824
1st Qu.: 0.01651	1st Qu.: 0.005372
Median : 0.02853	Median : 0.014125
Mean : 0.02777	Mean : 0.011902
3rd Qu.: 0.04138	3rd Qu.: 0.021442
Max. : 0.08053	Max. : 0.030896

Table 3.1: Summary for GDP and employment growth rates

As shown in the summary in Table 3.1, when the economic recession struck the Netherlands in 2008, the real GDP growth rate in 2009 decreased by -0.0367614, thus over 3%. Prior to

<sup>11</sup>Real GDP growth rate =  $(\text{Nominal GDP}_t - \text{Nominal GDP}_{t-1}) / \text{Nominal GDP}_{t-1}$

<sup>12</sup>Employment growth rate =  $(\text{Total employment}_t - \text{Total employment}_{t-1}) / \text{Total employment}_{t-1}$

<sup>13</sup>A, Table A.2.

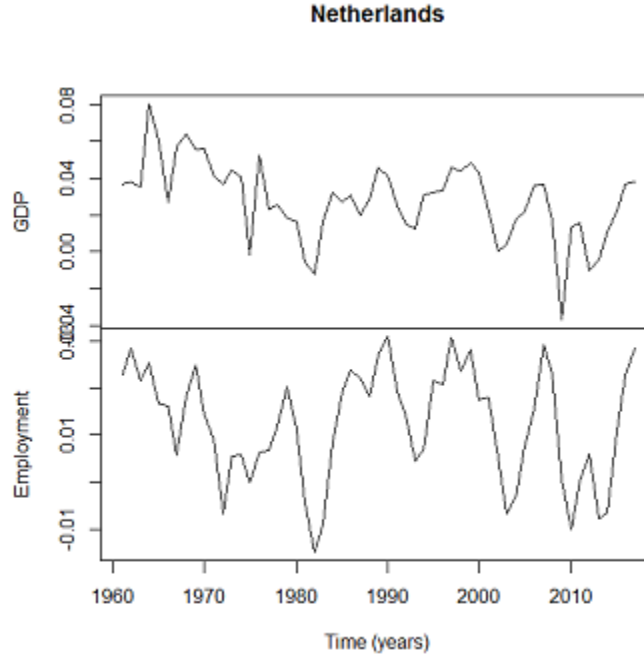


Figure 3.1: Time series for GDP and employment

this recession, the lowest real GDP growth rate can be traced back to the recession in the early eighties. In the plots above, employment does not seem to have a trend. The GDP series has a slight downward trend till the early eighties. After this trend, GDP stabilized until it experienced a sharp drop around 2008. Overall this series does not seem to follow a linear trend. The Great Recession in the late 2000s could posit abnormalities in the time series, demanding special attention to process the data during this period.

### 3.1 Tests

A model for series  $y_t$  is linear if it can be expressed as

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t \quad (3)$$

This is an autoregressive (AR) process  $y_t$  of order  $p$  ( $AR(p)$ ).  $u_t$  is an unobservable white noise process with mean zero and time invariant variance  $E(u_t^2) = \sigma_u^2$  and the  $\alpha_i$  are fixed coefficients that can be estimated by least squares. Denoting by  $L$  the lag operator ( $L$ ) $y_t = y_{t-1}$ , one can write the process (3) more compactly as

$$(1 - \alpha_1 L - \dots - \alpha_p L^p)y_t = u_t \quad \text{or} \quad \alpha(L)y_t = u_t \quad (4)$$

with  $\alpha(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$ . The process is said to be stationary if all roots of the characteristic equation  $\alpha(L) = 0$  exceed unity in absolute value. On the contrary the process is non-stationary if any of the roots is unity. Several tests have been developed for investigating stationarity. They are commonly called unit root tests.

Before proceeding to any unit root tests, one must conclude about the drift (constant) and linear trend of the series. According to the graphical clue in Figure 3.1, after the data transformation, the overall processes do not seem to have a linear trend. GDP fluctuates around 0.02 above zero, and employment around 0.005 above zero. This may suggest a random walk model with drift. The estimation equation including a drift parameter  $v$  then becomes

$$y_t = v + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t \quad (5)$$

### 3.1.1 Augmented Dickey-Fuller (ADF) test

An augmented Dickey-Fuller (ADF) test tests the null hypothesis of a unit root against the alternative of stationarity. If the process (4) is non-stationary, it is integrated at order one  $\alpha(1) = 1 - \alpha_1 - \dots - \alpha_p = 0$ . Assuming no linear trend, we reparameterize the model by subtracting  $y_{t-1}$  on both sides of (3):

$$\begin{aligned} y_t - y_{t-1} &= -y_{t-1} + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t \\ y_t - y_{t-1} &= (-1 + \alpha_1) y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + u_t \\ \Delta y_t &= -\alpha(1) y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + u_t \\ \Delta y_t &= \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + u_t \end{aligned} \quad (6)$$

where  $\phi = -\alpha(1)$  and  $\alpha_j^* = -(\alpha_{j+1} + \dots + \alpha_p)$ . The  $t$ -statistic of the coefficient  $\phi$  is obtained from OLS estimation of (6) to test  $H_0 : \phi = 0$  versus  $H_1 : \phi > 0$  [Fuller (1976) and Dickey Fuller (1979)]. The ADF test on the two processes can be performed using function `ur.df` in package `urca`.

The process  $y_t, t = 1, 2, \dots, T$  can be decomposed into:

$$y_t = r_t + \epsilon_t \quad (7)$$

where  $r_t$  is a random walk, and

$$r_t = r_{t-1} + u_t \quad (8)$$

Assume  $u_t \sim iid(0, \sigma_u^2)$  and  $\epsilon_t$  is a stationary process. If  $\sigma_u^2 = 0$ ,  $r_t$  is a constant,  $y_t$  is composed of a constant and a stationary error  $\epsilon_t$ , and is stationary. Thus,  $H_0 : \sigma_u^2 = 0$  versus  $H_1 : \sigma_u^2 \neq 0$  is equivalent to the test hypothesis.

Now consider  $\epsilon_t$  as autocorrelated errors, in particular AR(1) errors of the form

$$\epsilon_t = \rho \epsilon_{t-1} + \gamma_t \quad (9)$$

with the  $\gamma_t$  iid. Under the null hypothesis that  $u_t = 0$ ,  $y_t$  becomes a random walk when the parameter  $\rho$  approaches unity. As a result, the ADF test inadequately rejects the null hypothesis of unit root of  $y_t$ , depending on how close  $\rho$  is to unity [Kwiatkowski et al., 1992].

### 3.1.2 Elliott, Rothenberg & Stock Unit Root (ERS) test

To improve the power of the unit root test, Elliott, Rothenberg & Stock (1996) proposed to locally detrend or demean the time series, known as a Dickey-Fuller GLS (DF-GLS) test, which takes the serial correlation of  $\epsilon_t$  into consideration. Consider remodelling the parameter  $\rho$  in (9) as  $\rho = 1 - \frac{\bar{c}}{T}$  and filter  $y_t$  with  $1 - \frac{\bar{c}}{T}L$  where  $L$  is the lag operator, i.e.  $\bar{y}_t = y_t - (\frac{\bar{c}}{T})y_{t-1}$ . The test detects unit root by first demeaning the data  $y_t$  into  $\bar{y}_t$ . Working with  $\bar{y}$  gains the unit root test a 50% power. Depending on different specifications of  $r_t$  in (7),  $\bar{c}$  takes different values. This thesis uses predetermined  $\rho = 0.85$  suggested by Hwang and Schmdit (1996) for annual data to optimize the power (as in point optimal P-test).

This is an efficient unit root test because the ERS test has substantial power over the ADF unit root test, especially when the roots of the characteristic equation approach unity. An ERS P-test was performed for the two processes using function `ur.ers` in package `urca`.

## 3.2 Sensitivity to lags

The results of these tests are very sensitive to the number of lags i.e. order of the AR processes included in their respective R functions. It is very likely that some tests will indicate the presence of a unit root, while other will not. A sensible choice in the lag selection is crucial to the test precision. A few generic methods of orders selection are used in this thesis.

### 3.2.1 Automatic lags

These two tests use default lags in the R functions by means of leaving the lags argument empty for `ur.df`. The ADF function chooses the optimal lag by the AIC criterion. R automatically generates output with indications of lag numbers. Function `ur.ers` does not admit a lag argument or generates outputs with lags. The results are summarized in Table 3.2.

GDP	Lag length	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
ADF	1	retain unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root
employment	Lag length	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
ADF	1	reject unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root

Table 3.2: Default lag and results of ADF and ERS tests.

As can be seen under the results for GDP, different tests at significance level  $\alpha = 0.01$  indicate inconsistent presence of unit root and stationarity. The ERS P-test rejects the unit root hypothesis while the ADF test fails to do so. Although the ERS test is with higher power than the ADF test, the inconsistent suggestions of (non)stationarities in the table may necessitate a better input of lags in the tests.

### 3.2.2 OLS estimation

GDP and employment are autoregressive processes of unknown order  $p$ . The OLS estimation approach was motivated by Hamilton (1994) in Chapter 17.7 to estimate (3). A  $t$ -test of the

highest lag  $\alpha_p = 0$  can then be compared with the critical value. If the null hypothesis of this lag being zero is accepted, the  $F$ -test of the joint null hypothesis that both  $\alpha_p = 0$  and  $\alpha_{p-1} = 0$  can be compared with the usual  $F(2, T - k)$  distribution. The procedure continues until the joint null hypothesis that  $\alpha_p = 0, \alpha_{p-1} = 0, \dots, \alpha_1 = 0$  is rejected.

This function was implemented in R <sup>14</sup> with an upper bound  $\bar{p} = 5$  to estimate the optimal lags for the two processes. An optimal number of lags zero was found for both GDP and employment processes when no linear trend but only a constant was included in the models. With  $p = 0$  for both processes, the results of the tests are shown in Table 3.3.

GDP	Lag length	p = 0.01	p = 0.05	p = 0.1
ADF	0	reject unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root
employment	Lag length	p = 0.01	p = 0.05	p = 0.1
ADF	0	retain unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root

Table 3.3: OLS estimation for lags and results of the ADF and ERS tests.

Using this approach and its optimal  $p = 0$ , employment shows inconsistency of the presence of unit root between the two stationarity tests. The optimal  $p = 0$  seems dubious, as it is nearly impossible for GDP and employment to have no lags. Hamilton's approach is probably a heuristic approach that is not optimal in finding lags. This is only a simple technique that one might consider to seek lag lengths.

### 3.2.3 Pacf and lags

Let  $x_t$  be a time series, `acf` computes the correlation between  $x_t$  and itself at different lagged values, and `pacf` computes correlation that results after removing the effect of any correlations due to the terms at shorter lags [Cowpertwait & Metcalfe, 2009]. The correlogram containing (p)acf were plotted for the two series in Figure 3.2 using functions `acf` and `pacf` in package `stats`.

As reported by the Box-Jenkins method (1970), `pacf` plotting has an important role in identifying the AR lags. Referring back to  $AR(p)$  process in (3), the  $k$ th coefficient is the partial autocorrelation at lag  $k$ . For  $AR(p)$ , the coefficients  $\alpha_k$  and the corresponding partial autocorrelations are zero for all  $k > p$ . In Figure 3.2, the two functions for both series have departure from zero at the 95% confidence interval indicated by the dashed lines. GDP's first lag in the partial ACF (top right) plot is significant, indicating an underlying  $AR(1)$  process, corresponding to an optimal lag number zero. Employment's first and second lags in the partial autocorrelation plot are significant, indicating an underlying  $AR(2)$  process, corresponding to an optimal lag of one. Using these lags and repeating all of the unit root and stationarity tests, the results are provided in Table 3.4.

Among the three approaches that select the optimal lags, the current approach gives the seemingly most consistent results of the tests. Recalling the introduction of this section, there is no definite answers as to how to select lags. These three approaches merely give bit of an idea of how the AR order can be.

---

<sup>14</sup>G.2



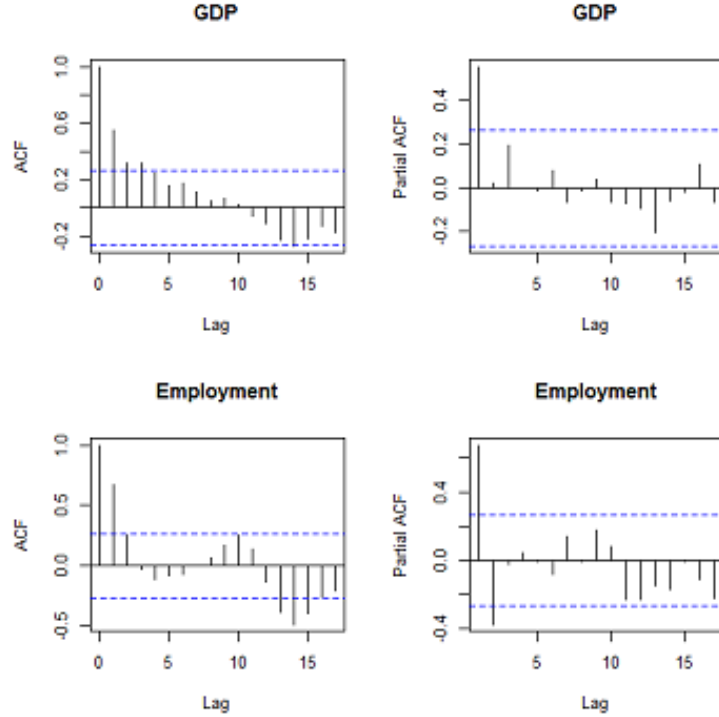


Figure 3.2: acf and pacf for GDP and employment

GDP	Lag length	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
ADF	0	reject unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root
employment	Lag length	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
ADF	1	reject unit root	reject unit root	reject unit root
ERS	/	reject unit root	reject unit root	reject unit root

Table 3.4: Lags determined by pacf and results of the ADF and ERS tests.

### 3.3 Residuals diagnostics: complete versus reduced data

Among the three approaches mentioned in the previous section, the pacf approach produces the most consistent results of stationarity for the tests. Since there is no one definite method for lag selection, one may use pacf to choose the AR order for the two processes and fit (5) into the data. To examine if this regression equation could explain the data well, one may look at  $u_t$ . The residuals for test ADF are accessible using package **CADFtest**. Any violations of the white noise assumptions can be checked by the residuals (p)acf, QQ-plot, serial correlation, and homo(hetero)scedasticity plotted in Figure 3.3 and 3.4.

In Figure 3.3 no white noise assumptions seem to be violated for the employment residuals. The acf and pacf do not exceed the 95% boundary; residuals in the QQ plot form roughly a straight line against their theoretical quantiles; the serial correlations for the neighbouring residuals is approximately zero; and the residuals against the fitted value  $\hat{y}_t$  do not seem problematic. For employment, there are no evident violations against the white noise assumption. However the residuals of GDP in Figure 3.4 exhibit light skewness to the left in its QQ-plot. The variability of residuals seem to increase with the values of  $\hat{y}_t$ . The irregularity of the residuals could

be attributed to two main recessions in the early eighties and 2008 when the volatility of the residual variance became inconsistent.

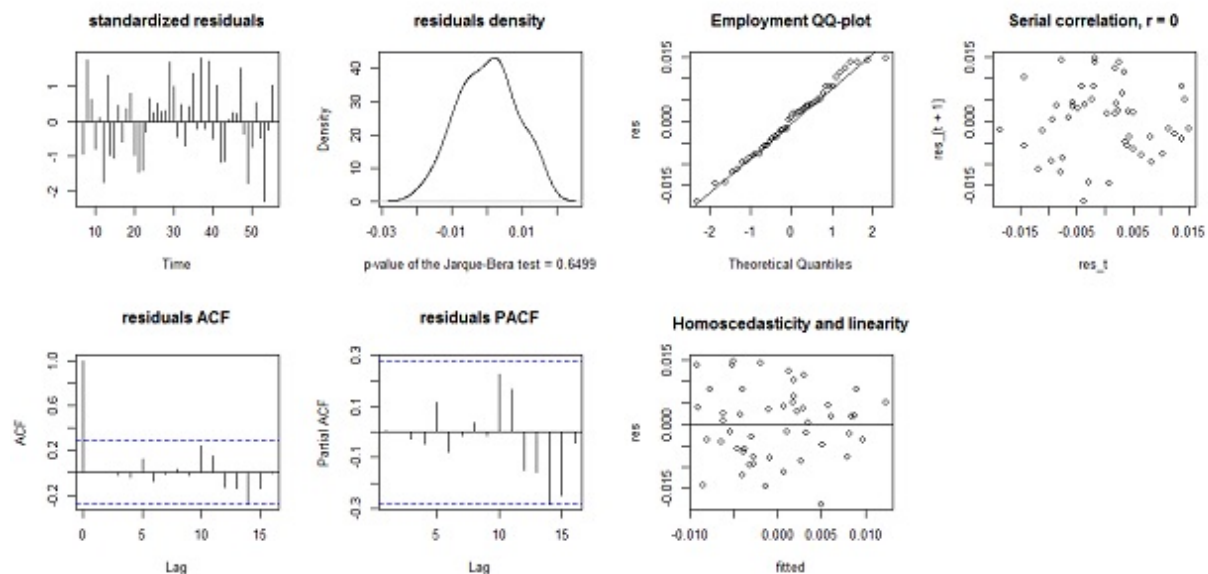


Figure 3.3: Residual diagnostics for employment, 1961-2015.

Data are further split into two parts with the first part containing years 1961 till 2006 before the recession, and the second part containing years 2007 till 2015. The recession in the eighties was not considered as a criterion for data split as to prevent the sample size from becoming small. Using data from years 1961 till 2006 without the data from the recession period might free GDP from residual violations. The same tests with lags determined by the pacf approach were performed again on the pre-recessional data, and residuals were plotted in Figure 3.5. The resulted graphs for years 1961 till 2006 show improvements for the GDP residuals, especially for its residual normality. The increasing residual variability in Figure 3.4 is slightly improved after removing the Great Recession data. The improvements cannot remove the doubts about residuals violation and about the stationarity of the GDP process even with recession data excluded.

### 3.4 Summary

Conclusions about optimal lag and stationarity are impossible due to the inherent problems of the various stationarity/unit root tests and lag selection techniques. Previously in the unit root tests, if the optimal lags were selected based on the pacf plots, the hypothesis of unit root was rejected by ACF and ERS. This result appeared to be consistent for both series in both data sets. Apart from the issues over optimal lag and stationarity, the problematic residuals make the the model fit skeptical. Reducing the data improved the residual conditions, but this eventually does not remove the doubts about stationarity. These techniques simply shed some light on the stationarity and lag conditions for the two series, with no clear-cut conclusions.

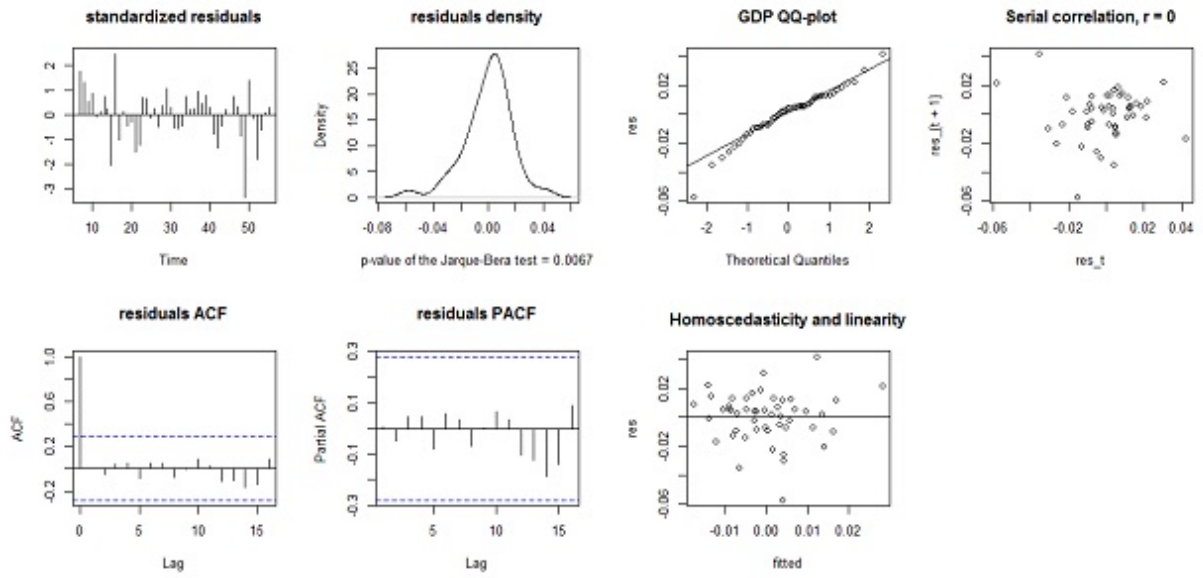


Figure 3.4: Residual diagnostics for GDP, 1961-2015.

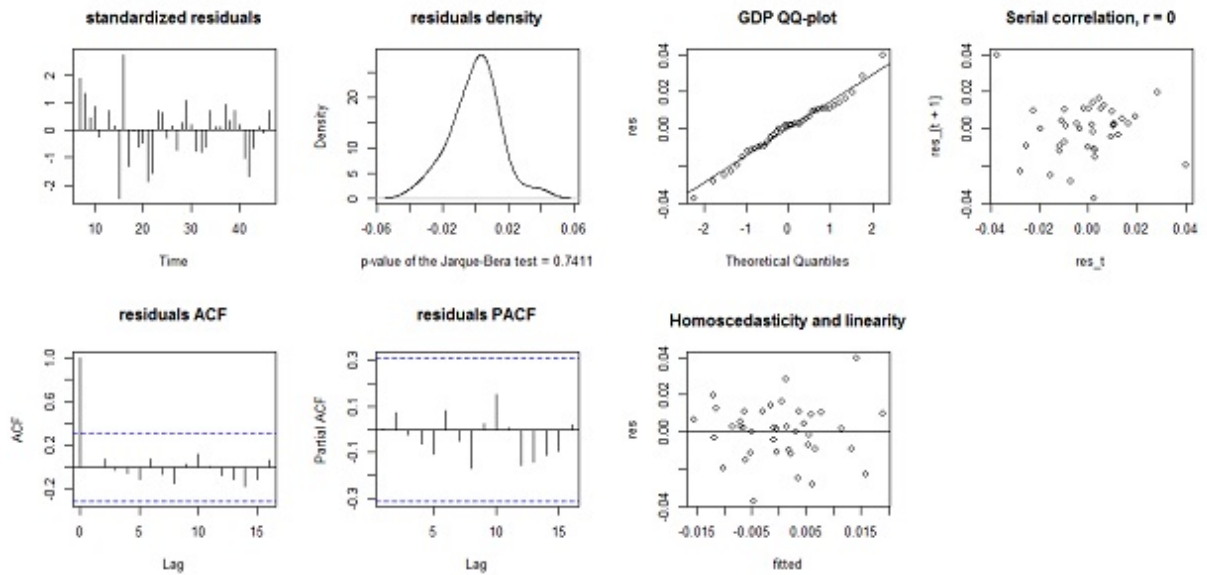


Figure 3.5: Residual diagnostics for GDP, 1961-2006.

## 4 Vector autoregression

From the economics perspective, GDP and employment growth rates cannot be determined as exogenous as they are mutually dependent. It is preferable to model them simultaneously, and for this purpose we consider a vector autoregressive (VAR) process. In the two-variable case, the time path of GDP can be affected by the past realizations of employment and the other way around. Additionally the VAR model allows for several useful tools to examine the relation between the variables, such as the impulse response analysis, which will be introduced later in this chapter.

The basic VAR( $p$ ) process can be written as:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \quad (10)$$

where  $y_t = (y_{1t}, \dots, y_{Kt})'$  is a  $(K \times 1)$  random vector, the  $A_i$  are fixed  $(K \times K)$  coefficient matrices. Finally,  $u_t = (u_{1t}, \dots, u_{Kt})'$  is a  $K$ -dimensional white noise with  $E(u_t) = 0$ ,  $E(u_t u_t') = \Sigma_u$ , and  $E(u_t u_s') = 0$  for  $s \neq t$ . For the current analysis,  $K = 2$ .<sup>15</sup>

### 4.1 VAR lags

A suitable order  $p$  of the VAR( $p$ ) model can be determined from the data by several methods. The function `VARselect` in package `vars` offers the selection criteria: Akaike's information Criterion [Akaike, 1969], Hannan-Quinn Criterion [Hannan and Quinn, 1979], Schwarz Criterion [Schwarz, 1978], and Akaike's final prediction error [Akaike, 1970]. In Table 4.1 and Table 4.2, all of the selection criteria indicate a lag order of three for the data with and without the data from the recessional period. Lag order  $p = 3$  was set for the VAR analysis of the two series.

selection					
AIC(n)	HQ(n)	SC(n)	FPE(n)		
3	3	3	3		
criteria					
	1	2	3	4	5
AIC(n)	-1.7840e+01	-1.8046e+01	-1.8401e+01	-1.8301e+01	-1.8176e+01
HQ(n)	-1.7753e+01	-1.7901e+01	-1.8197e+01	-1.8039e+01	-1.7855e+01
SC(n)	-1.7611e+01	-1.7664e+01	-1.7865e+01	-1.7613e+01	-1.7334e+01
FEP(n)	1.7878e-08	1.4562e-08	1.0240e-08	1.1358e-08	1.2964e-08

Table 4.1: VAR lags for data 1961-2015.

selection					
AIC(n)	HQ(n)	SC(n)	FPE(n)		
3	3	3	3		
criteria					
	1	2	3	4	5
AIC(n)	-1.806700e+01	-1.8183e+01	-1.8441e+01	-1.8352e+01	-1.8214e+01
HQ(n)	-1.7975e+01	-1.8031e+01	-1.8228e+01	-1.8078e+01	-1.7879e+01
SC(n)	-1.7816e+01	-1.7765e+01	-1.7856e+01	-1.7600e+01	-1.7294e+01
FEP(n)	1.4250e-08	1.2706e-08	9.8583e-09	1.0863e-08	1.2627e-08

Table 4.2: VAR lags for data 1961-2006.

<sup>15</sup>Details of estimation method in B.

## 4.2 VAR for the complete versus reduced data

A VAR model was fitted to the complete and reduced data sets with  $p = 3$ . The differences in residuals  $u_t$  generated by VAR using either complete or reduced data can be compared to validate the data split. The improvements of the residuals of VAR could be attributed to the data split, which essentially prevents the VAR model from fitting the Great Recession data. The VAR residuals can be plotted as a comparison to the residual diagnostics of the AR processes. This may provide some ideas of how residuals' improvement is related to the endogeneity of the two variables.

A third-order vector autoregression VAR(3) model incorporating a drift  $v$  was constituted for the complete data, and (10) becomes:

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + u_t \quad (11)$$

This model is fitted to the complete data with the function `VAR` in package `vars`. The VAR model can be summarized as <sup>16</sup>:

$$G_t = 0.0081 + 0.5049G_{t-1} + 0.0401E_{t-1} + 0.0063G_{t-2} - 0.4336E_{t-2} + 0.2375G_{t-3} + 0.2322E_{t-3} + u_{1t}$$

$$E_t = 0.0050 + 0.2566G_{t-1} + 0.5711E_{t-1} - 0.1061G_{t-2} + 0.0627E_{t-2} - 0.1261G_{t-3} - 0.1222E_{t-3} + u_{2t}$$

The same model with  $p = 3$  was fitted to the reduced data. The VAR model can be summarized as:

$$G_t = 0.0119 + 0.4635G_{t-1} + 0.1939E_{t-1} + 0.0200G_{t-2} - 0.3338E_{t-2} + 0.1422G_{t-3} + 0.1265E_{t-3} + u_{1t}$$

$$E_t = 0.0071 + 0.1959G_{t-1} + 0.6074E_{t-1} - 0.1043G_{t-2} + 0.1375E_{t-2} - 0.1244G_{t-3} - 0.2137E_{t-3} + u_{2t}$$

The VAR parameter coefficients provide the best model fitted to the data, but the coefficients of the lagged variables per se contain little interpretable information. The estimated parameters are used to produce good forecasts, but the estimations do not give meanings to the causal relationships of the variables and their past values. The signs of the VAR estimators decide the dynamics at the same time point for the two dimensions, but the coefficients should not be used to directly say anything about the effects at different time points (lags) the two variables have for each other. Our data are observations of the economy, we cannot interpret the parameter estimations the same way as we do it in an experiment. A VAR model is used to fit the data for the parameter estimations, which further constitute analyses that study the dynamic relations between the variables. Therefore the above estimated models are not of main interest in terms of forecasts and interpretations. <sup>17</sup>

## 4.3 VAR stability

An important characteristic of VAR( $p$ ) is its stability. This means VAR( $p$ ) is stationary with time invariant means, variances, and covariances structure (Pfaff, 2006).

---

<sup>16</sup>Estimation method in B.

<sup>17</sup>Conclusion derived from a private discussion with the CPB econometrician Dr. Rob Luginbuhl.

Any VAR( $p$ ) can be written as VAR(1) [Lutkepohl, 2006]:

$$Y_t = \boldsymbol{\nu} + \mathbf{A}Y_{t-1} + U_t \quad (12)$$

where

$$Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad \boldsymbol{\nu} = \begin{bmatrix} v \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_K & 0 & \dots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix}, \quad U_t = \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The characteristic polynomial of  $\mathbf{A}$  is  $\det(\mathbf{A} - \lambda I_K)$  or  $\det(\lambda I_K - \mathbf{A})$ , where  $\lambda$  are the eigenvalues of  $\mathbf{A}$ . The reverse characteristic polynomial is

$$\det(I_K - \mathbf{A}\lambda) = \det(I_K - A_1\lambda - \dots - A_p\lambda^p) \quad (13)$$

if  $\det(I_K - A_1\lambda - \dots - A_p\lambda^p) \neq 0$  for  $|\lambda| \leq 1$ , VAR( $p$ ) is stable. This condition has been satisfied.  $\lambda$  can be found using the function `roots` in package `vars`.

#### 4.4 VAR Residual diagnostics

Several tests were performed on the VAR residuals. The null hypotheses of these tests are: the residuals exhibit no autoregressive conditional heteroscedasticity (ARCH test) [Engle(1982), Hamilton (1994), and Lutkepohl (2006)]; have zero autocorrelations (Portmanteau test) [Ljung Box (1978)]; have skewness and kurtosis being zero (JB normality tests) [Bera and Jarque (1980), (1981), Jarque-Bera (1987), Lutkepohl (2006)]. The R functions `serial.test`, `arch.test`, and `normality.test` in package `vars` can be used for these tests. Results in Table 4.3 and Table 4.4 show that the null hypothesis of the Kurtosis test for the VAR residuals derived from the complete data is rejected at  $p = 0.04$ . This indicates that the probability distribution of the residuals is heavy-tailed and not normal. All of the other hypotheses are retained, that is, the errors are uncorrelated and homoscedastic.

Diagnostics for the residuals are plotted in Figure 4.1 to 4.4. In the **diagram of fit and residuals** for the series in these figures, the top part is the VAR model fit with the observed values indicated by the solid black line and the fitted values by the dashed blue line. The real values of the two variables are adequately captured by the fitted values. The plot underneath shows that the VAR residuals seem to be a white noise series that fluctuates around a red line at  $y = 0$ . In **ACF Residuals** and **PACF Residuals**, none of the acf or pacf have autocorrelations between the residuals and their lagged values, indicating white noise VAR residuals.

Comparing the residuals fitting respectively to complete versus reduced data we see that the GDP residuals have improvements especially with regards to its **QQ-plot**. The **homoscedasticity and linearity** plot has a less dense scatter in the center region. Removing the data from the recession period generates slightly better VAR residuals. For employment, there are no obvious changes in the appearance of the VAR residuals. This is not surprising as the employment series is not as problematic as the GDP series during the recessional period.

Comparing VAR residuals in Figure 4.1 and 4.3 to AR residuals in Figure 3.4 and 3.5, no clear improvements can be observed. The VAR residuals are not essentially improved in terms of

the residual hypothesis. However splitting the data and use only the non-recessional part does improve the model fit for both AR and VAR.

test	statistic	pvalue
Portmanteau Test (asymptotic)	30.15	0.99
ARCH (multivariate)	46.33	0.41
JB-Test (multivariate)	6.70	0.15
Skewness only (multivariate)	0.47	0.788
Kurtosis only (multivariate)	6.22	0.04

Table 4.3: Diagnostic tests of VAR(3) for the complete data.

test	statistic	pvalue
Portmanteau Test (asymptotic)	37.83	0.92
ARCH (multivariate)	57.76	0.09
JB-Test (multivariate)	2.43	0.65
Skewness only (multivariate)	0.76	0.68
Kurtosis only (multivariate)	1.67	0.43

Table 4.4: Diagnostic tests of VAR(3) for the reduced data.

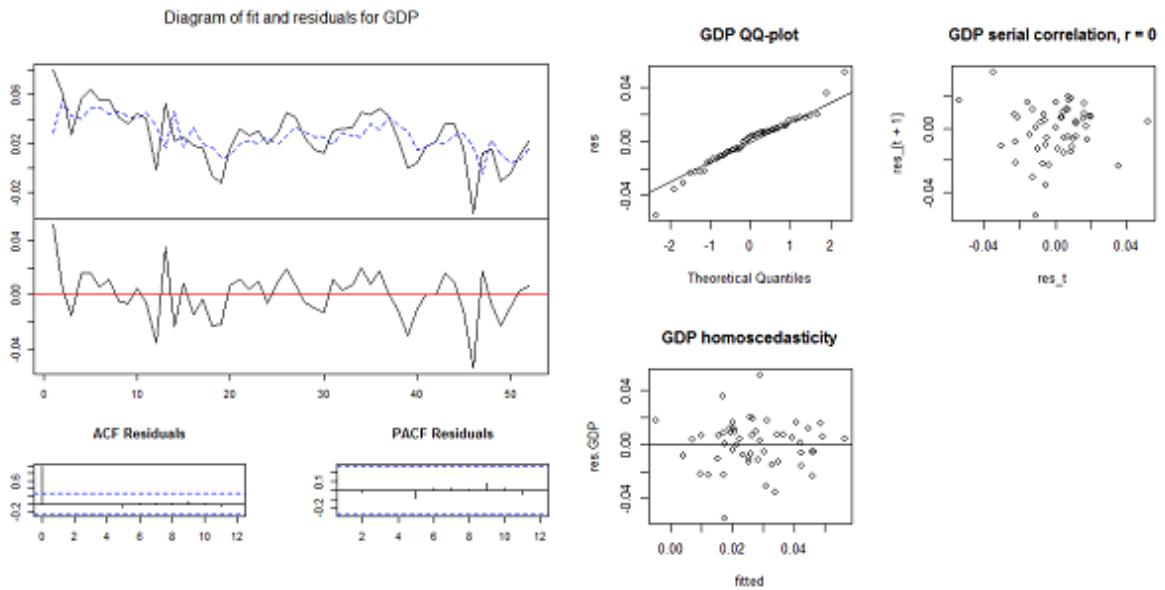


Figure 4.1: VAR fit and residuals diagnostics for GDP (complete data).

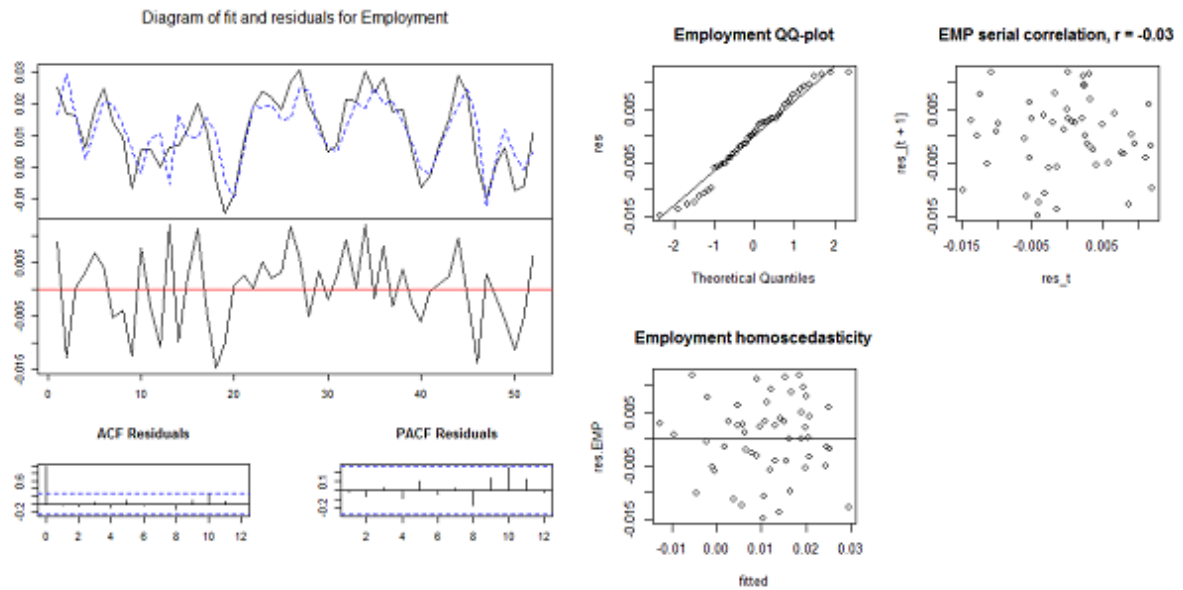


Figure 4.2: VAR fit and residuals diagnostics for employment (complete data).

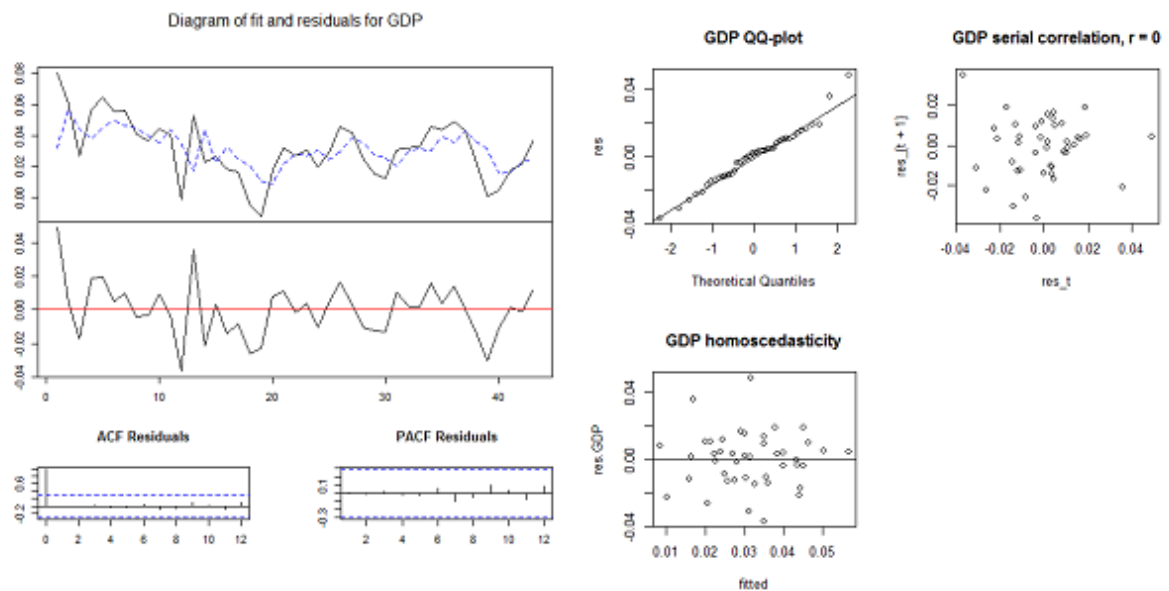


Figure 4.3: VAR fit and residuals diagnostics for GDP (reduced data).



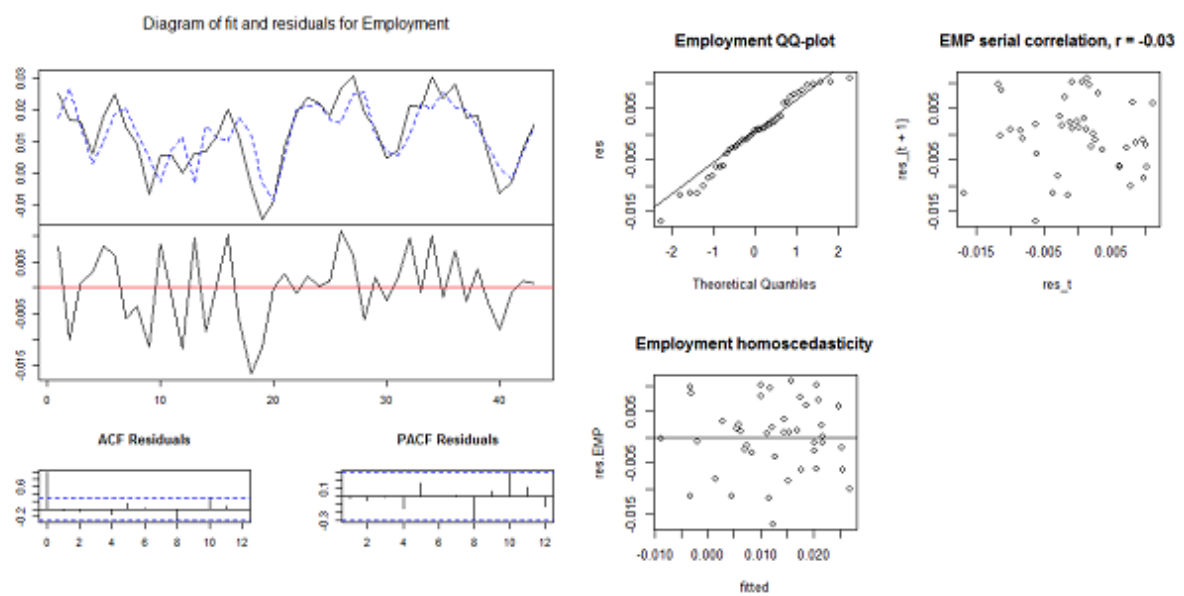


Figure 4.4: VAR fit and residuals diagnostics for employment (reduced data).

## 4.5 Impulse response analysis

The goal of bringing out the VAR system is to find out the interrelatedness of the VAR variables, which can be extrapolated from the autoregressive coefficients of the VAR system. Consider forcing a shock onto the VAR system by changing the residual of one or both variables, for one period, and then returning it to zero thereafter. The time path whereby the variables return to the equilibrium is called the impulse response of the VAR [Greene, 2008]. In the current analysis, impulse response analysis measures how many units in GDP and employment change in reaction to one positive unit change in their residuals  $u_t$  respectively.

In Section 4.3, we know that the VAR(1) representation for VAR( $p$ ) is (12). Under the stability assumption, the process  $Y_t$  has the moving average (MA) representation (see Lutkepohl 2006)

$$Y_t = \mu + \sum_{p=0}^{\infty} \mathbf{A}^p U_{t-p} \quad (14)$$

The vector moving average (VMA) representation for the  $K$ -dimensional VAR process  $y_t$  can be obtained by premultiplying (14) by the  $(K \times Kp)$  matrix  $J := [I_K : 0 : \dots : 0]$

$$\begin{aligned} y_t &= JY_t \\ &= \mu + \sum_{p=0}^{\infty} \Phi_p u_{t-p} \end{aligned} \quad (15)$$

where  $\Phi_p = J\mathbf{A}^p J'$ , and  $u_t$  is a white noise process with covariance matrix  $\Sigma_u$ . For our two-dimensional VAR model, this equation can be further partitioned as

$$y_t = \begin{bmatrix} y_{t,1} \\ y_{t,2} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

The coefficients  $\Phi$  of the VMA representation can be used to generate the effects of the residuals  $u$  on the entire time paths of  $y_t$ . For example, the coefficient  $\Phi$  with  $L = 0$  is the instantaneous impact of a one-unit change in  $u_t$  on  $y_t$ , and  $\Phi$  with  $L = 1$  is the response of a one-unit change on  $u_t$  in  $y_{t+1}$ . The  $\Phi$  are called the impulse response functions (IRF) [Enders, 1995].

### 4.5.1 Orthogonalization by Choleski decomposition

Given that the variables in the VAR system are interrelated to each other, and that the shocks are used to explain such relations, the impulse response analysis assumes that the shocks only happen to one variable at a time [Lutkepohl, 2006]. The error terms  $u_{1t}$  and  $u_{2t}$  are often instantaneously correlated, that is  $\Sigma_u$  is not diagonal. This suggests that a shock in one variable is accompanied by a shock in another variable. Therefore the VAR impulse response analysis is under criticisms: if the VAR residuals  $u_t$  are correlated, the underlying shocks are also correlated. One cannot force  $\Sigma_u$  to be diagonal, as this will obscure the actual relation between the variables [Lutkepohl, 2006].

One solution to isolate the shocks in correlated components of  $u_t$  is Choleski decomposition. Using a Choleski decomposition of the covariance matrix  $\Sigma_u$  generates a lower triangular matrix  $\mathbf{P}$  that satisfies  $\Sigma_u = \mathbf{P}\mathbf{P}'$ . In that case the vector  $\epsilon_t = \mathbf{P}^{-1}u_t$  gives orthogonalized shocks, as  $\Sigma_\epsilon = \mathbf{P}^{-1}\Sigma_u\mathbf{P}^{-1'} = I_K$ . Because the components of  $\epsilon_t$  are orthogonal, a change in one

component has no effect on the other components. Therefore (15) can be transformed into (see Lutkepohl 2004)

$$y_t = \mu + \sum_{p=0}^{\infty} \Phi_p \mathbf{P} \mathbf{P}^{-1} u_{t-p} = \mu + \sum_{p=0}^{\infty} \theta_p \epsilon_{t-p} \quad (16)$$

where  $\theta_p = \Phi_p \mathbf{P}$ . For the current analysis,  $K = 2$  and this equation can be specified as

$$\begin{bmatrix} y_{t,1} \\ y_{t,2} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \theta_{0,11} & \theta_{0,12} \\ \theta_{0,21} & \theta_{0,22} \end{bmatrix} \begin{bmatrix} \epsilon_{t,1} \\ \epsilon_{t,2} \end{bmatrix} + \begin{bmatrix} \theta_{1,11} & \theta_{1,12} \\ \theta_{1,21} & \theta_{1,22} \end{bmatrix} \begin{bmatrix} \epsilon_{t-1,1} \\ \epsilon_{t-1,2} \end{bmatrix} + \dots$$

The  $jk$ -th element of  $\theta_p$  represents the effect on variable  $j$  of a unit change in shock in the  $k$ -th variable that has occurred  $p$  periods ago. For example, a unit change in the shock of GDP given  $p$  results in

$$\begin{aligned} \theta_p &= \begin{bmatrix} \theta_{0,11} & \theta_{0,12} \\ \theta_{0,21} & \theta_{0,22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \theta_{0,11} \\ \theta_{0,21} \end{bmatrix} \\ \theta_p &= \begin{bmatrix} \theta_{1,11} & \theta_{1,12} \\ \theta_{1,21} & \theta_{1,22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \theta_{1,11} \\ \theta_{1,21} \end{bmatrix} \\ &\vdots \end{aligned}$$

A unit change in the shock of employment given  $p$  results in

$$\begin{aligned} \theta_p &= \begin{bmatrix} \theta_{0,11} & \theta_{0,12} \\ \theta_{0,21} & \theta_{0,22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \theta_{0,12} \\ \theta_{0,22} \end{bmatrix} \\ \theta_p &= \begin{bmatrix} \theta_{1,11} & \theta_{1,12} \\ \theta_{1,21} & \theta_{1,22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \theta_{1,12} \\ \theta_{1,22} \end{bmatrix} \\ &\vdots \end{aligned}$$

The orthogonalized impulse response functions with  $\theta_p$  plotted against  $p$  can visually represent the behavior of the series in response to the shocks. Orthogonalized impulse responses of the GDP/employment system are plotted in Figure 4.5, in which the four graphs give the entries of the  $2 \times 2$  matrices  $\theta_p$  as a function of  $p$ . An orthogonalized GDP shock in the bottom-left graph is seen to induce the employment to increase by one period (year) and then it tapers off to zero.

In addition to a lower triangular matrix  $\mathbf{P}$ , many matrices  $\mathbf{P}$  satisfy  $\Sigma_u = \mathbf{P} \mathbf{P}'$ , but there is some degree of arbitrariness when constructing shocks in this manner. Sims (1981) recommends trying different orderings of the variables, the resulting different recursive structures can be used to check the robustness of the results. That is to say, a different ordering of the variables should produce different  $\mathbf{P}$  matrices with different effects of the shocks on the system. Following this idea GDP and employment variables are switched in order, the IRF for this new ordering is plotted in Figure 4.6. It is obvious that the two figures have different impulse responses.

Which set of impulse responses actually reflects the ongoing of the variables is one of the main obstacles in a VAR system [Lutkepohl, 2006]. The non-unique impulse responses leads to Lucas critique (1976), which is translated by Enders (1995) as being mechanical, merely summarizing the dynamics of the data, being devoid of any economic content. The effects of shocks are improperly identified in a VAR system, unless there is a justification from subject matter theory for a specific recursive structure  $\mathbf{P}$  [Enders, 1995].

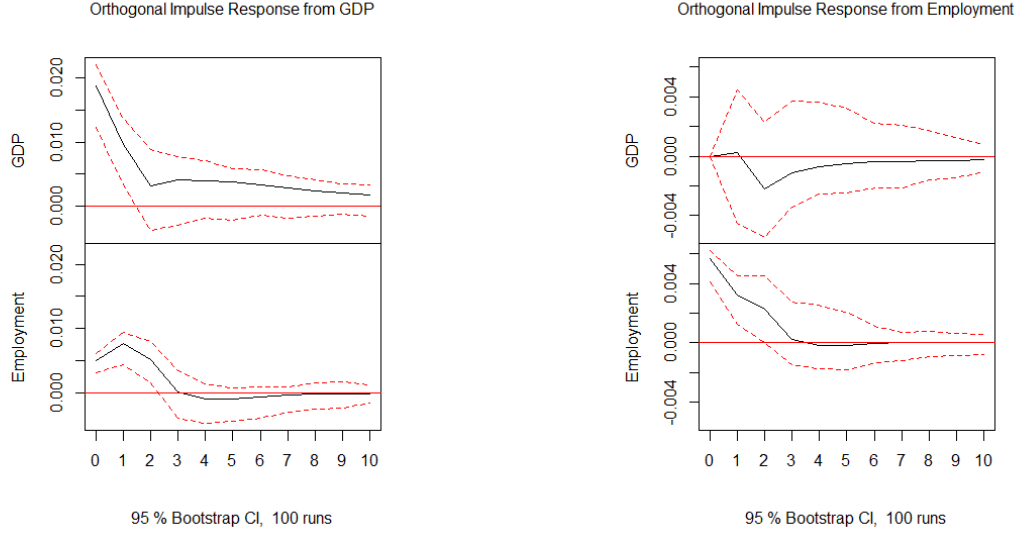


Figure 4.5: Orthogonalized impulse responses of the GDP/employment system. Impulse responses  $\theta_{p,11}$ ,  $\theta_{p,21}$ ,  $\theta_{p,12}$ , and  $\theta_{p,22}$  each corresponds the top-left, bottom-left, top-right, and bottom-right graphs.

Another obstacle of VAR is that sometimes it is not practical to try all alternative orderings as the number of variables increases. Yet this is not of concern for the current setting where  $K = 2$ .

#### 4.6 Summary

Given the unclear conclusion about their stationarity, the two processes are modelled by VAR as an alternative option for the AR models. Unlike AR models, VAR considers the variables simultaneously, which is expected to improve the model fit. However this is not shown in the VAR residuals, at far as the current data is concerned. Excluding the data from the recessionary time improves the model fit, as in the case in of AR modelling. In the impulse response analysis, the underlying shocks are not uncorrelated due to the correlated components of the VAR residuals. To solve this problem, Choleski decomposition is used, but is associated with doubts about its mechanical usage. It is difficult to claim that the impulse response analysis of VAR is able to provide much of a theoretically justified result.

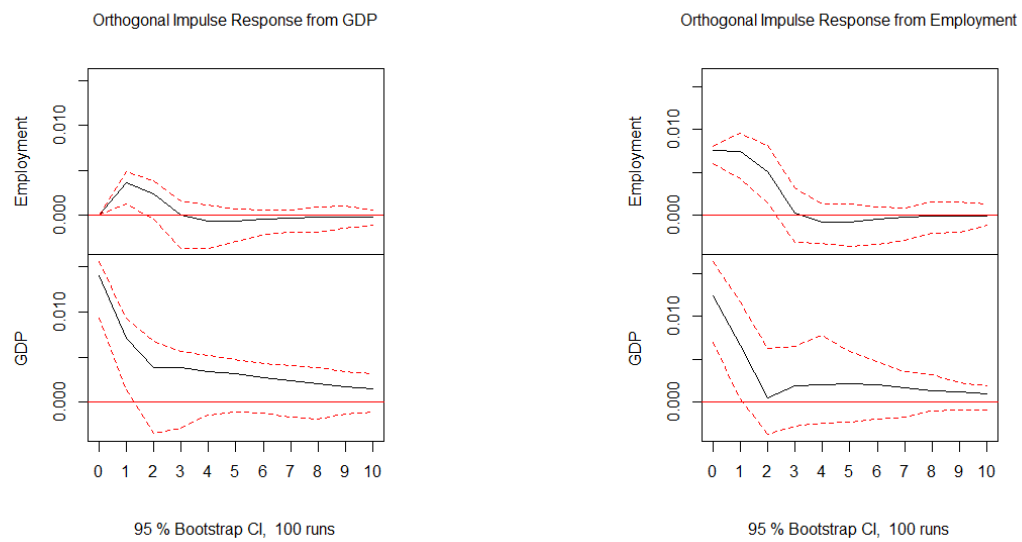


Figure 4.6: Orthogonalized impulse responses of the GDP/employment system using different ordering.

## 5 Structural vector autoregression

The VAR structure produces more than one set of impulse responses and it is not possible to determine which set of impulse responses reflect the ongoing between the variables in a VAR system. An alternative structural vector autoregression (SVAR) model can be considered. SVAR uses nonsample information to specify unique shocks, hence unique impulse responses.

An SVAR is the structural form of VAR. Recall the VAR equations in (10), an SVAR representation is:

$$Ay_t = A_1^*y_{t-1} + \dots + A_p^*y_{t-p} + B\epsilon_t \quad (17)$$

where  $A_j^* = AA_j$  ( $j = 1, \dots, p$ ) and  $\epsilon_t = B^{-1}Au_t$ . The correlated components of the VAR residuals  $u_t$  will be transformed into uncorrelated structural residuals  $\epsilon_t$ , which will have a diagonal covariance matrix  $\Sigma_\epsilon$  with proper choice of  $A$  and  $B$ . There are three conventional models that can be used to make such choices, and below are the representations of these three models made by Lutkepohl (2006):

- A model, for which  $B$  is an identity matrix  $I_K$ ,  $u_t = A^{-1}\epsilon_t$ , and  $\Sigma_u = A^{-1}\Sigma_\epsilon A^{-1'}$ . Normalizing the variances of the structural residuals to one, i.e., assuming  $\epsilon_t \sim (0, I_K)$ ,  $\Sigma_u = A^{-1}A^{-1'}$ . Given  $GDP_t$  and  $employment_t$  on the left-hand side of equation (17), the introduction of  $A$  matrix models the instantaneous relations between the variables directly, and gives additional contemporaneous endogenous variables.

- B model, for which  $A$  is an identity matrix  $I_K$ ,  $u_t = B\epsilon_t$ , and  $\Sigma_u = B\Sigma_\epsilon B'$ . Normalizing the variances of the structural residuals to one, i.e., assuming  $\epsilon_t \sim (0, I_K)$ ,  $\Sigma_u = BB'$ . The introduction of  $B$  matrix simplifies the error structure and make the elements of the error vector uncorrelated. This is a direct approach to orthogonalize the structural innovations  $\epsilon_t$  from the reduced form residuals  $u_t$ . This model and Choleski decomposition differ from each other: the former emphasizes the theoretical justification for the orthogonalization, and the latter has its orthogonalization done by the mechanical application of the Cholesky decomposition. In principle, for an SVAR B model,  $B$  can have other forms of restrictions, the triangular form is just an example.

- AB model, for which  $u_t = A^{-1}B\epsilon_t$  and  $\Sigma_u = A^{-1}B\Sigma_\epsilon B'A^{-1'}$ . Normalizing the variances of the structural innovations to one, i.e., assuming  $\epsilon_t \sim (0, I_K)$ ,  $\Sigma_u = A^{-1}BB'A^{-1'}$ . This model is rarely used in practice.

Assuming the structural innovations uncorrelated, one can take the observed  $\Sigma_u$  to identify the structural parameters  $A$  and/or  $B$  since they are not readily observable. Thereby the underlying structural model can be uncovered from the reduced-formed VAR model, while preserving the assumed independent error structure.

### 5.1 Parameter restriction and identification

The SVAR system generally suffers from the problem of underidentification, because VAR parameters are outnumbered by SVAR parameters. More specifically, the number of nonredundant elements of  $\Sigma_u$  is equal to  $K(K+1)/2$  and it is not possible to identify more than this number of parameters of the structural form. Take A model of SVAR for instance, the total number of elements of  $A$  is  $K^2$ . Therefore  $K^2 - K(K+1)/2 = K(K-1)/2$  would be the minimal

number of restrictions remained to be imposed on the SVAR parameters. Sims (1986) and Bernanke (1986) used economic theories for this restriction, which is regarded as non-recursive identification of shocks. Recall that Choleski decomposition and its recursive orthogonolization of shocks create arbitrariness in the impulse responses. The non-recursive structure imposes the restrictions on the A and/or B matrix and resolves the arbitrary problems.

Three ways to restrict the matrices are as follows:

- For A model, restrictions are placed on A matrix while B is offset to  $I_K$ , with  $K = 2$  in this analysis;
- For B model, restrictions are placed on B matrix while A is offset to  $I_K$ , with  $K = 2$  in this analysis;
- For AB model, restrictions are placed on A and B matrices.

In the current context, the variance covariance matrix  $\Sigma_u$  for GDP and employment is:

$$\Sigma_u = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = A^{-1}B\Sigma_\epsilon B'A^{-1'}$$

Take the example of an A model where B is an identity matrix, and assume  $\epsilon_t \sim (0, I_K)$ :

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1'}$$

The symmetry of  $\Sigma_u$  that  $\sigma_{u_{12}} = \sigma_{u_{21}}$  provides three equations to determine four unknown values. The exact identification of A requires a minimal number of  $K(K-1)/2 = 1$  restrictions placed on A elements to uniquely identify the system. More than one restriction on A would be over-identifying. To implement the SVAR models and their overidentification tests, the elements of the model(s) are set to restriction numbers and free parameters to be estimated. Using the current example, one element of A can be restricted to zero whereas the remaining three elements are free parameters. The output will contain the estimations of these three elements. Note that the restriction should not be arbitrary, but is determined by economic information. If the element that represents the relation between the variables is considered economically insignificant, it should be restricted. The restricted models are the input in function **SVAR** under package **SVAR** that generates estimated model parameters. The overidentification tests and results are available in the summary of the SVAR model.

## 5.2 Overidentification tests and parameter estimation

Section 5.1 explained why SVAR parameter restrictions are necessary. The structural restrictions are made by putting at least  $K(K-1)/2$  numbers of restrictions on the relevant entries in the matrix. The remaining entries are the matrix elements that need to be estimated.

Identifying more than  $K(K-1)/2$  restrictions, is called over-identification. Overidentification suffers from Sims' (1980) criticisms of the "incredible identification restrictions" inherent in structural models. Therefore, a likelihood ratio identification test is computed to test if the overidentifying restrictions are valid. <sup>18</sup>

---

<sup>18</sup>See C for details of the SVAR parameter estimations and overidentification tests.

### 5.2.1 Tests for the complete data

In the literature of Blanchard and Perotti (2002), it is standard to introduce the restrictions on the contemporary matrix  $A$ . For the current data,  $K = 2$ , at least  $K(K-1)/2 = 1$  restriction can be placed on  $A$ . Overidentification tests can be performed on an  $A$  model with three restrictions and two restrictions respectively. In these two tables, the diagonal elements of  $A$  are restricted to be unity, this normalization enables us to write the  $k$ -th equation (17) with  $y_t$  as the left-hand variable. An additional restriction should be imposed on  $A$ . According to Lutkepohl (2006), such restrictions cannot be determined by statistical methods, but have to come from nonsample sources specified by the analyst. Our main interest is how employment reacts to changes in GDP, thus this relationship represented by NA is estimated. Consequently the additional restriction for employment-GDP is zero (Table 5.1). In a two-restriction model (Table 5.2), the relations between the two variables can be estimated in both directions.

A model	GDP	Employment
GDP	1	0
Employment	NA	1

Table 5.1: SVAR three parameters restrictions.

A model	GDP	Employment
GDP	1	NA
Employment	NA	1

Table 5.2: SVAR two parameters restrictions.

Under the null hypothesis that the restrictions are valid, overidentification tests are performed. The null hypothesis is rejected at a 0.05 significance level for two and three restrictions. The results suggest that the over-identifying restrictions are not valid for estimating the SVAR parameters. A just identification model should be carried out.

	GDP	Employment
GDP	1	NA
Employment	NA	NA

Table 5.3: SVAR one parameter restriction.

For a just-identified model, one restriction is placed on GDP's relation with itself. The restriction of  $A$  is illustrated in Table 5.3, its estimating results of  $A$  and  $B$  are summarized in Table 5.4.

### 5.2.2 Tests for the reduced data

All conditions being equal, the same numbers of restrictions were done for the reduced data. The validity of the over-identifying restrictions is rejected at a 0.05 significance level, suggesting invalid three-restriction and two-restriction models. A just-identified model with the same restriction as in Table 5.3 is carried out and its results of  $A$  and  $B$  are summarized in Table 5.5.

The estimated  $A$  and  $B$  parameters in these two tables are used to produce the impulse responses for the SVAR system in the following section.



Estimated A matrix		GDP	Employment
GDP		1.00	130.9
Employment		-71.31	119.3
Estimated standard errors for A matrix		GDP	Employment
GDP		0.000	13.00
Employment		6.994	21.79
Estimated B matrix		GDP	Employment
GDP		1	0
Employment		0	1
Covariance matrix of reduced form residuals (*100)		GDP	Employment
GDP		0.035089	0.009372
Employment		0.009372	0.005693

Table 5.4: SVAR estimations for one restriction, 1961-2015.

Estimated A matrix		GDP	Employment
GDP		1.00	133.9
Employment		-72.67	110.9
Estimated standard errors for A matrix		GDP	Employment
GDP		0.000	14.60
Employment		7.837	23.86
Estimated B matrix		GDP	Employment
GDP		1	0
Employment		0	1
Covariance matrix of reduced form residuals (*100)		GDP	Employment
GDP		0.031215	0.008184
Employment		0.008184	0.005454

Table 5.5: SVAR estimations for one restriction, 1961-2006.

### 5.3 Structural impulse response analysis

Having found the uncorrelated structural shocks and the corresponding  $\mathbf{A}$  and  $\mathbf{B}$  matrices, Lutkepohl (2006) generalized the impulse response coefficients in (15) to obtain the impulse response coefficients for an SVAR model:

$$\theta_p = \Phi_p \mathbf{A}^{-1} \mathbf{B}, \quad p = 0, 1, 2, \dots \quad (18)$$

where  $\Phi_p$  represents the impulse response coefficients for the correlated shocks. For an  $\mathbf{A}$  model, an MA representation based on the structural shocks is

$$y_t = \mu + \sum_{p=0}^{\infty} \Phi_p \mathbf{A}^{-1} \mathbf{A} u_{t-p} = \mu + \sum_{p=0}^{\infty} \theta_p \epsilon_{t-p} \quad (19)$$

where  $\theta_p = \Phi_p \mathbf{A}^{-1}$ . As introduced in Section 4.5.1, impulse response function  $\theta_p$  can be plotted against  $p$  to visually illustrate the behavior of the series in response to shocks. The structural impulse responses for the two variables are plotted in Figure 5.1 and 5.2 with the function `irf` in package `vars`. The input of function `irf` are  $\mathbf{A}$  and  $\mathbf{B}$  obtained from previous section's just-identified SVAR model.

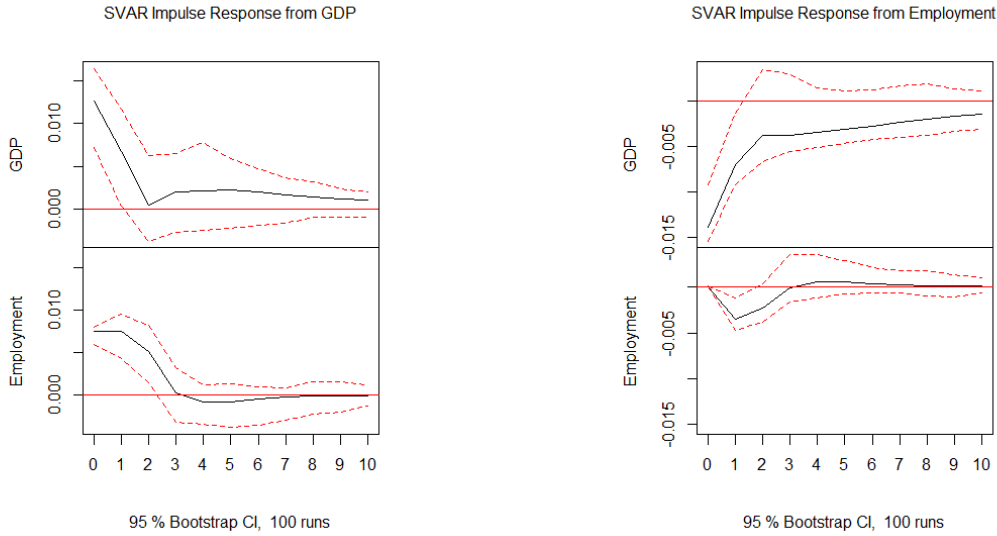


Figure 5.1: SVAR IRF for GDP(left) and employment(right) with restriction on GDP-GDP: 1961-2015.

To interpret the plots, recall that the  $jk$ -th element of  $\theta_p$  represents the effect on variable  $j$  of a unit change in shock in the  $k$ -th variable that occurred  $p$  periods ago and can be written as

$$\theta_p = \begin{bmatrix} \theta_{p,11} & \theta_{p,12} \\ \theta_{p,21} & \theta_{p,22} \end{bmatrix} \begin{bmatrix} \epsilon_{t-p,1} \\ \epsilon_{t-p,2} \end{bmatrix}$$

For the complete data, the left plot of Figure 5.1 shows the responses of the variables to a positive unit GDP growth rate shock. The dotted lines indicate 95% bootstrapped confidence

interval calculated using Hall's method (1992) with 100 replications. One unit shock in GDP at  $p = 0$ ,  $p = 1$ , and  $p = 2$  each causes the value of  $G_t$  to change by 0.0126, 0.0066, and 0.0004 units, and each causes the value of  $E_t$  to change by  $7.5444e - 03$ ,  $7.5460e - 03$ , and  $5.1565e - 03$  units. The right plot of Figure 5.1 shows the responses of the variables to a positive unit employment growth rate shock. One unit shock in employment at  $p = 0$ ,  $p = 1$ , and  $p = 2$  each causes the value of  $G_t$  to change by  $-0.0138$ ,  $-0.0069$ , and  $-0.0038$  units, and each causes the value of  $E_t$  to change by  $1.0579e - 04$ ,  $-3.4919e - 03$ , and  $-2.3111e - 03$  units.

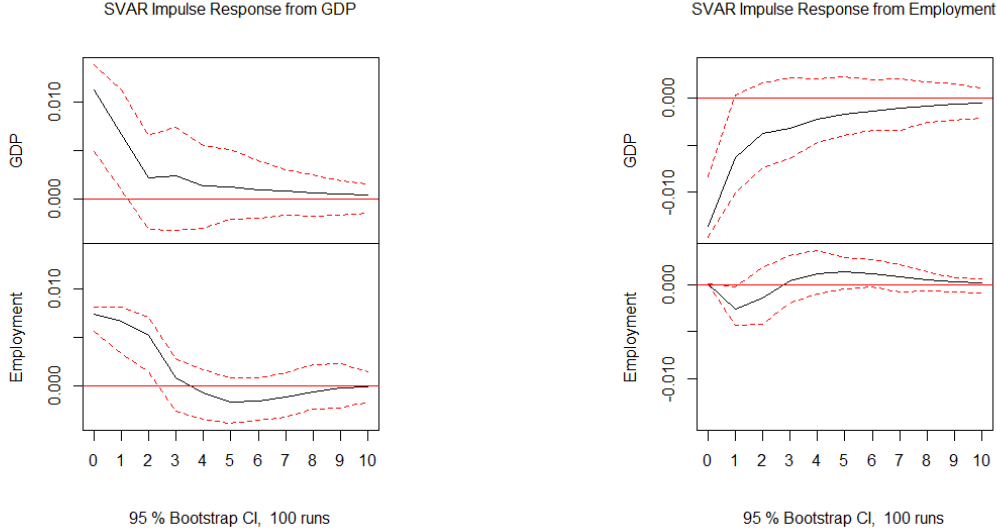


Figure 5.2: SVAR IRF for GDP(left) and employment(right) with restriction on GDP-GDP: 1961-2006

For the reduced data, the left plot of Figure 5.2 shows the responses of the variables to a positive unit GDP growth rate shock. One unit shock in GDP at  $p = 0$ ,  $p = 1$ , and  $p = 2$  each causes the value of  $G_t$  to change by 0.0112, 0.0066, and 0.0021 units, and each causes the value of  $E_t$  to change by  $7.384390e - 03$ ,  $6.692443e - 03$ , and  $5.208862e - 03$  units. The right plot of Figure 5.2 shows the responses of the variables to a positive unit employment growth rate shock. One unit shock in employment at  $p = 0$ ,  $p = 1$ , and  $p = 2$  each causes the value of  $G_t$  to change by  $-0.0136$ ,  $-0.0062$ , and  $-0.0037$  units, and each causes the value of  $E_t$  to change by 0.0001,  $-0.0026$ , and  $-0.0013$  units.

According to the estimated SVAR model for the complete data, a positive shock to the GDP growth rate increases the employment growth rate immediately, then increases it for about two years with a maximal response after one year ( $7.5460e - 03$ ), and gradually drives down the employment growth rate. According to the estimated SVAR model for the reduced data, a positive GDP growth rate shock increases the employment growth rate immediately, and the increase lasts for about two years until the effect gradually returns to the baseline. An employment growth rate shock for both data sets decreases GDP growth rate immediately until it returns to baseline. The responses are in line with the economic theory: when recessions occur, the employment growth rate does not decrease immediately due to the lagging effect. Furthermore, the contrast between the estimated SVAR model fitted to the complete and reduced data seems to precisely capture the dynamics of GDP-employment growth rates proposed by the CPB findings (2014). When the recession is absent, the employment rate should return to baseline with no

fluctuations. When the recession is present, the employment growth rate increases within one year after the recession probably due to the economic factors such as labour hoarding and so on explained in Chapter one. The growth rate drops afterwards, probably showing a larger scale of impact from the recession and a smaller effect from these confounding factors. However take into the consideration that the CPB analysis used a different set of data, thus any comparisons formed are not strictly justified.

## 5.4 Summary

VAR models are not ideal regarding their recursive structure of restrictions and arbitrary generation of uncorrelated shocks. As a consequence the impulse responses of the VAR systems are not unique. Non-recursive restrictions for the SVAR models are made in this analysis with an A model that directly models the instantaneous relations between GDP and employment. The overidentifying restrictions on A matrix are not valid, necessitating just-identified models for both the complete and reduced data. Parameters in A and B are estimated and form the structural shocks. Finally the impulse responses are formed for SVAR, and compared with the CPB findings. Once again splitting the data is necessary, which is evident in their impulse responses: the impulse responses generated by an SVAR model fitted to the complete data resemble the findings of the CPB, and this is not the case with the reduced data. When recession is present, the Dutch employment rate could be influenced by an opposite force imposed on it by the GDP rate and other factors. If the recession is absent, such opposite force is not expected to exist.

## 6 Identification of shocks via a change in volatility

SVAR restrictions are essential for unique parameter identifications. In the previous section, the structural parameters in  $A$  and  $B$  were estimated so that the structural shocks  $\epsilon_t$  are linearly transformed from  $u_t$ . The models must be restricted with a minimal number of restrictions, or this process suffers from underidentification i.e., SVAR parameters outnumber VAR parameters. However over-identifying restrictions are not necessarily valid and must be tested by the overidentification tests, which complicate the estimation process.

Furthermore, another problem is associated with the just-identified SVAR system. Although the relations between the two variables are justified by economic theories, information may still be lost through the restrictions of the impact matrices  $A$  and  $B$ . In Section 5.2.1, one parameter restriction in Table 5.3 resulted in the loss of information on the relation between GDP and itself.

Lutkepohl (2012) proposed another approach to reveal the uncorrelated unique shocks. His approach frees the identifying process from the problems mentioned above. The shocks identification utilizes the change in the series volatility, that is, the change before and after the Great Recession.

### 6.1 Identifying SVAR via changes in volatility

According to Lutkepohl (2012), another possibility for specifying the unique shocks is available when the  $u_t$  distribution, and thus the  $\epsilon_t$  distribution, changes during the sample period. For  $t = 1, \dots, T_1$  from the first part of the sample,  $E(u_t u_t') = \Sigma_1$ , and for  $t > T_1$ ,  $E(u_t u_t') = \Sigma_2$ , and  $\Sigma_1 \neq \Sigma_2$ . In the current analysis the first part of the sample consists of observations from 1961 till 2006, and the second part consists of observations from 2007 till 2015. The distinct covariance matrices derived from the two samples increase the number of parameters for VAR.

Suppose all VAR parameters  $A_p$  are the same in two periods, the residual covariance matrices for the two samples can be decomposed into a matrix  $B$  and a diagonal matrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_K)$ ,

$$\Sigma_1 = BB' \quad \text{and} \quad \Sigma_2 = B\Lambda B'. \quad (20)$$

To uncover the structural shocks from the reduced form residuals, we have the linear transformation  $\epsilon_t = B^{-1}u_t$ , which are instantaneously uncorrelated as

$$E(\epsilon_t \epsilon_t') = \begin{cases} I_K, & t = 1, \dots, T_1 \\ \Lambda, & t > T_1 \end{cases}$$

The current bivariate system has

$$u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

The relations in (20) become

$$E(u_t u_t') = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} E(\epsilon_t \epsilon_t') \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}'$$

$$\begin{aligned}
\begin{bmatrix} \sigma_{1,1}^2 & \sigma_{12,1} \\ \sigma_{12,1} & \sigma_{2,1}^2 \end{bmatrix} &= \begin{bmatrix} b_{11}^2 + b_{12}^2 & b_{11}b_{21} + b_{12}b_{22} \\ b_{11}b_{21} + b_{12}b_{22} & b_{21}^2 + b_{22}^2 \end{bmatrix} \\
\begin{bmatrix} \sigma_{1,2}^2 & \sigma_{12,2} \\ \sigma_{12,2} & \sigma_{2,2}^2 \end{bmatrix} &= \begin{bmatrix} \lambda_1 b_{11}^2 + \lambda_2 b_{12}^2 & \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} \\ \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} & \lambda_1 b_{21}^2 + \lambda_2 b_{22}^2 \end{bmatrix} \\
\sigma_{1,1}^2 &= b_{11}^2 + b_{12}^2 \\
\sigma_{12,1} &= b_{11}b_{21} + b_{12}b_{22} \\
\sigma_{2,1}^2 &= b_{21}^2 + b_{22}^2 \\
\sigma_{1,2}^2 &= \lambda_1 b_{11}^2 + \lambda_2 b_{12}^2 \\
\sigma_{12,2} &= \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} \\
\sigma_{2,2}^2 &= \lambda_1 b_{21}^2 + \lambda_2 b_{22}^2
\end{aligned}$$

With  $K = 2$ , the six equations above would have a unique (apart from ordering) set of solutions for  $b_{11}, b_{12}, b_{21}, b_{22}, \lambda_1, \lambda_2$ , if  $\lambda_1 \neq \lambda_2$  with a specific order e.g.  $\lambda_1 < \lambda_2$  [Lanne et al., 2010]. Since the first sample has variance  $I_K$ , the variance  $\Lambda$  of the second sample is the variance relative to the first sample. This means that the requirement of  $\lambda_1 \neq \lambda_2$  can be satisfied when the changes in the variance of the two variables are not the same. In fact only the residual variance of one variable has to change [Lutkepohl, 2012]. We splitted the sample in Chapter 4 and saw slightly different variability of the GDP residuals while the employment residuals seem to remain homoscedastic. The VAR residual plots Figure 4.1 till 4.4 give a rough idea regarding the variance status across the two variables.

As mentioned earlier in this section, Lutkepohl's approach expands the SVAR model and the VAR model it is based on. Compared to the SVAR system, Lutkepohl's approach expands the covariance matrices with additional parameters  $b$  and new parameters  $\Lambda$ . Parameter restrictions, whether over-identifying or just-identifying the system, are deemed unnecessary. This prevents the loss of economic content caused by any parameter restrictions. Once the ordering of the variables is fixed, Lutkepohl's approach utilizes the heterogeneous changes across the VAR residuals to uniquely characterizes the structural shocks. Besides these contributions, Lutkepohl's two covariance matrices are both diagonal, the structural shocks are instantaneously uncorrelated.

## 6.2 Implementation of the approach

Using the equation wise OLS in (11), the residuals  $u_t$  are obtained. Assume  $u_t \sim N(0, \Sigma_m)$ ,  $\Sigma_m$  are the  $u_t$  covariance matrices of different samples i.e., time regimes  $\mathcal{T}_m$

$$\hat{\Sigma}_m = \frac{1}{T_m - T_{m-1}} \sum_{t \in \mathcal{T}_m} \hat{u}_t \hat{u}_t', \quad (21)$$

where  $T_m - T_{m-1}$  is the number of observations in sample  $\mathcal{T}_m$ . Next the estimates  $\hat{\Sigma}_m$  are used in a GLS estimator

$$\hat{\beta} = \left( \sum_{t=1}^T Z_{t-1} Z_{t-1}' \otimes \hat{\Sigma}_m^{-1} \right)^{-1} \left( \sum_{t=1}^T (Z_{t-1} \otimes \hat{\Sigma}_m^{-1}) y_t \right), \quad (22)$$

where  $Z_{t-1} = (1, y'_{t-1}, \dots, y'_{t-p})$ .  $\hat{\beta}$  contains the estimates for the time-invariant VAR parameters in the form of  $\text{vec}[\nu, A_1, \dots, A_p]$ .

$\hat{\beta}$  is then used to substitute  $y_t - (Z'_t \otimes I_K)\hat{\beta}$  for  $u_t$  in the maximum likelihood

$$\log \ell(\beta, \sigma) = -\frac{KT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |\Sigma_m| - \frac{1}{2} \sum_{t=1}^T u'_t \Sigma_m^{-1} u_t, \quad (23)$$

where  $\Sigma_m$ , are replaced with  $BB'$  and  $B\Lambda_m B'$ . Estimation of the parameters  $B$  and  $\Lambda$  can be done by the concentrated maximum likelihood optimized with respect to  $b$  and  $\lambda$ .

The MLE of  $B$  and  $\Lambda$  obtained this way can be used to substitute  $\Sigma_m$  to produce  $\beta$ . Therefore the estimation process iterates until the values for  $b$ ,  $\lambda$ , and  $\beta$  converge.<sup>19</sup>

### 6.3 Simulation test

Simulation test was done to validate this optimization function. A two dimensional  $u_t$  was simulated and incorporated with a fixed set of  $\beta$  parameters to simulate data. From the simulated  $u_t$  the covariance structures in (21) can be formed. From here on the iterative algorithm is done to generate the converged  $b$ ,  $\lambda$ , and  $\beta$ .  $\beta$  is compared with the true parameters, i.e. the fixed set of  $\beta$  used to initiate the simulation. Repeating the simulation a sufficient number of times, the estimates of  $\beta$  are collected at the end of each simulation.

Each simulation of varying sample size were tested on the optimization function. Increasing the size of the sample should decrease the difference between the true parameter and the mean of the parameter estimates. Asymptotically, the standard deviations of the parameter estimates should decrease to zero as the sample size goes to infinity.<sup>20</sup> The parameters estimated by these simulation trials are compared with the true parameters. Such simulations help explore how sample size influences to the estimation accuracy.

With the sample size increases and the number of simulations remains the same, standard deviations for all 14 parameter estimates decrease. Increasing sample size results in higher accuracy of estimation produced by this optimization function. Therefore inputs of large sample sizes are preferred for this optimization function for relatively accurate parameter estimations.

An accurate result derives from a combination of large sample size (300) simulated 1000 times. The distribution of the simulation estimates are plotted against the 14 true parameter values (marked in red) in Figure 6.1 till 6.4. The global maxima of the estimations' distributions all adequately capture the true parameters.

<sup>19</sup>A detailed description of this estimation procedure can be found in D.

<sup>20</sup>In practice, we tried a sample of size 55, which was simulated 250 times, and a sample of size 300 simulated 250 times respectively. Test results can be found in Table E.1 and Table E.2 in E.

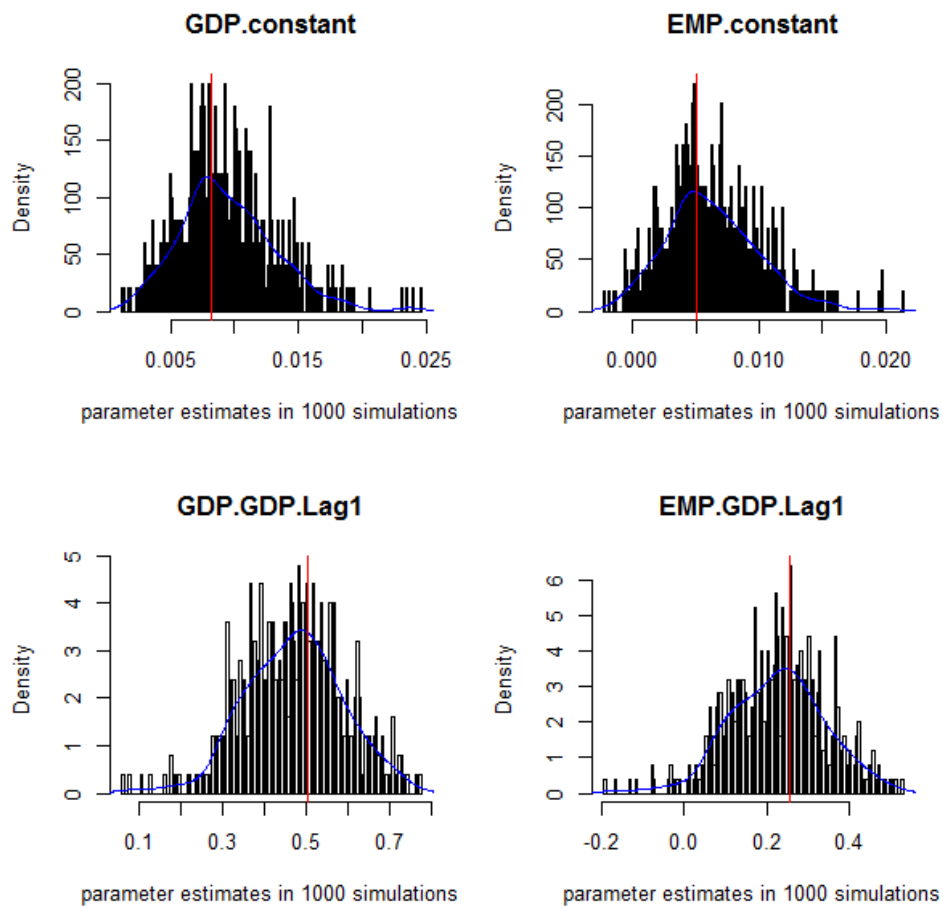


Figure 6.1: Parameter estimates and true parameters for GDP.constant, EMP.constant, GDP.GDP.Lag1, EMP.GDP.Lag1. Sample size 300, 1000 simulations.



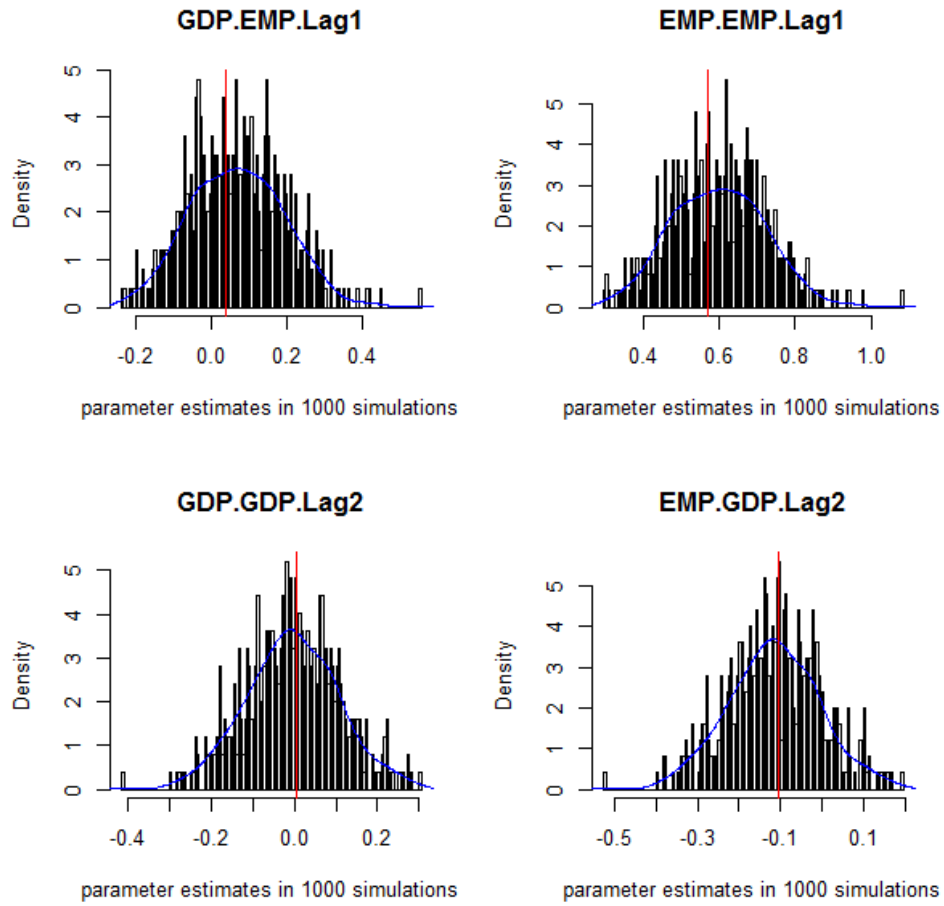


Figure 6.2: Parameter estimates and true parameters for GDP.EMP.Lag1, EMP.EMP.Lag1, GDP.GDP.Lag2, EMP.GDP.Lag2. Sample size 300, 1000 simulations.

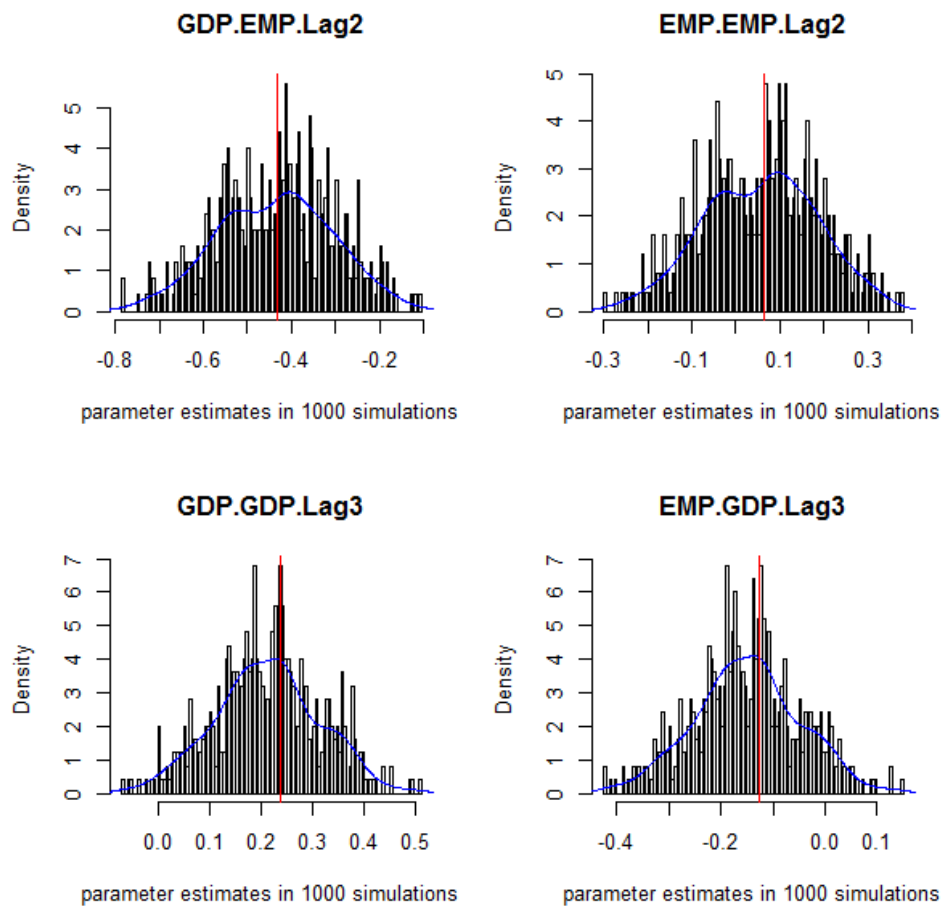


Figure 6.3: Parameter estimates and true parameters for GDP.EMP.Lag2, EMP.EMP.Lag2, GDP.GDP.Lag3, EMP.GDP.Lag3. Sample size 300, 1000 simulations.

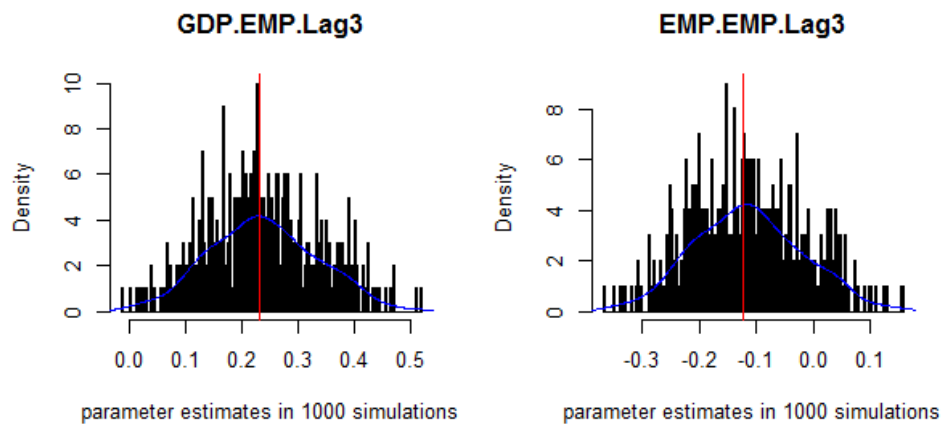


Figure 6.4: Parameter estimates and true parameters for GDP.EMP.Lag3, EMP.EMP.Lag3. Sample size 300, 1000 simulations.

## 6.4 Parameter estimates

Having proved the reliability of the optimization function, the parameter estimates for the data are estimated and listed in Table 6.1.

VAR parameter estimates	
GDP.constant	0.0094732
EMP.constant	0.0059390
GDP.GDP.Lag1	0.4942354
EMP.GDP.Lag1	0.2304458
GDP.EMP.Lag1	0.1007533
EMP.EMP.Lag1	0.5877956
GDP.GDP.Lag2	0.0085358
EMP.GDP.Lag2	-0.1052553
GDP.EMP.Lag2	-0.3881558
EMP.EMP.Lag2	0.0971135
GDP.GDP.Lag3	0.1981883
EMP.GDP.Lag3	-0.1263884
GDP.EMP.Lag3	0.1943895
EMP.EMP.Lag3	-0.1638035

Table 6.1: Optimized parameter estimates for  $\beta$  for the data.

Therefore the VAR representation is:

$$G_t = 0.0094 + 0.4942G_{t-1} + 0.1007E_{t-1} + 0.0085G_{t-2} - 0.3881E_{t-2} + 0.1981G_{t-3} + 0.1943E_{t-3} \quad (24)$$

$$E_t = 0.0059 + 0.2304G_{t-1} + 0.5877E_{t-1} - 0.1052G_{t-2} + 0.0971E_{t-2} - 0.1263G_{t-3} - 0.1638E_{t-3} \quad (25)$$

The estimated parameters for  $b$  and  $\lambda$  are reported in Table 6.2.

b parameter estimates		$\lambda$ parameter estimates	
9.579433e-03	0.001170346	1.904369	0.000000
-9.090616e-05	-0.005835668	0.000000	2.274337

Table 6.2: Optimized parameter estimates for  $b$  and  $\lambda$

## 6.5 Summary

Lutkepohl's approach exploits heteroscedasticity of the residual variances. By subsetting the VAR residuals in two parts, two  $\Sigma_u$  are obtained and lead to the increase of the VAR parameters. Any additional restrictions on the SVAR parameters will become over-identifying. Different volatility changes also make it possible for the parameters to be uniquely identified, as long as the ordering of the variables are fixed. Shocks obtained this way are instantaneously uncorrelated. The optimization and simulation are done for this approach, with the conclusion that this technique becomes accurate when the sample size becomes large.

## 7 Methods comparison

By means of comparing the fitted values, one can identify the better approach between identifying shocks via SVAR versus changes in volatility. (S)VAR models generate different sets of autoregressive parameters than Lutkepohl's approach, the fitted values of the two approaches differ. In econometrics, the fitted values are referred to as in-sample forecasts, whereas the predictions into the future are termed as out-of-sample forecasts. Lutkepohl's approach aims to reveal the structural shocks using two regimes, which might improve both in-sample and out-of-sample predictions compared to the VAR and SVAR predictions.

### 7.1 Predictions: VAR and SVAR

The in-sample forecast are fitted values of the VAR model. The out-of-sample forecast can be achieved by predicting the short-term behaviour of the variables of the VAR model [Lutkepohl, 2004]. At forecast origin  $T = 2015$ , an  $h$ -step ahead forecast is obtained recursively through:

$$y_{T+h|T} = A_1 y_{T+h-1|T} + \dots + A_p y_{T+h-p|T} \quad (26)$$

starting with  $h = 1$ . The MSE matrix of an  $h$ -step forecast is

$$\Sigma_y(h) = E(y_{T+h} - y_{T+h|T})(y_{T+h} - y_{T+h|T})' = \Sigma_{j=0}^{h-1} \Phi_j \Sigma_u \Phi_j' \quad (27)$$

Based on the previous analysis for the two variables, we assume that  $u_t \sim \text{iid } N(0, \sigma_u^2)$ , and then the forecast errors are also multivariate normal. The forecast intervals can be established as:

$$[y_{k,T+h|T} - c_{1-\alpha/2} \sigma_k(h), y_{k,T+h|T} + c_{1-\alpha/2} \sigma_k(h)] \quad (28)$$

$y_{k,T+h|T}$  denotes the  $k$ th component of  $y_{T+h|T}$ , and  $\sigma_k(h)$  denotes the square root of the  $k$ th diagonal component of  $\Sigma_y(h)$ , that is,  $\sigma_k(h)$  is the standard deviation of the  $h$ -step forecast error for the  $k$ th component of  $y_t$ . This out-of-sample prediction can be implemented using the R function `predict` in package `vars`, the input of the function is the VAR model. The predictions and their confidence intervals for the following years 2016-2020 are recorded in Table 7.1, where the forecast step is  $h = 5$ .

	GDP forecast	lower	upper	CI	EMP forecast	lower	upper	CI
2016	0.0197	-0.0169	0.0565	0.0367	0.0168	0.0021	0.0316	0.0147
2017	0.0152	-0.0260	0.0565	0.0413	0.0175	-0.0044	0.0395	0.0220
2018	0.0171	-0.0248	0.0591	0.0419	0.0138	-0.0108	0.0384	0.0246
2019	0.0185	-0.0242	0.0613	0.0428	0.0122	-0.0124	0.0368	0.0246
2020	0.0198	-0.0236	0.0633	0.0435	0.0117	-0.0129	0.0364	0.0247

Table 7.1: Out-of-sample (S)VAR forecast for the following five years using the complete data.

In Table 7.1, the predictive interval are derived from (28). The intervals are fairly wide and become uninformative as the prediction year increases. Therefore only a five-year forecast was made.

To validate the model, one should compare in-sample estimations against the sample realizations. In-sample forecasts at  $t+1$  were estimated with the (S)VAR model containing realizations up to  $t$ . The estimation window was expanded such that the sample increases by one observation in each estimation iteration <sup>21</sup>. In Figure 7.2, most of the in-sample forecasts (marked in red) starting 1964 look decent for employment, although in Figure 7.1 the VAR and SVAR models were unable to capture the big decrease in GDP in the mid seventies, early eighties and in 2009. This is not surprising as VAR and SVAR do not have intrinsic mechanisms to capture big jumps.

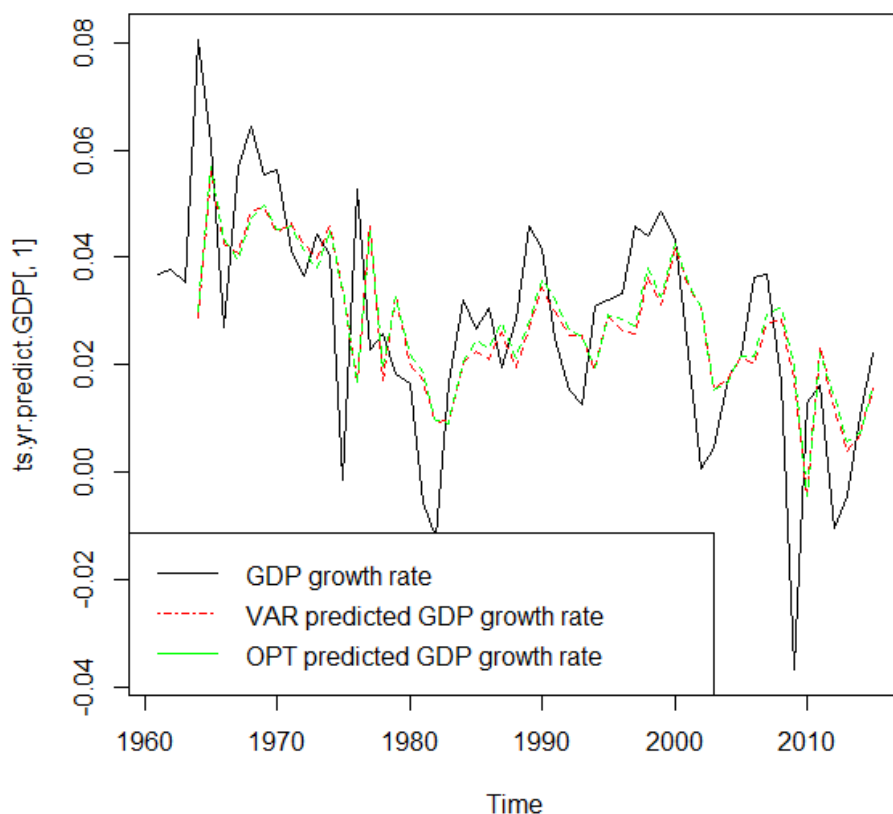


Figure 7.1: In-sample predictions for GDP.

## 7.2 Predictions: identification of shocks via a change in volatility

The VAR equations in (24) and (25) can similarly be used for predictions to be made for  $t+1$  in-sample forecasts <sup>22</sup>. The in-sample forecasts are simply fitted values of VAR. Predictions with these fitted values were plotted in green in Figure 7.1 and 7.2.

Compared to the (S)VAR fitted values, the fitted values derived by Lutkepohl's optimization method show no obvious improvement in in-sample forecasts. From the plots, it is very difficult

<sup>21</sup>Estimations in Table F.1, F.

<sup>22</sup>Estimations in Table F.2, F.

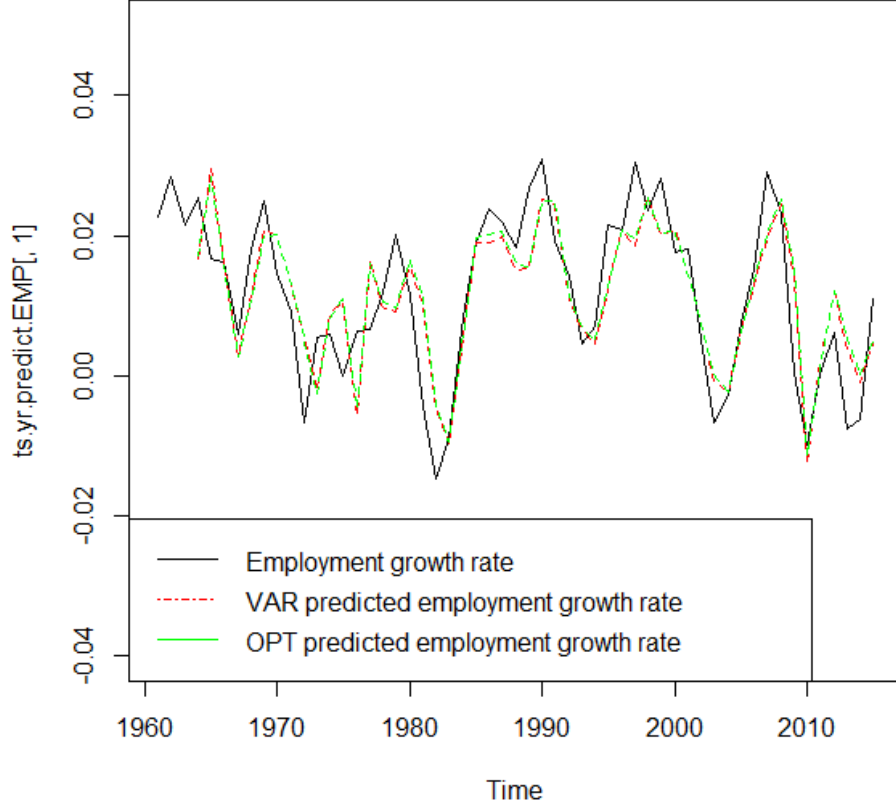


Figure 7.2: In-sample predictions for employment.

to judge which model performs better around the recession time in the early 2000. The in-sample predictions for the two approaches do not seem to have utter differences as their residual sum of squares are very close with  $RSS_{\text{var}} = 0.01835212$  and  $RSS_{\text{opt}} = 0.01846032$ .

### 7.3 Comparison of the predictions through simulation

One can find out which technique is better by comparing the residual sum of squares for their in-sample predictions. Simulate data  $y_{\text{sim}}$  using the technique in Section 6.4, and incorporate  $y_{\text{sim}}$  respectively with the VAR parameters results in  $\hat{y}_{\text{var}}$ . Using the same in-sample data  $y_{\text{sim}}$  and incorporating them with the optimized  $\beta$  parameters achieved by Lutkepohl's method results in fitted values  $\hat{y}_{\text{opt}}$ . Compare the VAR fitted values with the simulated data  $RSS_{\text{var}} = (\hat{y}_{\text{var}} - y_{\text{sim}})^2$ , and similarly  $RSS_{\text{opt}} = (\hat{y}_{\text{opt}} - y_{\text{sim}})^2$ . After summing each method's RSS for the two variables, we find  $RSS_{\text{opt}} < RSS_{\text{var}}$ . Although in Section 7.2 Lutkepohl's model fits the real data less adequately than the VAR model, the reason could very well be the limited sample size. As demonstrated in Section 6.4, increasing the sample size increases the accuracy of the parameter estimates, which improves the fitted values. In this simulation test, with a large sample of 300 data simulated for each variable, the residual sum of squares for Lutkepohl's estimation model is smaller than that for the VAR model.

For out-of-sample prediction comparisons, simulate in-sample data  $y_{\text{in.sim}}$  and out-of-sample data  $y_{\text{out.sim}}$  for  $h = 5$ . Incorporate VAR and the optimized parameters estimates with the simulated in-sample data  $y_{\text{in.sim}}$  to sequentially generate out-of-sample estimations  $\hat{y}_{\text{var}}$  and  $\hat{y}_{\text{opt}}$ , each with  $h = 5$ . Calculate the mean squared error  $\text{MSE}_{\text{var}} = \frac{1}{h} \sum (y_{\text{sim}} - \hat{y}_{\text{var}})^2$  and  $\text{MSE}_{\text{opt}} = \frac{1}{h} \sum (y_{\text{sim}} - \hat{y}_{\text{opt}})^2$ . Calculate the MSE for VAR and compare it with the MSE for Lutkepohl's model. We found that  $\text{MSE}_{\text{opt}} < \text{MSE}_{\text{var}}$ .

We want to compare Lutkepohl's model to the VAR model, as far as their predictive behavior is concerned. Suppose Lutkepohl's model is really describing the true conditions of the series, in that case VAR is an incorrect model. However we need to take into account that when we have limited data, we are not able to estimate them accurately enough using Lutkepohl's correct model. With these extra noise, even if Lutkepohl's model is the right model, it is worse than VAR in terms of predictions. Then one needs to make decisions based on empirical experiences. Lutkepohl's model is more general, because it has volatility matrices and gives more freedom to the analysis. If the covariance matrices are the same, Lutkepohl's model is just a VAR model with no special specifications. Therefore this approach is not less valid than (S)VAR.

Lutkepohl (2012) stated in his paper that this approach is a purely technical way of identifying unique shocks. It is comparable to Choleski decomposition, which only obtains orthogonal shocks but does not lead automatically to economically meaningful shocks. According to Lutkepohl the statistical identification is handy but the method just normalizes the shocks in a mathematical way. The shocks identified this way need not to correspond to the shocks in the variables. The interpretation of this method requires economic insights.



## 8 Discussion

In this thesis two goals were achieved. To recap the research questions, one is to find out the strength and lag of employment reaction to changes in GDP, the other is to produce a supplementary approach to explain the reaction size and lag in asymmetrical economic times. The comparisons between (S)VAR and Lutkepohl's model are done using in-sample and out-of-sample predictions.

The first goal was achieved by an SVAR model, based on which an IRF was made. The impact of economic production shocks in the Netherlands is reflected in a positive employment growth rate that decreases in size till lag two. In this context this means the impact lasts for two years upon the economic shock. The decrease differs for the complete and reduced data, and the difference is highlighted by their impulse responses. For an SVAR model built on data that includes the data from the recessionary period, positive unit shock in GDP is related to nearly unchanged employment growth rate one year following this shock. The decrease of the employment growth rate then returns to equilibrium after about two years. From such impulse responses, one can observe a consistency with the CPB findings: compared to other countries, the Netherlands seems to show a difference in the employment reaction to the GDP dynamics. In other countries, the decline in employment rate followed the shocks in GDP immediately, whereas in the Netherlands this reaction has a one-year delay. An SVAR built on the reduced data is "clean" as it may be more trustworthy in revealing the relation between the two variables free of disturbances from the Great Recession. For an SVAR model built on data that excludes the data from the recessionary period, the employment growth rate continuously decreases until it returns to equilibrium.

VAR and SVAR modeled the two variables. To a certain extent they described the dynamics in these two variables using the impulse response function. Their results confirmed the CPB's hypothesis about the difference in the Netherlands compared to other countries in terms of the relation between GDP and employment upon the Great Recession. However with regards to how this difference came about, or economic implications for this difference, VAR and SVAR do not have further explanations. Economic insights are needed to interpret the results found in light of this difference.

The CPB mainly focused on major socio-economic variables for such explanations. It is said that labour hoarding, temporary employment, self-employment, and real wages played an opposite role in the employment reaction to the crisis [CPB, 2014], which cushioned the effects economic shocks had on employment. Due to the small sample size and variables and the complexity of modelling, this research is limited in its economic implications. Including more national labour market variables may expand the economic purposes: components such as full-time/part time, male/female employment, and GDP components in trade, added value, savings and income of the two variables may be analyzed in detail. This all calls for further hypotheses and tests that may be linked to research at the CPB and other institutes in The Netherlands. Macroeconomics variables per se are not the only means through which explanations for this national difference can be derived.

In addition to macroeconomics, what can be looked at is NUT 2 sectoral and regional explanations. The Netherlands is a country with most of its jobs in Sector three - a tertiary industry. People employed in the tertiary industry are more flexible to economic turns/shocks, and thus more resilient to economic shocks [Kazutoshi, 1992]. This may indeed impede employment's reaction to GDP in times of bad economics because the people employed can relatively easily adjust to different working roles compared to those in the primary and secondary industries.

Regional economics variables may shed light at the divergence according to another CPB's publication (2015): The mismatch between labour supply and labour demand finds its roots in e.g., the relocation of business activities or regional differences in employment growth.

Recently it was announced that The Netherlands has the strongest after-crisis growth in Europe. Up to 2015 the GDP growth was 1.9% [Badir, 2016]. Perhaps as an open economy majorly dependent on imports and exports (and re-exports) implies that in bad times decline hits its employment harder, later, with more impact, while in good times an equally fast positive effect may be present for this reason. However the data run till 2015, which is too short to see resilience/recovery in the Netherlands already, as only the last two years actually show renewed growth again.

The second goal of the thesis to build a model based the existing ones was achieved by implementing Lutkepohl's (2012) approach in R. Lutkepohl has Markov switching in residual volatility – one of his other approaches utilizing volatility change – implemented in MATLAB, but the current approach is not known in any programming languages. This contribution may help the CPB and other economic research institutes improve the VAR model in times of asymmetrical economy. Prior to the adoption of this approach, one should observe the time series to see if the volatility and the subsequent residual variances change. More specifically, if the changes satisfy what was addressed in this paper.

As stated before, this approach only performs in a mechanical way. It may not predict better in separate economic regimes when the sample size is small, but increasing the sample size can improve the accuracy of this model. Therefore it can be considered as a good model with more freedom in the choice of the residual covariances. In this paper, Lutkepohl's model does not have improvement over traditional models in prediction behavior. Given that the limitations of the sample could disrupt the result of this model, it is strongly suggested to expand the data size for further usage.

Since the beginning of the Great Recession, few models were designed in particular to account for the recessional abnormality, either for the employment growth or other labour market variables. Because of the Great Recession, GDP growth became unprecedentedly low, which may call for new legitimate models for various reactions to this change. In Eurostat and ECB (2016) one simple model can be found to identify the changing cyclicity of euro area employment over the course of the crisis and the subsequent recovery period. Such a simple quantitative model requires the support from more in-depth qualitative input, which are far-fetched for this thesis.

To conclude, the thesis achieves its goals in part. Further research for modelling and economic interpretations should be made to render in-depth explanations for the different development of GDP and employment in the Netherlands.

## 9 Appendix

### A Data and descriptive

	Time	GDP Volume	Market Prices	Total Employment
1	1960		1.46652e+11	4347823
2	1961		1.51373e+11	4446754
3	1962		1.57649e+11	4572996
4	1963		1.63499e+11	4670939
5	1964		1.76550e+11	4789397
6	1965		1.86599e+11	4869676
7	1966		1.92234e+11	4948973
8	1967		2.02910e+11	4977865
9	1968		2.16242e+11	5067969
10	1969		2.27962e+11	5194255
11	1970		2.40948e+11	5269608
12	1971		2.51114e+11	5317323
13	1972		2.59429e+11	5281253
14	1973		2.71592e+11	5310131
15	1974		2.82362e+11	5340996
16	1975		2.82107e+11	5339997
17	1976		2.96547e+11	5373839
18	1977		3.03421e+11	5409664
19	1978		3.10958e+11	5474178
20	1979		3.16673e+11	5584872
21	1980		3.22000e+11	5648393
22	1981		3.20346e+11	5623376
23	1982		3.16237e+11	5540016
24	1983		3.21795e+11	5490237
25	1984		3.31833e+11	5536004
26	1985		3.40647e+11	5641567
27	1986		3.51294e+11	5776231
28	1987		3.57793e+11	5902820
29	1988		3.68457e+11	6010670
30	1989		3.84961e+11	6172216
31	1990		4.01047e+11	6362913
32	1991		4.10877e+11	6485418
33	1992		4.17410e+11	6577025
34	1993		4.22685e+11	6607104
35	1994		4.35400e+11	6654059
36	1995		4.49565e+11	6796896
37	1996		4.64355e+11	6938011
38	1997		4.85317e+11	7150151
39	1998		5.06758e+11	7318368
40	1999		5.31859e+11	7523966
41	2000		5.54909e+11	7655736
42	2001		5.66692e+11	7794048
43	2002		5.67279e+11	7835167

	Time	GDP Volume	Market Prices	Total Employment
44	2003		5.69384e+11	7782833
45	2004		5.79443e+11	7760500
46	2005		5.92475e+11	7818250
47	2006		6.13847e+11	7938583
48	2007		6.36539e+11	8169167
49	2008		6.47350e+11	8358250
50	2009		6.22968e+11	8361000
51	2010		6.31174e+11	8277917
52	2011		6.41681e+11	8279583
53	2012		6.34896e+11	8330000
54	2013		6.32278e+11	8266500
55	2014		6.38669e+11	8214000
56	2015		6.52757e+11	8303357
57	2016		6.68768e+11	8392485
56	2017		6.86833e+11	8505093

Table A.1: Raw data summary

	Time	GDP	Employment
1	1961	0.0368965	0.0227539
2	1962	0.0376802	0.0283898
3	1963	0.0352190	0.0214177
4	1964	0.0805316	0.0253606
5	1965	0.0609474	0.0167620
6	1966	0.0270398	0.0162837
7	1967	0.0567697	0.0058379
8	1968	0.0643122	0.0181009
9	1969	0.0553588	0.0249186
10	1970	0.0563992	0.0145069
11	1971	0.0412088	0.0090549
12	1972	0.0363747	-0.0067835
13	1973	0.0443552	0.0054681
14	1974	0.0402760	0.0058124
15	1975	-0.0015890	-0.0001871
16	1976	0.0525609	0.0063374
17	1977	0.0227682	0.0066667
18	1978	0.0257681	0.0119257
19	1979	0.0181460	0.0202210
20	1980	0.0163962	0.0113738
21	1981	-0.0055277	-0.0044289
22	1982	-0.0122474	-0.0148238
23	1983	0.0166230	-0.0089854
24	1984	0.0319823	0.0083360
25	1985	0.0266208	0.0190684
26	1986	0.0306322	0.0238700
27	1987	0.0195287	0.0219155
28	1988	0.0285342	0.0182709
29	1989	0.0457270	0.0268765

	Time	GDP	Employment
30	1990	0.0415551	0.0308961
31	1991	0.0247886	0.0192530
32	1992	0.0152937	0.0141250
33	1993	0.0124630	0.0045734
34	1994	0.0306948	0.0071067
35	1995	0.0321292	0.0214661
36	1996	0.0332070	0.0207617
37	1997	0.0457297	0.0305765
38	1998	0.0440040	0.0235263
39	1999	0.0486219	0.0280935
40	2000	0.0432725	0.0175134
41	2001	0.0221094	0.0180664
42	2002	0.0004743	0.0052758
43	2003	0.0043831	-0.0066794
44	2004	0.0176845	-0.0028696
45	2005	0.0215217	0.0074415
46	2006	0.0362603	0.0153913
47	2007	0.0368183	0.0290459
48	2008	0.0168897	0.0231460
49	2009	-0.0367614	0.0003290
50	2010	0.0129011	-0.0099370
51	2011	0.0160876	0.0002013
52	2012	-0.0103609	0.0060893
53	2013	-0.0043975	-0.0076230
54	2014	0.0104535	-0.0063509
55	2015	0.0221179	0.0108787

Table A.2: Data summary

## B VAR estimation method

For a sample of size  $T$ , define:

$Y = (y_1, \dots, y_T)$ ,  $B = (A_1, \dots, A_p)$ ,  $Z_t = [y_t, \dots, y_{t-p+1}]'$ ,  $Z = (Z_0, \dots, Z_{T-1})$ ,  $U = (u_1, \dots, u_T)$ ,  $y = \text{vec}(Y)$ ,  $\beta = \text{vec}(B)$ ,  $b = \text{vec}(B')$ ,  $u = \text{vec}(U)$ .

The VAR model in (10) can be written as

$$Y = BZ + U \quad (29)$$

or

$$y = (Z' \otimes I_K)\beta + u \quad (30)$$

The covariance matrix of  $u$  is

$$\Sigma_u = I_T \otimes \Sigma_u \quad (31)$$

The multivariate LS estimation (or GLS estimation) of  $\beta$  means to choose the estimator that minimizes  $u'(I_T \otimes \Sigma_u)^{-1}u$ .

The LS estimator is  $\hat{\beta} = ((ZZ')^{-1}Z \otimes I_K)y$ .

Thus the VAR parameter estimates can be produced,

$$\hat{B} = B + UZ'(ZZ')^{-1} \quad (32)$$

## C Estimation of (over-identifying) structural parameters

The validity of overidentifying restrictions can be tested with an LR test [see Lutkepohl (2006)]:

$$LR = T(\ln|\tilde{\Sigma}_u^*| - \ln|\tilde{\Sigma}_u|) \quad (33)$$

$T$  is the number of observations. The statistic has a  $\chi^2$  distribution with degrees of freedom equals the number of restrictions exceeding  $K(K-1)/2$  - that is, the number of overidentifications - is computed.

$\tilde{\Sigma}_u$  in equation (33) can be obtained through the transformations of  $Y = [y_1, \dots, y_T]$  and  $Z = [Y_0, \dots, Y_{T-1}]$ :

$$\tilde{\Sigma}_u = T^{-1}(Y - \hat{A}Z)(Y - \hat{A}Z)' \quad (34)$$

and  $\tilde{\Sigma}_u^*$  in equation (33) can be obtained by:

$$\tilde{\Sigma}_u^* = \tilde{A}^{-1}\tilde{B}\tilde{B}'\tilde{A}'^{-1} \quad (35)$$

$\tilde{A}$ ,  $\tilde{B}$  in equation (35) are the ML estimators of the structural parameters and  $A$  and  $B$ :

$$\ln L_c(A, B) = constant + \frac{T}{2}\ln|A|^2 - \frac{T}{2}\ln|B|^2 - \frac{T}{2}tr(A'B^{-1'}B^{-1}A\tilde{\Sigma}_u) \quad (36)$$

Maximization with respect to  $A$  and  $B$  has to be done by numerical optimization, or direct maximization. For this thesis we use direct maximization.

## D Non-linear estimation of Lutkepohl's parameters

- Step 1: Obtain residuals  $u_t$  by estimating equationwise OLS of VAR;
- Step 2: Residuals obtained according to the data split;
  - $\hat{u}_{1t}$ : VAR residuals from 1961 - 2006
  - $\hat{u}_{2t}$ : VAR residuals from 2007 - 2015
- Step 3: Set  $Z_{t-1}$  as a combination of 1 and regressors for two dimensions;

$$Z1_{t-1} = (1, y'_{t-1}, \dots, y'_{t-p}) \quad (37)$$

- Step 4: Calculate the covariance matrices  $\Sigma$  for different regimes;

$$\hat{\Sigma}_1 = \frac{1}{T_{B-1} - 1 + 1} \sum_{t \in \text{regime1}} \hat{u}1_t \hat{u}1'_t \quad (38)$$

$$\hat{\Sigma}_2 = \frac{1}{T - T_B + 1} \sum_{t \in \text{regime2}} \hat{u}2_t \hat{u}2'_t \quad (39)$$

- Step 5:  $\beta$ : autoregressive estimators for different regimes;

$$\hat{\beta} = \left( \sum_{t=1}^{T_B-1} Z_{t-1} Z'_{t-1} \otimes \hat{\Sigma}_1^{-1} + \sum_{t=T_B}^T Z_{t-1} Z'_{t-1} \otimes \hat{\Sigma}_2^{-1} \right)^{-1} \left( \sum_{t=1}^{T_B-1} (Z_{t-1} \otimes \hat{\Sigma}_1^{-1}) y1_t + \sum_{t=T_B}^T (Z_{t-1} \otimes \hat{\Sigma}_2^{-1}) y2_t \right) \quad (40)$$

- Step 6: Construct new  $u_t$ ;

$$u_t = y_t - (Z'_t \otimes I_K) \hat{\beta} \quad (41)$$

- Step 7: Derive starting values of  $\Lambda$  and  $B$ ;

Generate elements  $\lambda$  and  $b$  from the  $\Sigma$  matrices obtained in equations (38) and (39).

$$\begin{aligned} \sigma_{1,1}^2 &= b_{11}^2 + b_{12}^2 \\ \sigma_{12,1} &= b_{11}b_{21} + b_{12}b_{22} \\ \sigma_{2,1}^2 &= b_{21}^2 + b_{22}^2 \\ \sigma_{1,2}^2 &= \lambda_1 b_{11}^2 + \lambda_2 b_{12}^2 \\ \sigma_{12,2} &= \lambda_1 b_{11}b_{21} + \lambda_2 b_{12}b_{22} \\ \sigma_{2,2}^2 &= \lambda_1 b_{21}^2 + \lambda_2 b_{22}^2 \end{aligned}$$

- Step 8: Concentrated log-likelihood optimized with respect to  $\sigma$ ;

$$\log \ell(\beta, \sigma) = -\frac{KT}{2} \log 2\pi - \frac{1}{2} \left( \sum_{t=1}^{T_B-1} \log |\Sigma_1| + \sum_{t=T_B}^T \log |\Sigma_2| \right) - \frac{1}{2} \left( \sum_{t=1}^{T_B-1} u'_t \Sigma_1^{-1} u_t + \sum_{t=T_B}^T u'_t \Sigma_2^{-1} u_t \right) \quad (42)$$

Numerical optimization using starting values of  $b$  and  $\lambda$  from the previous step. Negative log-likelihood of equation (40) produces MLE for the elements of  $\Lambda$  and  $B$ , which constitute new  $\Sigma_1$  and  $\Sigma_2$ .

- Step 9: Replace  $\hat{\Sigma}_1$  and  $\hat{\Sigma}_2$  in equation (40) with the new  $\Sigma_1$  and  $\Sigma_2$  in the previous step;

- Step 10: Repeat step 5 through step 8, until the elements of the consecutive  $\Sigma$  and  $\beta$  converge.

## E Results of simulation test

Simulations are built on the optimization process in D:

- Step 11: Simulate  $u_t$  using the two covariance matrices from fitting the VAR model, combine the simulated residuals with fixed auto-regressive parameters  $A$ 's to generate data;
- Step 12: Repeat step 3 through step 10 for each simulation;
- Step 13: Collect the parameter  $A$ 's estimates, and compare them with the fixed true parameters.

Steps 1 to 10 generated the parameters for the empirical data. In steps 11 to 13 data were simulated to test the reliability and the validity of the optimization function. Simulations returned global maxima of the parameters, regardless of the initial residuals simulated to replace  $u_t$  in step 2. The parameter estimations under varying simulations and sample sizes are recorded in the tables below.

	true parameters	est.mean	est.sd
GDP.constant	0.0081	0.0156	0.0167
EMP.constant	0.0050	0.0126	0.0163
GDP.GDP.Lag1	0.5049	0.3106	0.4433
EMP.GDP.Lag1	0.2565	0.0734	0.4427
GDP.EMP.Lag1	0.0401	0.1514	0.4086
EMP.EMP.Lag1	0.5711	0.6617	0.4045
GDP.GDP.Lag2	0.0063	-0.0561	0.3934
EMP.GDP.Lag2	-0.1061	-0.1685	0.3818
GDP.EMP.Lag2	-0.4336	-0.3856	0.3838
EMP.EMP.Lag2	0.0627	0.0917	0.3762
GDP.GDP.Lag3	0.2374	0.1051	0.3054
EMP.GDP.Lag3	-0.1261	-0.2447	0.3060
GDP.EMP.Lag3	0.2322	0.2193	0.2827
EMP.EMP.Lag3	-0.1222	-0.1215	0.2774

Table E.1: Sample size 55, 250 simulations.

	true parameters	est.mean	est.sd
GDP.constant	0.0081	0.0089	0.0040
EMP.constant	0.0050	0.0058	0.0040
GDP.GDP.Lag1	0.5049	0.4824	0.1219
EMP.GDP.Lag1	0.2565	0.2346	0.1218
GDP.EMP.Lag1	0.0401	0.0605	0.1268
EMP.EMP.Lag1	0.5711	0.5891	0.1270
GDP.GDP.Lag2	0.0063	-0.0063	0.1076
EMP.GDP.Lag2	-0.1061	-0.1179	0.1075
GDP.EMP.Lag2	-0.4336	-0.4374	0.1308
EMP.EMP.Lag2	0.0627	0.0546	0.1302
GDP.GDP.Lag3	0.2374	0.2310	0.1076
EMP.GDP.Lag3	-0.1261	-0.1311	0.1080
GDP.EMP.Lag3	0.2322	0.2308	0.0997
EMP.EMP.Lag3	-0.1222	-0.1204	0.1003

Table E.2: Sample size 300, 250 simulations.



## F In-sample forecasts

	GDP realizations	GDP forecast	EMP realizations	EMP forecast
1961	0.0368	NA	0.0227	NA
1962	0.0376	NA	0.0283	NA
1963	0.0352	NA	0.0214	NA
1964	0.0805	0.0287	0.0253	0.0166
1965	0.0609	0.0563	0.0167	0.0295
1966	0.0270	0.0424	0.0162	0.0162
1967	0.0567	0.0406	0.0058	0.0025
1968	0.0643	0.0485	0.0181	0.0113
1969	0.0553	0.0494	0.0249	0.0208
1970	0.0563	0.0445	0.0145	0.0198
1971	0.0412	0.0462	0.0090	0.0131
1972	0.0363	0.0423	-0.0067	0.0056
1973	0.0443	0.0393	0.0054	-0.0022
1974	0.0402	0.0458	0.0058	0.0089
1975	-0.0015	0.0337	-0.0001	0.0105
1976	0.0525	0.0169	0.0063	-0.0056
1977	0.0227	0.0459	0.0066	0.0164
1978	0.0257	0.0171	0.0119	0.0097
1979	0.0181	0.0328	0.0202	0.0090
1980	0.0163	0.0201	0.0113	0.0155
1981	-0.0055	0.0171	-0.0044	0.0103
1982	-0.0122	0.0093	-0.0148	-0.0047
1983	0.0166	0.0098	-0.0089	-0.0097
1984	0.0319	0.0202	0.0083	0.0057
1985	0.0266	0.0223	0.0190	0.0190
1986	0.0306	0.0208	0.0238	0.0188
1987	0.0195	0.0260	0.0219	0.0198
1988	0.0285	0.0195	0.0182	0.0151
1989	0.0457	0.0267	0.0268	0.0153
1990	0.0415	0.0343	0.0308	0.0250
1991	0.0247	0.0300	0.0192	0.0243
1992	0.0152	0.0254	0.0141	0.0108
1993	0.0124	0.0253	0.0045	0.0065
1994	0.0306	0.0189	0.0071	0.0046
1995	0.0321	0.0289	0.0214	0.0122
1996	0.0332	0.0264	0.0207	0.0205
1997	0.0457	0.0256	0.0305	0.0186
1998	0.0440	0.0363	0.0235	0.0253
1999	0.0486	0.0310	0.0280	0.0200
2000	0.0432	0.0418	0.0175	0.0208
2001	0.0221	0.0347	0.0180	0.0143
2002	0.0004	0.0308	0.0052	0.0079
2003	0.0043	0.0152	-0.0066	-0.0006
2004	0.0176	0.0172	-0.0028	-0.0023
2005	0.0215	0.0212	0.0074	0.0063
2006	0.0362	0.0201	0.0153	0.0130
2007	0.0368	0.0275	0.0290	0.0194
2008	0.0168	0.0283	0.0231	0.0245
2009	-0.0367	0.0174	0.0003	0.0140
2010	0.0129	-0.0048	-0.0099	-0.0127
2011	0.0160	0.0233	0.0002	0.0016
2012	-0.0103	0.0120	0.0060	0.0118
2013	-0.0043	0.0039	-0.0076	0.0037
2014	0.0104	0.0068	-0.0063	-0.0010
2015	0.0221	0.0154	0.0108	0.0046

Table F.1: In-sample (S)VAR forecast for years 1964 - 2015.

	GDP realizations	GDP forecast	EMP realizations	EMP forecast
1961	0.0368	NA	0.0227	NA
1962	0.0376	NA	0.0283	NA
1963	0.0352	NA	0.0214	NA
1964	0.0805	0.0320	0.0253	0.0171
1965	0.0609	0.0559	0.0167	0.0295
1966	0.0270	0.0447	0.0162	0.0181
1967	0.0567	0.0411	0.0058	0.0034
1968	0.0643	0.0461	0.0181	0.0106
1969	0.0553	0.0481	0.0249	0.0205
1970	0.0563	0.0445	0.0145	0.0183
1971	0.0412	0.0463	0.0090	0.0133
1972	0.0363	0.0422	-0.0067	0.0069
1973	0.0443	0.0370	0.0054	-0.0017
1974	0.0402	0.0417	0.0058	0.0074
1975	-0.0015	0.0316	-0.0001	0.0082
1976	0.0525	0.0171	0.0063	-0.0060
1977	0.0227	0.0411	0.0066	0.0134
1978	0.0257	0.0185	0.0119	0.0093
1979	0.0181	0.0313	0.0202	0.0068
1980	0.0163	0.0217	0.0113	0.0138
1981	-0.0055	0.0193	-0.0044	0.0097
1982	-0.0122	0.0113	-0.0148	-0.0033
1983	0.0166	0.0084	-0.0089	-0.0099
1984	0.0319	0.0161	0.0083	0.0029
1985	0.0266	0.0199	0.0190	0.0151
1986	0.0306	0.0213	0.0238	0.0156
1987	0.0195	0.0277	0.0219	0.0180
1988	0.0285	0.0232	0.0182	0.0151
1989	0.0457	0.0290	0.0268	0.0154
1990	0.0415	0.0361	0.0308	0.0246
1991	0.0247	0.0336	0.0192	0.0241
1992	0.0152	0.0294	0.0141	0.0121
1993	0.0124	0.0277	0.0045	0.0078
1994	0.0306	0.0203	0.0071	0.0050
1995	0.0321	0.0281	0.0214	0.0114
1996	0.0332	0.0270	0.0207	0.0186
1997	0.0457	0.0275	0.0305	0.0171
1998	0.0440	0.0383	0.0235	0.0246
1999	0.0486	0.0345	0.0280	0.0206
2000	0.0432	0.0439	0.0175	0.0215
2001	0.0221	0.0372	0.0180	0.0153
2002	0.0004	0.0329	0.0052	0.0090
2003	0.0043	0.0173	-0.0066	-0.0003
2004	0.0176	0.0168	-0.0028	-0.0020
2005	0.0215	0.0191	0.0074	0.0049
2006	0.0362	0.0188	0.0153	0.0101
2007	0.0368	0.0268	0.0290	0.0165
2008	0.0168	0.0301	0.0231	0.0226
2009	-0.0367	0.0218	0.0003	0.0138
2010	0.0129	0.0006	-0.0099	-0.0102
2011	0.0160	0.0206	0.0002	0.0011
2012	-0.0103	0.0115	0.0060	0.0106
2013	-0.0043	0.0045	-0.0076	0.0008
2014	0.0104	0.0064	-0.0063	-0.0027
2015	0.0221	0.0137	0.0108	0.0034

Table F.2: In-sample Lutkepohl's model forecast for years 1964 - 2015.

## G Source of R code used in this thesis

### G.1 R code Chapter 2

```
#####  
# Load necessary packages  
#####  
  
library(dse)  
library(strucchange)  
library(knitr)  
library(tseries)  
library(vars)  
library(xlsx)  
library(urca)  
library(CADFtest)  
library(forecast)  
library(car)  
library(dynlm)  
  
#####  
# Sort data  
#####  
  
# Load raw data  
load("data_year.RData")  
colnames(data.year) <- c("Country",  
                        "Time",  
                        "GrossDomesticProductValueMarketPrices",  
                        "GrossDomesticProductVolumeMarketPrices",  
                        "GrossDomesticProductDeflatorMarketPrices",  
                        "UnemploymentLevel",  
                        "UnemploymentRate",  
                        "LabourProductivityOfTheTotalEconomy",  
                        "TotalEmployment",  
                        "TotalEmploymentNationalAccountsBasis",  
                        "LabourForce",  
                        "HoursWorkedPerEmployeeTotalEconomy",  
                        "JobVacanciesTotalNewVacanciesFlow",  
                        "JobVacanciesTotalUnfilledVacanciesStock",  
                        "FinancialNetWorthOfNonfinancialCorporations",  
                        "DebtOfNonfinancialCorporations",  
                        "ShortTermFinancialAssetsToShortTermLiabilitiesRatioOf  
NonfinancialCorporations",  
                        "DebtToEquityRatioOfNonfinancialCorporations",  
                        "DebtToTotalFinancialAssetsRatioOfNonfinancialCorporations",  
                        "FinancialNetWorthOfFinancialCorporations",  
                        "DebtOfFinancialCorporations",  
                        "ShortTermFinancialAssetsToShortTermLiabilitiesRatioOf
```

```

FinancialCorporations",
"DebtToEquityRatioOfFinancialCorporations",
"NonfinancialCorporationsDebtToGrossOperatingSurplus")

# Variables index.
varind <- which(colnames(data.year) == "GrossDomesticProductValueMarketPrices"):
  which(colnames(data.year) == "HoursWorkedPerEmployeeTotalEconomy")

# Country names.
country.names <- as.vector(unique(data.year$Country))

# Replace all the coma in the original data with points.
data.year[, varind] <- apply(data.year[, varind], 2, function(x)
  {as.numeric(gsub(",", ".", x, fixed = TRUE))})

#####
# Create new data
#####

# Create first difference GDP (fdGDP).
data.year$fdGDP <- c(NA, diff(data.year$GrossDomesticProductVolumeMarketPrices))

# Create growth rate GDP (GrowthGDP).
nGDP <- length(data.year$GrossDomesticProductVolumeMarketPrices)
data.year$GrowthGDP <- (data.year$fdGDP /
  c(NA, data.year$GrossDomesticProductVolumeMarketPrices[-nGDP]))

# Create GDP deflator 2012 (GDPDeflatorMarketPrices 2012).
base.year <- which(data.year$Time == 2012) # Find row index for year 2012.
baseyear2012 <- data.year[base.year, "GrossDomesticProductDeflatorMarketPrices"]
# Find deflator 2012 for all countries.
mat <- NULL
for(i in 1:34){ # Create deflator for different countries with base year 2012.
  deflator.country <- data.year[data.year$Country == country.names[i],
    "GrossDomesticProductDeflatorMarketPrices"]
  mat <- c(mat, deflator.country / baseyear2012[i])
}
data.year$GDPDeflatorMarketPrices2012 <- mat

# Create new GDP (NewGDPVolumeMarketPrices).
data.year$NewGDPVolumeMarketPrices <- data.year$GrossDomesticProductValueMarketPrices
/ data.year$GDPDeflatorMarketPrices2012

# Create first difference new GDP (fdNewGDP).
data.year$fdNewGDP <- c(NA, diff(data.year$NewGDPVolumeMarketPrices))

# Create new growth rate GDP (NewGrowthGDP).
nNewGDP <- length(data.year$NewGDPVolumeMarketPrices)
data.year$NewGrowthGDP <- (data.year$fdNewGDP /

```

```

c(NA, data.year$NewGDPVolumeMarketPrices[-nNewGDP]))

# Create difference between old and new growth rate GDP (diffGrowthGDP).
data.year$diffGrowthGDP <- data.year$NewGrowthGDP - data.year$GrowthGDP

# Create employment rate.
data.year$EmpRate <- c(data.year$TotalEmployment / data.year$LabourForce)

# Create first difference total employment (fdEmp).
data.year$fdEmp <- c(NA, diff(data.year$TotalEmployment))

# Create growth rate total employment (GrowthEmp).
nemp <- length(data.year$TotalEmployment)
data.year$GrowthEmp <- (data.year$fdEmp / c(NA, data.year$TotalEmployment[-nemp]))

# Create first difference total employment national accounts (fdempNA).
data.year$fdEmpNA <- c(NA, diff(data.year$TotalEmploymentNationalAccountsBasis))

# Create growth rate total employment national accounts (GrowthEmpNA).
nempNA <- length(data.year$TotalEmploymentNationalAccountsBasis)
data.year$GrowthEmpNA <- (data.year$fdEmpNA /
                          c(NA, data.year$TotalEmploymentNationalAccountsBasis[-nempNA]))

# Create first difference unemployment level (fdUnempLevel).
data.year$fdUnempLevel <- c(NA, diff(data.year$UnemploymentLevel))

# Create growth rate unemployment level (GrowthUnempLevel).
nunempL <- length(data.year$UnemploymentLevel)
data.year$GrowthUnempLevel <- (data.year$fdUnempLevel /
                              c(NA, data.year$UnemploymentLevel[-nunempL]))

# Create first difference unemployment rate (fdUnempRate).
data.year$fdUnempRate <- c(NA, diff(data.year$UnemploymentRate))

# Create first difference labour force (fdLabourForce).
data.year$fdLabourForce <- c(NA, diff(data.year$LabourForce))

# Create growth rate labour force (GrowthLabourForce).
nlabourforce <- length(data.year$LabourForce)
data.year$GrowthLabourForce <- (data.year$fdLabourForce /
                              c(NA, data.year$LabourForce[-nlabourforce]))

#####
# Output created data
#####

# Remove the first growth rates for each country.
country.names <- as.vector(unique(data.year$Country))
ind <- c(which(colnames(data.year) == "fdGDP"), #

```

```

which(colnames(data.year) == "GrowthGDP"), #
which(colnames(data.year) == "fdNewGDP"), #
which(colnames(data.year) == "NewGrowthGDP"), #
which(colnames(data.year) == "diffGrowthGDP"), #
which(colnames(data.year) == "fdEmp"), #
which(colnames(data.year) == "GrowthEmp"), #
which(colnames(data.year) == "fdEmpNA"), #
which(colnames(data.year) == "GrowthEmpNA"), #
which(colnames(data.year) == "EmpNASum"), #
which(colnames(data.year) == "fdUnempLevel"), #
which(colnames(data.year) == "GrowthUnempLevel"), #
which(colnames(data.year) == "fdUnempRate"), #
which(colnames(data.year) == "fdLabourForce"), #
which(colnames(data.year) == "GrowthLabourForce"))

for(i in country.names){
  data.year[data.year$Country == i, ][1, c(ind)] <- NA
}

# Reorganize columns.
data.year <- data.year[, c("Country", "Time", "GrossDomesticProductValueMarketPrices",
  "GrossDomesticProductVolumeMarketPrices",
  "GrossDomesticProductDeflatorMarketPrices",
  "GDPDeflatorMarketPrices2012", "NewGDPVolumeMarketPrices",
  "fdGDP", "GrowthGDP", "fdNewGDP", "NewGrowthGDP",
  "diffGrowthGDP", "EmpRate", "TotalEmployment", "fdEmp",
  "GrowthEmp", "TotalEmploymentNationalAccountsBasis",
  "fdEmpNA", "GrowthEmpNA", "HoursWorkedPerEmployeeTotalEconomy",
  "UnemploymentLevel", "fdUnempLevel", "GrowthUnempLevel",
  "UnemploymentRate", "fdUnempRate",
  "LabourProductivityOfTheTotalEconomy", "LabourForce",
  "fdLabourForce", "GrowthLabourForce",
  "JobVacanciesTotalNewVacanciesFlow",
  "JobVacanciesTotalUnfilledVacanciesStock",
  "FinancialNetWorthOfNonfinancialCorporations",
  "DebtOfNonfinancialCorporations",
  "ShortTermFinancialAssetsToShortTermLiabilitiesRatioOf
NonfinancialCorporations",
  "DebtToEquityRatioOfNonfinancialCorporations",
  "DebtToTotalFinancialAssetsRatioOfNonfinancialCorporations",
  "FinancialNetWorthOfFinancialCorporations",
  "DebtOfFinancialCorporations",
  "ShortTermFinancialAssetsToShortTermLiabilitiesRatioOf
FinancialCorporations",
  "DebtToEquityRatioOfFinancialCorporations",
  "NonfinancialCorporationsDebtToGrossOperatingSurplus")]

# Data for the Netherlands.

```

```

data.year.NL <- data.year[data.year$Country == "Netherlands", ]

# Export data.
write.xlsx(data.year, file = "data_year_aggregate.xlsx", sheetName="R export",
           col.names = TRUE, row.names = FALSE)
write.xlsx(data.year.NL, file = "data_year_aggregate_NL.xlsx", sheetName="R export",
           col.names = TRUE, row.names = FALSE)
save.image("M:/p_arb/BBP_wgh/Xinzhu/data.year.RData")
save.image("M:/p_arb/BBP_wgh/Xinzhu/data.year.NL.RData")

#####
# Data description
#####

load("M:/p_arb/BBP_wgh/Xinzhu/data.year.RData")
load("M:/p_arb/BBP_wgh/Xinzhu/data.year.NL.RData")

kable(data.year.NL[, c(2,4,14)])
kable(summary(data.year.NL[-c(56:57), c(4, 14)]))

org.ts.gdp <- ts(data.year.NL[-c(56:57), 4], start = 1961, end = 2015, freq = 1)
plot(org.ts.gdp, xlab = "Time (years)", ylab = "", main = "Nominal GDP", col = "red")

org.ts.emp <- ts(data.year.NL[-c(56:57), 14], start = 1961, end = 2015, freq = 1)
plot(org.ts.emp, xlab = "Time (years)", ylab = "", main = "Total employment", col = "green")

```

## G.2 R code Chapter 3

```

#####
# Load data
#####

NL <- na.omit(data.frame(data.year.NL$Time,
                        data.year.NL$NewGrowthGDP,
                        data.year.NL$GrowthEmp))
colnames(NL) <- c("Time", "GDP", "Employment")

kable(NL[-c(56:57), ])
kable(summary(NL[-c(56:57), 2:3]))

ts.yr <- ts(NL[-c(56:57), 2:3], start = head(NL$Time, n = 1),
           end = tail(NL$Time, n = 1), freq = 1)
plot(ts.yr, xlab = "Time (years)", ylab = c("GDP", "Employment"), main = "Netherlands")

#####
# Automatic lags
#####

```

```

summary(ur.df(NL[-c(56:57), "GDP"], type = "drift"))
summary(ur.ers(NL[-c(56:57), "GDP"], type = "P-test", model = "constant"))
summary(Arima(NL[-c(56:57), "GDP"], order = c(1, 0, 0), include.mean = TRUE))

summary(ur.df(NL[-c(56:57), "Employment"], type = "drift"))
summary(ur.ers(NL[-c(56:57), "Employment"], type = "P-test", model = "constant"))

#####
# OLS estimation
#####

op.lag.drift <- function(lags, X){

  z <- diff(X)
  n <- length(z)
  z.diff <- embed(z, lags + 1)[,1]
  z.lag.1 <- X[(lags + 1):n]
  k <- lags + 1
  z.diff.lag <- embed(z, lags + 1)[, 2:k]
  z.diff.lag <- as.matrix(z.diff.lag)
  tt <- (lags+1):n

  # t-test for the lowest lag.
  mod1 <- lm(z.diff ~ 1 + z.lag.1 + z.diff.lag)
  p <- summary(mod1)$coefficients[2 + lags, 4]

  if(p > 0.05){
    print(paste("p value for", rownames(summary(mod1)$coefficients)[2+lags], "is", p))

    # F-test for the full model and models with removed low lags.
    out <- NULL
    for(i in 1:(lags - 2)){
      z.diff.lag.dp <- z.diff.lag[, -c((lags - i): lags)]
      mod0 <- lm(z.diff ~ 1 + z.lag.1 + z.diff.lag.dp)
      mod.nolag <- lm(z.diff ~ 1 + z.lag.1)
      out <- rbind(out, anova(mod0, mod1))
    }
    return(rbind(out, anova(mod.nolag, mod1)))

  }else{
    return(summary(mod1))
  }

}

op.lag.drift(lags = 5, X = NL[-c(56:57), "GDP"])
op.lag.drift(lags = 5, X = NL[-c(56:57), "Employment"])

op.lag.drift(lags = 5, X = NL[-c(47:57), "GDP"])

```



```

op.lag.drift(lags = 5, X = NL[-c(47:57), "Employment"])

summary(ur.df(NL[-c(56:57), "GDP"], type = "drift", lags=0))
summary(ur.ers(NL[-c(56:57), "GDP"], type = "P-test", model = "constant"))

summary(ur.df(NL[-c(56:57), "Employment"], type = "drift", lags=0))
summary(ur.ers(NL[-c(56:57), "Employment"], type = "P-test", model = "constant"))

#####
# Pacf and lags
#####

par(mfrow = c(2,2))
acf(NL[-c(56:57), "GDP"], main = "GDP")
pacf(NL[-c(56:57), "GDP"], main = "GDP")
acf(NL[-c(56:57), "Employment"], main = "Employment")
pacf(NL[-c(56:57), "Employment"], main = "Employment")

summary(ur.df(NL[-c(56:57), "GDP"], type = "drift", lags=0))
summary(ur.ers(NL[-c(56:57), "GDP"], type = "P-test", model = "constant"))

summary(ur.df(NL[-c(56:57), "Employment"], type = "drift", lags=1))
summary(ur.ers(NL[-c(56:57), "Employment"], type = "P-test", model = "constant"))

# Unit root tests for the reduced data.

summary(CADFtest(NL[-c(47:57), "GDP"], type = "drift", max.lag.y = 5,
                 criterion = c("AIC")))
summary(ur.df(NL[-c(47:57), "GDP"], type = "drift", lags = 0))
summary(ur.ers(NL[-c(47:57), "GDP"], type = "P-test", model = "constant"))

summary(CADFtest(NL[-c(47:57), "Employment"], type = "drift", max.lag.y = 5,
                 criterion = c("AIC")))
summary(ur.df(NL[-c(47:57), "Employment"], type = "drift", lags = 1))
summary(ur.ers(NL[-c(47:57), "Employment"], type = "P-test", model = "constant"))

#####
# Residuals diagnostics: complete versus reduced data
#####

# GDP residual diagnostics

plot(CADFtest(NL[-c(56:57), "GDP"], type = "drift", max.lag.y = 5, criterion = c("AIC")))

res <- as.numeric(resid(CADFtest(NL[-c(56:57), "GDP"], type = "drift",
                                max.lag.y = 5, criterion = c("AIC"))))
qqnorm(res, ylab = "res", main = "GDP QQ-plot")
qqline(res)
n <- length(res)

```

```

plot(res[-n], res[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("Serial correlation, r =", round(cor(res[-n], res[-1]), digits = 2)))

lags = 5 + 1
selectlags = c("BIC")
y <- NL[-c(56:57), "GDP"]
z <- diff(y)
n <- length(z)
x <- embed(z, lags)
z.diff <- x[, 1]
z.lag.1 <- y[lags:n]

critRes <- rep(NA, lags)
for (i in 2:(lags)) {
  z.diff.lag = x[, 2:i]
  result <- lm(z.diff ~ z.lag.1 + 1)
  critRes[i] <- AIC(result, k = switch(selectlags, AIC = 2, BIC = log(length(z.diff))))
  lags <- which.min(critRes)
}

z.diff.lag = x[, 2:lags]
result <- lm(z.diff ~ z.lag.1 + 1)
plot(as.numeric(fitted(result)), res, xlab = "fitted",
     main = "Homoscedasticity and linearity")
abline(h = 0)

# Employment residual diagnostics

plot(CADFTtest(NL[-c(56:57), "Employment"], type = "drift", max.lag.y = 5,
              criterion = c("AIC"))))

res <- as.numeric(resid(CADFTtest(NL[-c(56:57), "Employment"], type = "drift",
                                max.lag.y = 5, criterion = c("AIC"))))
qqnorm(res, ylab = "res", main = "Employment QQ-plot")
qqline(res)
n <- length(res)
plot(res[-n], res[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("Serial correlation, r =", round(cor(res[-n], res[-1]), digits = 2)))

lags = 5 + 1
selectlags = c("BIC")
y <- NL[-c(56:57), "Employment"]
z <- diff(y)
n <- length(z)
x <- embed(z, lags)
z.diff <- x[, 1]
z.lag.1 <- y[lags:n]

critRes <- rep(NA, lags)

```

```

for (i in 2:(lags)) {
  z.diff.lag = x[, 2:i]
  result <- lm(z.diff ~ z.lag.1 + 1 + z.diff.lag)
  critRes[i] <- AIC(result, k = switch(selectlags, AIC = 2,
                                     BIC = log(length(z.diff))))
  lags <- which.min(critRes)
}

z.diff.lag = x[, 2:lags]
result <- lm(z.diff ~ z.lag.1 + 1 + z.diff.lag)
plot(as.numeric(fitted(result)), res, xlab = "fitted", main = "Homoscedasticity
      and linearity")
abline(h = 0)

```

### G.3 R code Chapter 4

```

#####
# Vector lags
#####

VARselect(NL15, lag.max = 5, type = "const")
VARselect(NL06, lag.max = 5, type = "const")

varest15 <- VAR(NL15, p = 3, type = "const")
sum.varest15 <- summary(varest15)

varest06 <- VAR(NL06, p = 3, type = "const")
sum.varest06 <- summary(varest06)

#####
# VAR for the complete versus reduced data
#####

varest15 <- VAR(NL15, p = 3, type = "const")
sum.varest15 <- summary(varest15)

varest06 <- VAR(NL06, p = 3, type = "const")
sum.varest06 <- summary(varest06)

#####
# VAR covariance-stationarity
#####

roots(varest15)

roots(varest06)

ts.yr <- ts(NL15, start = head(NL$Time, n = 1), end = tail(NL$Time, n = 1), freq = 1)

```

```

plot(ts.yr, plot.type = "s", col = 2:3)
legend("topright", c("GDP growth rate", "Employment growth rate"), lty = 1,
      col = c("red", "green"))

ts.yr <- ts(NL06, start = head(NL$Time, n = 1), end = tail(NL$Time, n = 1), freq = 1)
plot(ts.yr, plot.type = "s", col = 2:3)
legend("topright", c("GDP growth rate", "Employment growth rate"), lty = 1,
      col = c("red", "green"))

#####
# VAR residuals diagnostics
#####

var2c.serial <- serial.test(varest15, type = "PT.asymptotic")
var2c.arch <- arch.test(varest15, multivariate.only = TRUE)
var2c.norm <- normality.test(varest15, multivariate.only = TRUE)

test <- c(var2c.serial$serial$method, var2c.arch$arch.mul$method,
          var2c.norm$jb.mul$JB$method, var2c.norm$jb.mul$Skewness$method,
          var2c.norm$jb.mul$Kurtosis$method)
statistic <- c(var2c.serial$serial$statistic, var2c.arch$arch.mul$statistic,
               var2c.norm$jb.mul$JB$statistic, var2c.norm$jb.mul$Skewness$statistic,
               var2c.norm$jb.mul$Kurtosis$statistic)
pvalue <- c(var2c.serial$serial$p.value, var2c.arch$arch.mul$p.value,
            var2c.norm$jb.mul$JB$p.value, var2c.norm$jb.mul$Skewness$p.value,
            var2c.norm$jb.mul$Kurtosis$p.value)
df <- data.frame(test, statistic, pvalue)
kable(df)

plot(varest15)

res.GDP <- as.numeric(resid(varest15)[, 1])
qqnorm(res.GDP, ylab = "res", main = "GDP QQ-plot")
qqline(res.GDP)
n.GDP <- length(res.GDP)
plot(res.GDP[-n.GDP], res.GDP[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("GDP serial correlation, r =", round(cor(res.GDP[-n.GDP],
                                                         res.GDP[-1]), digits = 2)))
plot(as.numeric(fitted(varest15)[, 1]), res.GDP, xlab = "fitted",
     main = "GDP homoscedasticity")
abline(h = 0)

res.EMP <- as.numeric(resid(varest15)[, 2])
qqnorm(res.EMP, ylab = "res", main = "Employment QQ-plot")
qqline(res.EMP)
n.EMP <- length(res.EMP)
plot(res.EMP[-n.EMP], res.EMP[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("EMP serial correlation, r =", round(cor(res.EMP[-n.EMP],
                                                         res.EMP[-1]), digits = 2)))

```

```

plot(as.numeric(fitted(varest15)[, 2]), res.EMP, xlab = "fitted",
     main = "Employment homoscedasticity")
abline(h = 0)

plot(varest06)

res.GDP <- as.numeric(resid(varest06)[, 1])
qqnorm(res.GDP, ylab = "res", main = "GDP QQ-plot")
qqline(res.GDP)
n.GDP <- length(res.GDP)
plot(res.GDP[-n.GDP], res.GDP[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("GDP serial correlation, r =", round(cor(res.GDP[-n.GDP],
                                                         res.GDP[-1]), digits = 2)))
plot(as.numeric(fitted(varest06)[, 1]), res.GDP, xlab = "fitted",
     main = "GDP homoscedasticity")
abline(h = 0)

res.EMP <- as.numeric(resid(varest06)[, 2])
qqnorm(res.EMP, ylab = "res", main = "Employment QQ-plot")
qqline(res.EMP)
n.EMP <- length(res.EMP)
plot(res.EMP[-n.EMP], res.EMP[-1], xlab = "res_t", ylab = "res_{t + 1}",
     main = paste("EMP serial correlation, r =", round(cor(res.EMP[-n.EMP],
                                                         res.EMP[-1]), digits = 2)))
plot(as.numeric(fitted(varest06)[, 2]), res.EMP, xlab = "fitted",
     main = "Employment homoscedasticity")
abline(h = 0)

#####
# Orthogonalized impulse response function
#####

IRF15 <- irf(varest15, n.ahead = 10, ortho = TRUE, cumulative = FALSE,
             boot = FALSE, seed = 12345)
plot(IRF15)

IRF06 <- irf(varest06, n.ahead = 10, ortho = TRUE, cumulative = FALSE,
             boot = FALSE, seed = 12345)

plot(IRF15)
plot(IRF06)

```

## G.4 R code Chapter 5

```

#####
# Overidentification tests and parameter estimation
#####

```

```

# Tests for the complete data
Amat <- diag(2)
Amat[2,1] <- NA
svarest15 <- SVAR(varest15, estmethod = c("direct"), Amat = Amat, max.iter = 1000)
summary(svarest15)

Amat <- diag(2)
Amat[1,2] <- Amat[2,1] <- NA
svarest15 <- SVAR(varest15, estmethod = c("direct"), Amat = Amat, max.iter = 1000)
summary(svarest15)

Amat <- diag(2)
Amat[1,1] <- Amat[1,2] <- Amat[2,2] <- NA
svarest15 <- SVAR(varest15, estmethod = c("direct"), Amat = Amat, max.iter = 1000,
                  lrtest = FALSE)
summary(svarest15)

Amat <- diag(2)
Amat[1,1] <- Amat[2,1] <- Amat[2,2] <- NA
svarest15 <- SVAR(varest15, estmethod = c("direct"), Amat = Amat, max.iter = 1000,
                  lrtest = FALSE)
summary(svarest15)

# Tests for the reduced data
Amat <- diag(2)
Amat[2,1] <- NA
svarest06 <- SVAR(varest06, estmethod = c("direct"), Amat = Amat, max.iter = 1000)
summary(svarest06)

Amat <- diag(2)
Amat[1,2] <- Amat[2,1] <- NA
svarest06 <- SVAR(varest06, estmethod = c("direct"), Amat = Amat, max.iter = 1000)
summary(svarest06)

Amat <- diag(2)
Amat[1,1] <- Amat[1,2] <- Amat[2,2] <- NA
svarest06 <- SVAR(varest06, estmethod = c("direct"), Amat = Amat, max.iter = 1000,
                  lrtest = FALSE)
summary(svarest06)

Amat <- diag(2)
Amat[1,1] <- Amat[2,1] <- Amat[2,2] <- NA
svarest06 <- SVAR(varest06, estmethod = c("direct"), Amat = Amat, max.iter = 1000,
                  lrtest = FALSE)
summary(svarest06)

#####
# Structural impulse response analysis

```

```
#####

# Plot SVAR IRF for complete data.
SIRF15 <- irf(svarest15, n.ahead = 10, ortho = TRUE, cumulative = FALSE, boot = FALSE,
              seed = 12345)
plot(SIRF15)

# Plot SVAR IRF for reduced data.
SIRF06 <- irf(svarest06, n.ahead = 10, ortho = TRUE, cumulative = FALSE, boot = FALSE,
              seed = 12345)
plot(SIRF06)
```

## G.5 R code Chapter 6

```
#####
# Implementation of the approach
#####

opt <- function(data, var.index, p, K, TB, random.start, h){
  # Gaussian log-likelihood function.
  #
  # Args:
  #   data: data
  #   var.index: variables index
  #   p: lags
  #   K: number of dimensions
  #   TB: time break point
  #
  # Returns:
  #   volchange: covariance matrices for different regimes and parameters
  #             estimates for the whole sample
  require(knitr)
  require(nleqslv)
  require(vars)

  if(random.start == FALSE){
    # Residuals obtained by estimating equationwise OLS of VAR.
    varest <- VAR(data[, var.index], p = p, type = "const")
    u <- rbind(matrix(rep(NA, p*K), ncol=K), matrix(residuals(varest), ncol = 2
                                                    )) # Sample VAR residuals.

    TB.index <- which(data$Year == TB)
    u1 <- u[(p+1):(TB.index-1),] # First regime VAR residuals.
    u2 <- u[(TB.index):nrow(data),] # Second regime VAR residuals.
  }

  if(random.start == TRUE){
    # Simulated residuals and data.
    u1 <- cbind(rnorm(200-p, mean=0.1, sd=sqrt(0.2)), rnorm(200-p, mean=0.4,
```

```

sd=sqrt(0.45)))
u2 <- cbind(rnorm(100+h, mean=-0.7, sd=sqrt(0.4)), rnorm(100+h, mean=-2,
sd=sqrt(0.5)))
u <- rbind(matrix(rep(NA, p*K), ncol=K), u1, u2)
A <- c(0.504914292, 0.040148983, 0.006361527, -0.433609669, 0.237462457,
0.232210695, 0.256568557, 0.571149051, -0.106114636, 0.062737666,
-0.126112307, -0.122209395)
const <- c(0.008180638 , 0.005025498)
y <- matrix(NA, nrow = 300+h, ncol = K)
y[1:p,1] <- rnorm(p, mean=0.03659856, sd=0.001257373)
y[1:p,2] <- rnorm(p, mean=0.02418713, sd=0.003700394)
# Data obtained by residuals and specific autoregressive parameter estimates
for(i in ((p+1):(300+h))){
  y[i,1] <- const[1] + A[1] * y[i-1, 1] + A[2] * y[i-1, 2] + A[3] * y[i-2, 1]
+ A[4] * y[i-2, 2] + A[5] * y[i-3, 1] + A[6] * y[i-3, 2] + u[i,1]
  y[i,2] <- const[2] + A[7] * y[i-1, 1] + A[8] * y[i-1, 2] + A[9] * y[i-2, 1]
+ A[10] * y[i-2, 2] + A[11] * y[i-3, 1] + A[12] * y[i-3, 2] + u[i,2]
}

data <- data.frame(Year = 1:(300+h), "GDP growth rate" = y[,1], "Employment
growth rate" = y[,2])
TB.index <- which(data$Year == TB)
}

# Covariance matrices for different regimes.
sigma1 <- 1 / (TB.index - 1) * matrix(rowSums(apply(u1, 1, function(x)
x %*% t(x))), nrow = 2, byrow = TRUE)
sigma2 <- 1 / (nrow(data) - TB.index + 1) *
matrix(rowSums(apply(u2, 1, function(x) x %*% t(x))), nrow = 2, byrow = TRUE)

# Beta
beta.gen <- function(cov1, cov2){
  # Beta generation function.
  #
  # Args:
  #   cov1: covariance matrix for the first regime.
  #   cov2: covariance matrix for the second regime.
  #
  # Returns:
  #   beta: vectorization of estimated beta containing the constants and
  #         autoregressive parameters.
  betal <- array(NA, dim = c((1+2*p)*2, (1+2*p)*2, nrow(data)))
  betar <- array(NA, dim = c((1+2*p)*2, 1, nrow(data)))

  for(t in (p+1):(nrow(data))){
    Z <- as.vector(unlist(c(1, data[t-1, var.index], data[t-2, var.index],

```



```

        data[t-3, var.index]))))

    if(t<TB.index){
      betal[,t] <- Z %*% t(Z) %x% solve(cov1)
      betar[,t] <- Z %x% solve(cov1) %*% matrix(unlist(data[t, var.index]),
                                                nrow = 2)
    }

    if(t>=TB.index){
      betal[,t] <- Z %*% t(Z) %x% solve(cov2)
      betar[,t] <- Z %x% solve(cov2) %*% matrix(unlist(data[t, var.index]),
                                                nrow = 2)
    }
  }

  betaL <- apply(betal[, (p+1):(nrow(data))], drop=FALSE, c(1,2), sum)
  betaR <- apply(betar[, (p+1):(nrow(data))], drop=FALSE, c(1,2), sum)
  beta <- solve(betaL, tol = 1e-50) %*% betaR
  return(beta)
}

beta <- beta.gen(sigma1, sigma2)

# Setting initial values for beta and sigma matrices.
beta.null <- matrix(rep(1, dim(beta)[1] * dim(beta)[2]), ncol = ncol(beta))
sigma1.null <- sigma2.null <- matrix(rep(1, dim(sigma1)[1] * dim(sigma1)[2]),
                                     ncol = ncol(sigma1))

# Iterative optimization.
while(any(abs(sigma1 - sigma1.null) >= 1e-5) ||
      any(abs(sigma2 - sigma2.null) >= 1e-5) ||
      any(abs(beta - beta.null) >= 1e-3)){

  beta <- beta.gen(sigma1, sigma2)

  # ut generated by yt - (Z' %x% I_K) %*% beta.
  ut <- array(NA, dim = c(1, K, nrow(data)))
  for(i in (p+1):nrow(data)){
    Zt <- as.vector(unlist(c(1, data[i, var.index], data[i-1, var.index],
                           data[i-2, var.index])))
    ut[,i] <- t(matrix(unlist(data[i, var.index]), nrow = 2) -
                t(Zt) %x% diag(K) %*% beta)
  }
  ut <- apply(ut, 2, I)

  # Numerical optimization.
  f <- function(pars){
    # Mirrored loglikelihood function for the model.

```

```

#
# Args:
# pars: initial paramters.
b11 <- pars[1] # b[1,1]
b12 <- pars[2] # b[1,2]
b21 <- pars[3] # b[2,1]
b22 <- pars[4] # b[2,2]
l11 <- pars[5] # lambda[1,1]
l22 <- pars[6] # lambda[2,2]
B <- matrix(c(b11, b12, b21, b22), nrow = 2, byrow = TRUE)
L <- matrix(c(l11, 0, 0, l22), nrow = 2, byrow = TRUE)
sigma1 <- B %%% t(B) # New covariance matrix for the first period.
sigma2 <- B %%% L %%% t(B) # New covariance matrix for the second period

ll <- rep(NA, nrow(data))

for(t in (p+1):nrow(data)){

  if(t < TB.index){
    ll[t] <- log(det(sigma1)) + as.vector(matrix(ut[t, ], nrow=1) %%%
                                              solve(sigma1) %%%
                                              matrix(ut[t, ], ncol=1))
  }

  if(t >= TB.index){
    ll[t] <- log(det(sigma2)) + as.vector(matrix(ut[t, ], nrow=1) %%%
                                              solve(sigma2) %%%
                                              matrix(ut[t, ], ncol=1))
  }
}

loglik <- -1/2 * (K * nrow(data)) * log(2 * pi) - 1/2 *
  sum(ll[(p+1):nrow(data)], p*det(sigma1))

# Negative log-likelihood.
-loglik
}

ini <- function(x){
  # Numerical optimization of b and lambda as initial values.
  #
  # Args:
  # x: unkonwn b and lambda elements.
  c(x[1] ^ 2 + x[2] ^ 2 - sigma1[1,1],
    x[1] * x[3] + x[2] * x[4] - sigma1[1,2],
    x[3] ^ 2 + x[4] ^ 2 - sigma1[2,2],
    x[5] * x[1] ^ 2 + x[6] * x[2] ^ 2 - sigma2[1,1],
    x[5] * x[1] * x[3] + x[6] * x[2] * x[4] - sigma2[1,2],
    x[5] * x[3] ^ 2 + x[6] * x[4] ^ 2 - sigma2[2,2])
}

```

```

}
x0 <- c(0.01,0.02,0.02,0.01,1.5,1.7)
ini.sol <- nleqslv(x0, ini, method = "Broyden")

# Starting b values that form the original covariance matrices,
# computed by ini function.
ini.B <- matrix(ini.sol$x[1:4], nrow = 2, byrow = TRUE)
# Starting lambda values that form the original covariance matrices,
# computed by ini function.
ini.L <- matrix(c(ini.sol$x[5], 0, 0, ini.sol$x[6]), nrow = 2, byrow = TRUE)

# Optimizing lambda and b elements and their standard errors using their
# generated starting values.
opm_mle <- optim(f, p = c(ini.B[1,1], ini.B[1,2], ini.B[2,1], ini.B[2,2],
                        ini.L[1,1], ini.L[2,2]),
               control = list(maxit = 5000), hessian = TRUE)
opm_se <- diag(sqrt(solve(opm_mle$hessian, tol = 1e-50)))

# Optimized B and lambda matrices.
B <- matrix(opm_mle$par[1:4], nrow = 2, byrow = TRUE)
L <- matrix(c(opm_mle$par[5], 0, 0, opm_mle$par[6]), nrow = 2, byrow = TRUE)

# New sigma.null matrices.
sigma1.null <- sigma1
sigma2.null <- sigma2

# New sigma matrices.
sigma1 <- B %*% t(B)
sigma2 <- B %*% L %*% t(B)

# New beta.null matrix.
beta.null <- beta

# New beta matrix.
beta <- beta.gen(cov1 = sigma1, cov2 = sigma2)
}

if(random.start == FALSE){
  volchange <- list(beta, sigma1, sigma2, B, L, opm_mle, opm_mle$convergence,
                  opm_se, u, data)
  names(volchange) <- c("beta", "sigma1", "sigma2", "B", "Lambda", "MLE",
                    "MLE convergence", "SE of the estimators", "residuals",
                    "data")
  rownames(volchange$beta) <- c("GDP.constant", "EMP.constant", "GDP.GDP.Lag1",
                              "EMP.GDP.Lag1", "GDP.EMP.Lag1", "EMP.EMP.Lag1",
                              "GDP.GDP.Lag2", "EMP.GDP.Lag2", "GDP.EMP.Lag2",
                              "EMP.EMP.Lag2", "GDP.GDP.Lag3", "EMP.GDP.Lag3",
                              "GDP.EMP.Lag3", "EMP.EMP.Lag3")
  colnames(volchange$beta) <- c("parameter estimates")
}

```

```

}

if(random.start == TRUE){
  true.beta <- matrix(c(const[1], const[2], A[1], A[7], A[2], A[8], A[3],
                        A[9], A[4], A[10], A[5], A[11], A[6], A[12]), ncol = 1)
  volchange <- list(true.beta, beta, B, L, u, y)
  names(volchange) <- c("true parameters", "beta", "B", "Lambda", "simulated
                        residuals", "simulated data")
  rownames(volchange$beta) <- c("GDP.constant", "EMP.constant", "GDP.GDP.Lag1",
                                "EMP.GDP.Lag1", "GDP.EMP.Lag1", "EMP.EMP.Lag1",
                                "GDP.GDP.Lag2", "EMP.GDP.Lag2", "GDP.EMP.Lag2",
                                "EMP.EMP.Lag2", "GDP.GDP.Lag3", "EMP.GDP.Lag3",
                                "GDP.EMP.Lag3", "EMP.EMP.Lag3")
  rownames(volchange$'true parameters') <- c("GDP.constant", "EMP.constant",
                                              "GDP.GDP.Lag1", "EMP.GDP.Lag1",
                                              "GDP.EMP.Lag1", "EMP.EMP.Lag1",
                                              "GDP.GDP.Lag2", "EMP.GDP.Lag2",
                                              "GDP.EMP.Lag2", "EMP.EMP.Lag2",
                                              "GDP.GDP.Lag3", "EMP.GDP.Lag3",
                                              "GDP.EMP.Lag3", "EMP.EMP.Lag3")
  colnames(volchange$beta) <- c("parameter estimates")
}

return(volchange)

}

#####
# Simulation test
#####

set.seed(2017)
n.sim <- 500
beta.sim <- matrix(NA, nrow = n.sim, ncol = 14)
u.sim <- array(NA, dim = c(300-3, 2, n.sim))
y.sim <- array(NA, dim = c(300, 2, n.sim))

for(i in 1:n.sim){
  out.sim <- try(opt(data = NA, var.index = 2:3, p = 3, K = 2, TB = 201,
                    random.start=TRUE, h=0))
  beta.sim[i,] <- try(getElement(out.sim, 'beta'))
  u.sim[,i] <- try(getElement(out.sim, 'simulated residuals'))
  y.sim[,i] <- try(getElement(out.sim, 'simulated data'))
  print(i)
}

beta <- as.numeric(beta.sim)
beta <- matrix(beta, ncol = 14)
beta <- na.omit(beta)

```

```

par(mfrow=c(2,2))
for(i in 1:14){
  hist(as.numeric(beta[c(1:499),i]), breaks = 5, prob = TRUE,
       xlab = "parameter estimates in 1000 simulations", main = rownames(out$beta)[i])
  lines(density(na.omit(beta[c(1:499), i])), col = "blue")
  abline(v = out.sim$`true parameters`[i], col = "red")
}

est.mean <- apply(beta, 2, mean)
est.sd <- apply(beta, 2, sd)
est.par <- cbind(out.sim$`true parameters`, est.mean, est.sd)
rownames(est.par) <- c("GDP.constant", "EMP.constant", "GDP.GDP.Lag1", "EMP.GDP.Lag1",
                      "GDP.EMP.Lag1", "EMP.EMP.Lag1", "GDP.GDP.Lag2", "EMP.GDP.Lag2",
                      "GDP.EMP.Lag2", "EMP.EMP.Lag2", "GDP.GDP.Lag3", "EMP.GDP.Lag3",
                      "GDP.EMP.Lag3", "EMP.EMP.Lag3")
colnames(est.par) <- c("true parameters", "est.mean", "est.sd")
kable(est.par)

#####
# Comparison of the predictions through simulation
#####

load("data.year.NL.RData")
NL <- na.omit(data.frame(data.year.NL$Time,
                        data.year.NL$NewGrowthGDP,
                        data.year.NL$GrowthEmp))
colnames(NL) <- c("Year", "GDP growth rate", "Employment growth rate")
dat <- NL[-c(56:57), ] # 1961-2015

varest <- VAR(dat[, 2:3], p = 3, type = "const")

out <- opt(data = dat, var.index = 2:3, p = 3, K = 2, TB = 2007, random.start=FALSE)
kable(out$beta)

# OPT in-sample forecast.
gdp.opt <- emp.opt <- NULL
for(i in 4:55){
  gdp.opt[i] <- out$beta[1] + out$beta[3] * dat[i-1,2] + out$beta[5] * dat[i-1,3] +
    out$beta[7] * dat[i-2,2] + out$beta[9] * dat[i-2,3] + out$beta[11] * dat[i-3,2] +
    out$beta[13] * dat[i-3,3]
  emp.opt[i] <- out$beta[2] + out$beta[4] * dat[i-1,2] + out$beta[6] * dat[i-1,3] +
    out$beta[8] * dat[i-2,2] + out$beta[10] * dat[i-2,3] + out$beta[12] * dat[i-3,2] +
    out$beta[14] * dat[i-3,3]
}

# VAR in-sample forecast
gdp.var <- c(rep(NA,3),fitted(varest)[,1])
emp.var <- c(rep(NA,3),fitted(varest)[,2])

```

```

predict.GDP <- data.frame(dat[,2], gdp.var, gdp.opt)
predict.EMP <- data.frame(dat[,3], emp.var, emp.opt)
ts.yr.predict.GDP <- ts(predict.GDP, start = 1961, end = 2015, freq = 1)
ts.yr.predict.EMP <- ts(predict.EMP, start = 1961, end = 2015, freq = 1)

#Plotting realisations and forecast values
plot(ts.yr.predict.GDP[,1], plot.type = "s")
lines(ts.yr.predict.GDP[,2], pch=22, lty=2, col="red")
lines(ts.yr.predict.GDP[,3], pch=22, lty=2, col="green")
legend("bottomleft", c("GDP growth rate", "VAR predicted GDP growth rate",
                      "OPT predicted GDP growth rate"), lty = c(1, 22),
      col = c("black", "red", "green"))

plot(ts.yr.predict.EMP[,1], plot.type = "s", ylim = c(-0.04, 0.05))
lines(ts.yr.predict.EMP[,2], pch=22, lty=2, col="red")
lines(ts.yr.predict.EMP[,3], pch=22, lty=2, col="green")
legend("bottomleft", c("Employment growth rate", "VAR predicted employment growth rate",
                      "OPT predicted employment growth rate"), lty = c(1, 22),
      col = c("black", "red", "green"))

sum((ts.yr.predict.GDP[-c(1:3),1] - ts.yr.predict.GDP[-c(1:3),2])^2) #RSS GDP var
sum((ts.yr.predict.GDP[-c(1:3),1] - ts.yr.predict.GDP[-c(1:3),3])^2) #RSS GDP opt

sum((ts.yr.predict.EMP[-c(1:3),1] - ts.yr.predict.EMP[-c(1:3),2])^2) #RSS employment var
sum((ts.yr.predict.EMP[-c(1:3),1] - ts.yr.predict.EMP[-c(1:3),3])^2) #RSS employment opt

```

## G.6 R code Chapter 7

```

#####
# Predictions: (S)VAR and identification of shocks via a change in volatility
#####

load("data.year.NL.RData")
NL <- na.omit(data.frame(data.year.NL$Time,
                        data.year.NL$NewGrowthGDP,
                        data.year.NL$GrowthEmp))
colnames(NL) <- c("Year", "GDP growth rate", "Employment growth rate")
dat <- NL[-c(56:57), ] # 1961-2015

varest <- VAR(dat[, 2:3], p = 3, type = "const")

out <- opt(data = dat, var.index = 2:3, p = 3, K = 2, TB = 2007, random.start=FALSE)
kable(out$beta)

# OPT in-sample predictions.
gdp.opt <- emp.opt <- NULL
for(i in 4:55){

```

```

gdp.opt[i] <- out$beta[1] + out$beta[3] * dat[i-1,2] + out$beta[5] * dat[i-1,3] +
  out$beta[7] * dat[i-2,2] + out$beta[9] * dat[i-2,3] + out$beta[11] * dat[i-3,2] +
  out$beta[13] * dat[i-3,3]
emp.opt[i] <- out$beta[2] + out$beta[4] * dat[i-1,2] + out$beta[6] * dat[i-1,3] +
  out$beta[8] * dat[i-2,2] + out$beta[10] * dat[i-2,3] + out$beta[12] * dat[i-3,2] +
  out$beta[14] * dat[i-3,3]
}

# VAR in-sample predictions.
gdp.var <- c(rep(NA,3),fitted(varest)[,1])
emp.var <- c(rep(NA,3),fitted(varest)[,2])

predict.GDP <- data.frame(dat[,2], gdp.var, gdp.opt)
predict.EMP <- data.frame(dat[,3], emp.var, emp.opt)
ts.yr.predict.GDP <- ts(predict.GDP, start = 1961, end = 2015, freq = 1)
ts.yr.predict.EMP <- ts(predict.EMP, start = 1961, end = 2015, freq = 1)

#Plotting realisations and forecast values
plot(ts.yr.predict.GDP[,1], plot.type = "s")
lines(ts.yr.predict.GDP[,2], pch=22, lty=2, col="red")
lines(ts.yr.predict.GDP[,3], pch=22, lty=2, col="green")
legend("bottomleft", c("GDP growth rate", "VAR predicted GDP growth rate",
  "OPT predicted GDP growth rate"), lty = c(1, 22),
  col = c("black", "red", "green"))

plot(ts.yr.predict.EMP[,1], plot.type = "s", ylim = c(-0.04, 0.05))
lines(ts.yr.predict.EMP[,2], pch=22, lty=2, col="red")
lines(ts.yr.predict.EMP[,3], pch=22, lty=2, col="green")
legend("bottomleft", c("Employment growth rate", "VAR predicted employment growth rate",
  "OPT predicted employment growth rate"), lty = c(1, 22),
  col = c("black", "red", "green"))

sum((ts.yr.predict.GDP[-c(1:3),1] - ts.yr.predict.GDP[-c(1:3),2])^2) #RSS GDP var
sum((ts.yr.predict.GDP[-c(1:3),1] - ts.yr.predict.GDP[-c(1:3),3])^2) #RSS GDP opt

sum((ts.yr.predict.EMP[-c(1:3),1] - ts.yr.predict.EMP[-c(1:3),2])^2) #RSS employment var
sum((ts.yr.predict.EMP[-c(1:3),1] - ts.yr.predict.EMP[-c(1:3),3])^2) #RSS employment opt

# Out-of-sample predictions using VAR
predictions15 <- predict(varest15, n.ahead = 10, ci = 0.95)

#####
# Comparison of the predictions through simulation
#####

h <- 10
p <- 3
#VAR coefficients
A <- c(0.504914292, 0.040148983, 0.006361527, -0.433609669, 0.237462457, 0.232210695,

```

```

0.256568557, 0.571149051, -0.106114636, 0.062737666, -0.126112307, -0.122209395)
const <- c(0.008180638, 0.005025498)
set.seed(2017)
out.sim <- opt(data = NA, var.index = 2:3, p = 3, K = 2, TB = 201,
               random.start = TRUE, h = 0)
#opt coefficients
beta.sim <- getElement(out.sim, 'beta')
#Simulated sample
sample.sim <- out.sim$`simulated data`
#Simulated predictions
pred.sim <- matrix(rnorm(2*h), ncol=2)
#Simulated sample and predictions
y <- rbind(sample.sim, pred.sim)

#In-sample predictions using VAR
pred <- y
for(i in 4:300){
  pred[i,1] <- const[1] + A[1] * pred[i-3, 1] + A[2] * pred[i-3, 2] + A[3] *
    pred[i-2, 1] + A[4] * pred[i-2, 2] + A[5] * pred[i-1, 1] + A[6] * pred[i-1, 2]
  pred[i,2] <- const[2] + A[7] * pred[i-3, 1] + A[8] * pred[i-3, 2] + A[9] *
    pred[i-2, 1] + A[10] * pred[i-2, 2] + A[11] * pred[i-1, 1] + A[12] * pred[i-1, 2]
}

#Accuracy of in-sample-prediction with VAR
RSS.var.gdp <- sum((y[1:300, 1] - pred[1:300, 1]) ^ 2)
RSS.var.emp <- sum((y[1:300, 2] - pred[1:300, 2]) ^ 2)
RSS.var <- RSS.var.gdp + RSS.var.emp

#In-sample predictions with opt
pred.opt <- y
for(i in 4:300){
  pred.opt[i,1] <- const[1] + beta.sim[1] * pred.opt[i-3, 1] + beta.sim[2] *
    pred.opt[i-3, 2] + beta.sim[3] * pred.opt[i-2, 1] + beta.sim[4] * pred.opt[i-2, 2]
  + beta.sim[5] * pred.opt[i-1, 1] + beta.sim[6] * pred.opt[i-1, 2]
  pred.opt[i,2] <- const[2] + beta.sim[7] * pred.opt[i-3, 1] + beta.sim[8] *
    pred.opt[i-3, 2] + beta.sim[9] * pred.opt[i-2, 1] + beta.sim[10] * pred.opt[i-2, 2]
  + beta.sim[11] * pred.opt[i-1, 1] + beta.sim[12] * pred.opt[i-1, 2]
}

#Accuracy of in-sample-prediction with VAR
RSS.opt.gdp <- sum((y[1:300, 1] - pred.opt[1:300, 1]) ^ 2)
RSS.opt.emp <- sum((y[1:300, 2] - pred.opt[1:300, 2]) ^ 2)
RSS.opt <- RSS.opt.gdp + RSS.opt.emp

#In-of-sample comparison between VAR and opt.
RSS.opt < RSS.var

#Out-of-sample predictions with VAR.
pred <- y

```



```

for(i in 301:310){
  pred[i,1] <- const[1] + A[1] * pred[i-1, 1] + A[2] * pred[i-1, 2] + A[3] *
    pred[i-2, 1] + A[4] * pred[i-2, 2] + A[5] * pred[i-3, 1] + A[6] * pred[i-3, 2]
  pred[i,2] <- const[2] + A[7] * pred[i-1, 1] + A[8] * pred[i-1, 2] + A[9] *
    pred[i-2, 1] + A[10] * pred[i-2, 2] + A[11] * pred[i-3, 1] + A[12] * pred[i-3, 2]
}
#Accuracy of out-of-sample-prediction with VAR
MSE.var.gdp <- sum((y[301:310, 1] - pred[301:310, 1]) ^ 2) / h
MSE.var.emp <- sum((y[301:310, 2] - pred[301:310, 2]) ^ 2) / h
MSE.var <- MSE.var.gdp + MSE.var.emp

#Out-of-sample predictions with opt.
pred.opt <- y
for(i in 301:310){
  pred.opt[i,1] <- beta.sim[1] + beta.sim[3] * pred.opt[i-1, 1] + beta.sim[5] *
    pred.opt[i-1, 2] + beta.sim[7] * pred.opt[i-2, 1] + beta.sim[9] * pred.opt[i-2, 2]
  + beta.sim[11] * pred.opt[i-3, 1] + beta.sim[13] * pred.opt[i-3, 2]
  pred.opt[i,2] <- beta.sim[2] + beta.sim[4] * pred.opt[i-1, 1] + beta.sim[6] *
    pred.opt[i-1, 2] + beta.sim[8] * pred.opt[i-2, 1] + beta.sim[10] * pred.opt[i-2, 2]
  + beta.sim[12] * pred.opt[i-3, 1] + beta.sim[14] * pred.opt[i-3, 2]
}

#Accuracy of out-of-sample prediction with optimization.
MSE.opt.gdp <- sum((y[301:310, 1] - pred.opt[301:310, 1]) ^ 2) / h
MSE.opt.emp <- sum((y[301:310, 2] - pred.opt[301:310, 2]) ^ 2) / h
MSE.opt <- MSE.opt.gdp + MSE.opt.emp

#Out-of-sample comparison between VAR and opt.
MSE.opt < MSE.var

```

## 10 References

- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21(1), 243-247.
- Akaike, H. (1970). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, 22(1), 203-217.
- Akkemik, K. A. (2007). The response of employment to GDP growth in Turkey: an econometric estimation. *Applied Econometrics and International Development*, 7(1).
- Bera, A. K., & Jarque, C. M. (1980). Efficient tests for normality, heteroscedasticity, and serial independence of regression residuals. *Economic Letters* 6(3), 255-259.
- Bernanke, B. (1986). Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy*. 1st ed. Cambridge. Mass.: National Bureau of Economic Research.
- Blanchard, O., & Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics* 107, 1329-1368.
- Box, G. E. P., George, E. P., Jenkins, G. M., & Gwilym, M. (1970). *Time Series Analysis: Forecasting and Control*. 1st ed. San Francisco: Holden-Day.
- Chinn, M., Ferrara, L., & Mignon, V. (2014). Explaining US employment growth after the great recession: the role of output-employment non-linearities. *Journal of Macroeconomics* 42, 118-129.
- CPB. (2014). Roads to recovery. *CPB Publishing*. Den Haag.
- CPB. (2014). The Dutch labour market during the Great Recession. *CPB Publishing*. Den Haag.
- CPB. (2015). Mismatch on the Dutch labour market in the Great Recession. *CPB Publishing*. Den Haag.
- Davis, B. (2009). What's a Global Recession? *The Wall Street Journal*.
- ECB Economic Bulletin, Issue 6 / 2016 - Article The employment-GDP relationship since the crisis.
- Elbourne, A., & Teulings, C. (2011). The potential of a small model. *CPB discussion paper*, 193.
- Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature*, 46(1), 3-56.
- Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4), 813-836.
- Enders, W. (1995). *Applied econometric time series*. 1st ed. New York: Wiley.
- Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica* 50, 987-1007.
- Fernandez-Villaverde, J., Rubio-Ramirez, J., Sargent T., & Watson M. (2007). ABCs (and Ds) of understanding VARs. *The American Economic Review*, 97(3), 1021-1026.

- Greene, W. H. (2008). *Econometric analysis*. 1st ed. Upper Saddle River. NJ: Pearson.
- Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. 1st ed. New York: Springer-Verlag.
- Hamilton, J. D. (1994). *Time series analysis*. 1st ed. Princeton N.J.: Princeton University Press.
- Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, 41, 190-195.
- Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review* 55, 163-172.
- Kazutoshi, K., (1992). *Employment security and labor market flexibility: an international perspective*. 1st ed. Detroit: Wayne State University Press.
- Kitov, I., & Kitov, O. (2011). Employment, unemployment and real economic Growth. *SSRN Electronic Journal*, doi:10.2139/ssrn.1894209
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3), 159-178.
- Lanne, M., Lütkepohl, H., Maciejowska, K. (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control*, 34(2), 121-131.
- Lucas, R (1976). Econometric Policy Evaluation: A Critique. The Phillips Curve and Labor Markets. *Carnegie-Rochester Conference Series on Public Policy*. 19-46.
- Lupi, C. (2009). Unit root CADF testing with R. *Journal of statistical software*, 71(2).
- Lütkepohl, H. (2004). *Applied time series econometrics*. 1st ed. Cambridge, UK: Cambridge University Press.
- Lütkepohl, H. (2012). Identifying Structural Vector Autoregressions Via Changes in Volatility. *SSRN Electronic Journal SSRN Journal*.
- Lutkepohl, H., Kratzig, M., Boreiko, D. (2006). *VAR analysis in JMulTi*. 1st ed.
- Lütkepohl, H. (2006). *New introduction to multiple time series analysis with 36 tables (3rd ed.)*. Berlin.
- Okun, A.M. (1962). Potential GNP: Its Measurement and Significance, Proceedings of the Business and Economics Statistics Section. *American Statistical Association*, 98-103.
- Metcalfe, A.V. (2009). *Introductory time series with R*. New York: Springer.
- Pfaff, B. (2006). *Analysis of integrated and cointegrated time series with R*. New York: Springer.
- Pfaff, B. (2008). VAR, SVAR and SVEC models: implementation within R packages vars. *Journal of statistical software*, 27(4).
- Raoufinia, K. (2016). Forecasting Employment Growth in Sweden Using a Bayesian VAR model. *National Institute of Economic Research*, 144.
- Schwarz, G. (1978), Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Sims, C. A. (1981). An autoregressive index model for the U.S. 1948-1975, in J. Kmenta & J. B. Ramsey, *Large-Scale Macro-Econometric Models*, 283-327.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48, 1-48.

- Sims, C. A. (1986). Are forecasting models usable for policy analysis? *Quarterly Review, Federal Reserve Bank of Minneapolis* 10, 2-16.
- Stock, J., & Watson, M. (2012). Disentangling the channels of the 2007-2009 recession. *Brooking Papers on Economic Activity*, forthcoming.
- Van der Vaart. (2013) *Time series*. Universiteit Leiden. Version 242013.