KU LEUVEN

FACULTY OF SCIENCE

The Variability of the Belgian Business Survey Indicator and its Predictive Power

Fabrice VAN BOECKEL

Co-Supervisor: Prof. G. Molenberghs
KU Leuven

Co-Supervisor: L. Van Belle
National Bank of Belgium

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Statistics

Academic year 2018-2019

Acknowledgement

Thanks my family and friends -

the National Bank of Belgium Laurent Van Belle, Jean Palate, David ... and all the members of the Department of Research and Development who where always there to answer any of my questions.

... -

KULeuven

"Statistics are the heart of democracy."

- Simeon Strunsky

Abstract

This Master Thesis explores the Variance of the Belgian Business Survey. Several finding concerning the nature and properties of the Variance are found as the bounds and relation with the mean.

In a second part, the predictive power of the variance is examined and it's found that ...

It's also the first time that à Markov Switching model is used in this context. It was showed that ...

Samenvatting

Deze master

Keywords

Business Surveys - Business Barometer - Trichotomous Observations - Survey Variance - Survey Volatility - Markov Switching -

Contents

CHAPTER 1

Introduction

Business Survey Indicator / Business Barometer / Business Confidence Indicator

A widespread method to predict the evolution of National Economies is the survey-based Business indicator. Belgium have been collecting this indicator for more than 60 years. This long evolution

- Talk about tradition of improving BSB

This Thesis is included in the continuity of a long tradition of papers proposing improvement and ways to add value to the Business Barometer (.....) will propose ways to add information to the Belgian Business Barometer, that could also be applied to others

Since 1968, the National Bank of Belgium publishes each month the national

CHAPTER 2

The Business Survey Indicator

This first chapter is a more general description of the Belgian Business Survey Indicator, that we will also call the Business Barometer. We will present it different calculations, the weighting that are applied and

explain the two types of weightings

2.1 History

In 1954 started the Business Survey of the National Bank of Belgium.

At the European level, in 1961, is launched the launched by the Commission decision of 15 November 1961. The first survey was the harmonised business survey in the manufacturing industry conducted in 1962. Since then, the sector coverage of the programme has widened considerably. The BCS programme was extended to the construction sector and to investment plans in the manufacturing sector in 1966, to consumers in 1972, to the retail trade in 1984, and to the services sector in 1996. More info see (Commission)

Almost 20 years later, in 1972, results where first synthesised in an indicator. It has since then been well known for been a very informative and useful indicator. For example in 1999 an article was published in the Wall Street Journal titled *"Euroland Discovers A Surprise Indicator: Belgian Confidence"* (Rhoads, 1999) that noted the accuracy and of it predictions.

Today, Nowcasting is a well studied subject and very important for public and private organisation to have an as clear as possible view of the state of the economy.

With the study of business cycles, that will be looked at more in details in part ...., nowcasting are two important objectives that will be looked at in this work.

Due to the structure of Belgian economy and ... it was also shown as been a good predictor for the European economy Vanhaelen et al. (2000)

changes in 2009 see later (section 2.3)

2.2 Sampling Method

...

2.3 Objective and Methodology

A large improvement of the indicator happened in 2009 which was explained in a working paper untitled "The National Bank of Belgium's new business survey indicator" (De Greef and Van Nieuwenhuyze)

Objectives :

Using specific questions to identify the business cycles and turning points

It can be included in what's now called nowcasting, which is now and short time forecasting of the evolution of the economy.

- only take a limited amount of questions into account, the most relevant ones (3-4 questions)
- inclusion of the services in the global indicator
- lighten smoothing method

The Business Cycles Theory

Before going further, a small summary of business cycle theory.

As already said, an important role of the Business Survey Indicator, is to identify business cycles.

The domain is very wide and a lot of books and articles where published about the subject.

cite burns 1946

see citations from De Greef

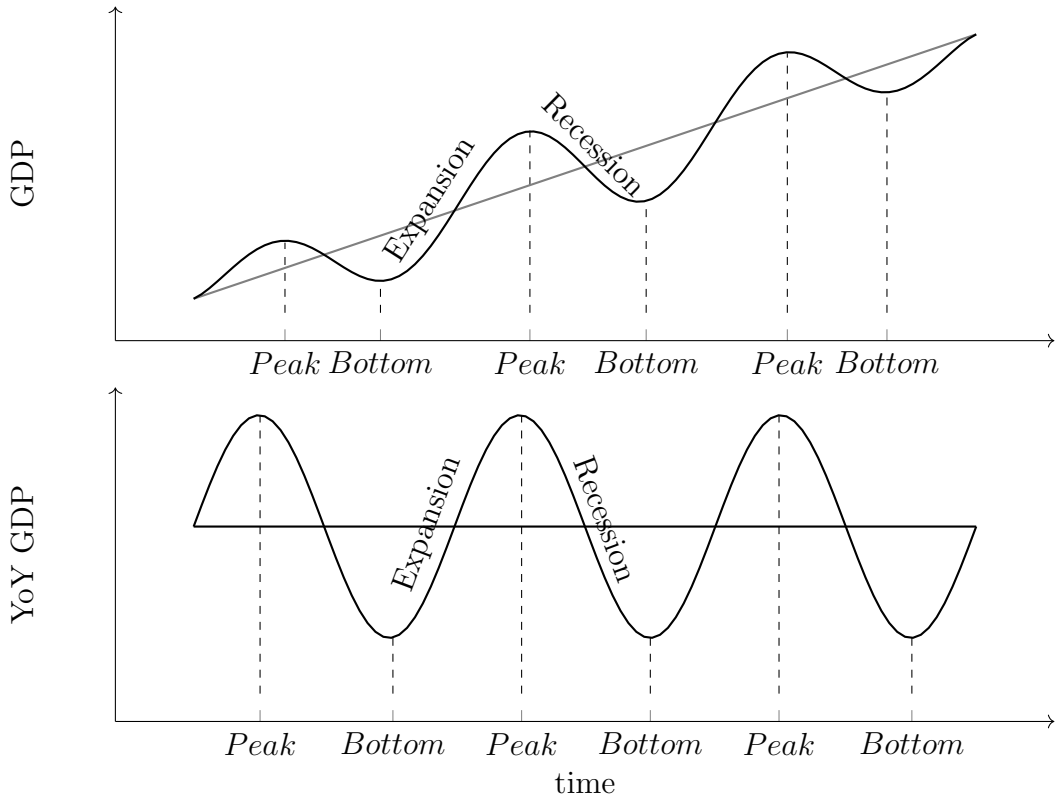


Figure 2.1: The Business Cycle theory of GDP and Year on Year GDP

Quality Criterion

- high correlation with Economical Indicators
- fluctuation that's mostly explained by the conjuncture
- predictive power for the future months
  more information can be found in De Greef and Van Nieuwenhuyze (2009)

2.3.1 Weighting procedure

According to the size of the company, measured different ways depending on the sector .....

2.3.2 Globalisation procedure

2.4 Questionnaire

General presentation

.... questionnaire can be found in appendix

Questions RS975

Questions taken into account for RS975: ....

originally question Q18, 27, 32 and 33, for simplicity numbered here as 1, 2, 3 and 4. see appendix page 48

2.5 Calculation of the Indicator

This section presents the method for calculating the indicator. The calculation in itself is standard, but the different ways to write it are important for interpretation and to better understand the following chapters.

2.5.1 Unweighted Indicator

The calculation of the unweighted indicator for a specific question at a specific time can be written as follow;

$$E(X) = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2.1}$$

where

x_i is the answer of the respondent i and can each take value -1, 0 and 1

n is the total of respondents

Since x can only take three different values, we can decompose it into

$$E(X) = \frac{\sum_{i=1}^{n} x_{+i} + \sum_{i=1}^{n} x_{Ni} + \sum_{i=1}^{n} x_{-i}}{n} \tag{2.2}$$

where

x_{+i}, x_{Ni} and x_{-i} are the positive(+), neutral (N) and negative (-) answers of the respondent i

n is the total of respondents

We know that $\sum_{i=1}^{n} x_{Ni} = 0$ so we can write

$$E(X) = \frac{\sum_{i=1}^{n} x_{+i}}{n} + \frac{\sum_{i=1}^{n} x_{-i}}{n} \tag{2.3}$$

$\sum_{i=1}^{n} x_{+i}/n$ is the proportion of positive answers and $\sum_{i=1}^{n} x_{-i}/n$ is the negative proportion of negative answer so for simplicity we write it

$$E(X) = \pi_+ - \pi_- \tag{2.4}$$

where π_+ and π_- are the proportion of respondents answering positive and negative. π is use here also in the probabilistic way as it can also be seen as the probability that a respondent answers positive, negative or neutral (π_0) with $\pi_+ + \pi_0 + \pi_- = 1$.

2.5.2 Weighted Indicator

Globalisation procedure

Weighting procedure

Calculation

$$E(X) = \frac{\sum_{i=1}^{n} (\omega_i p_i x_i)}{\sum_{i=1}^{n} \omega_i p_i} \tag{2.5}$$

where
x_i is the answer of the respondent i and can each take value -1, 0 and 1
p_i is the weight of the globalisation of the company i
ω_i is the weight of the company i

$$E(X) = \frac{\sum_{i=1}^{n} \omega_{+i} p_{+i} x_{+i} + \sum_{i=1}^{n} \omega_{Ni} p_{Ni} x_{Ni} + \sum_{i=1}^{n} \omega_{-i} p_{-i} x_{-i}}{n} \tag{2.6}$$

$$E(X) = \frac{\sum_{i=1}^{n} \omega_{+i} p_{+i} x_{+i}}{n} + \frac{\sum_{i=1}^{n} \omega_{-i} p_{-i} x_{-i}}{n} \tag{2.7}$$

$$E(X) = \pi_+ - \pi_- \tag{2.8}$$

where π_+ and π_- are the **weighted** proportion of respondents answering positive and negative. π is use here also in the probabilistic way as it can also be seen as the probability that a respondent answers positive, negative or neutral (π_0) with $\pi_+ + \pi_0 + \pi_- = 1$.

2.5.3 Properties

E(X) has -1 as lower bound and 1 as upper bound

2.5.4 Take different questions into account

The previous calculations where specific to each question. The published indicators are usually taking different survey questions into account. For example the Industry indicator that we will be interested in is composed of four questions that have all the same weight:

$$\text{Industry Business Indicator } = \frac{E(X_{Q1}) + E(X_{Q2}) + E(X_{Q3}) + E(X_{Q4})}{4} \tag{2.9}$$

where $E(X_{Q1})$, $E(X_{Q2})$, $E(X_{Q3})$ and $E(X_{Q4})$ are the different averages for question 18, 27, 32 and 33 (can be weighted or unweighted)

can also been seen as the mean of the answers for each participant at each period then combined together

$$\text{Industry Business Indicator } = \frac{\sum_{i=1}^{n} (x_{iQ1} + x_{iQ2} + x_{iQ3} + x_{iQ4})/4}{n} \tag{2.10}$$

CHAPTER 3

Variance of the Indicator

The variance is, with the mean, one of the first tool for Statisticians to study a certain variable.

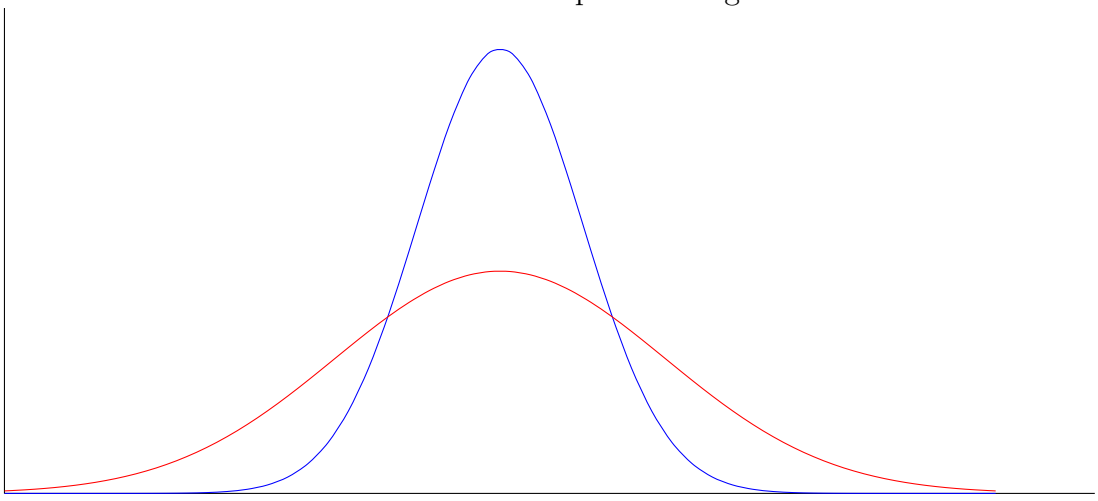
In the context of the Business Survey, the variance haven't been used much.

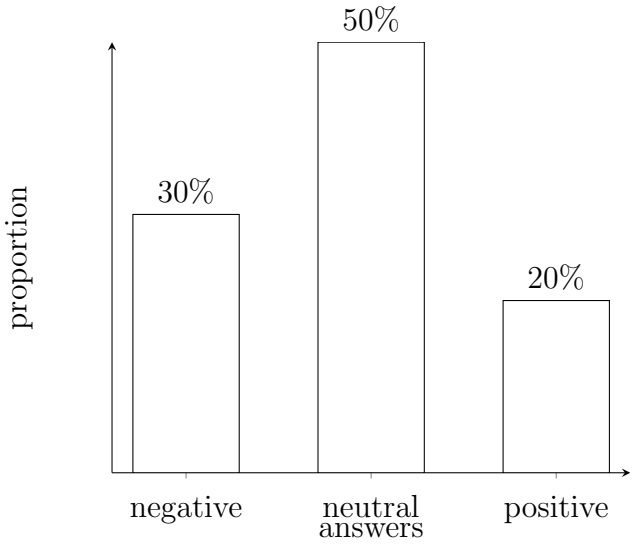
The added info of the variance

variance as a measure of the dispersion of the responses

In statistics, the variance has two m

Here we want to know to what extend respondents agree with each others.





difference sampling error and variance

here variance is a measure of the "dispersion" of the answers.

difference between nominal and continuous variable variance

As done for the Indicator, two different variances will be take into account here, the weighted and the unweighted variance of the indicator.

3.1 Variance of the Unweighted Indicator

The main variance

cite

$$\begin{aligned}
\mathrm{Var}(X) &= E\left[(X - E(X))^2\right] \\
&= E\left(X^2\right) - E\left(X\right)^2 \\
&= \left(\frac{\sum_{i=1}^{n} x_{+i}^2}{n}\right) + \left(\frac{\sum_{i=1}^{n} x_{Ni}^2}{n}\right) + \left(\frac{\sum_{i=1}^{n} x_{-i}^2}{n}\right) - E(X)^2 \\
&= \pi_+ + \pi_- - (\pi_+ - \pi_-)^2 \qquad (3.1)
\end{aligned}$$

Since $\left(\frac{\sum_{i=1}^{n} x_{Ni}^2}{n}\right) = 0$, $x_{+i}^2 = x_{+i}$, $x_{-i}^2 = x_{-i}$ and $E(X) = \pi_+ - \pi_-$

We then have several different ways to write the previous equation;

$$\begin{aligned}
\mathrm{Var}(X) &= \pi_+ + \pi_- - (\pi_+ - \pi_-)^2 \\
&= \pi_+ + \pi_- - E(X)^2 \qquad (3.2) \\
&= 1 - \pi_n - E(X)^2 \qquad (3.3)
\end{aligned}$$

Equation 3.1 is interesting
Equation 3.2
Equation 3.3

3.2 Generalization for Weighted Indicators / Variance of the Weighted Indicator

$$\mathrm{Var}(X) = \frac{1}{\sum \omega_i p_i} \sum_{i=1}^{N} \omega_i p_i (X_i - \bar{X})^2 \tag{3.4}$$

$$\begin{aligned} \mathrm{Var}(X) &= \pi_+ + \pi_- - (\pi_+ - \pi_-)^2 && (3.5)\\ &= \pi_+ + \pi_- - E(X)^2 && (3.6)\\ &= 1 - \pi_0 - E(X)^2 && (3.7) \end{aligned}$$

3.2.1 Properties

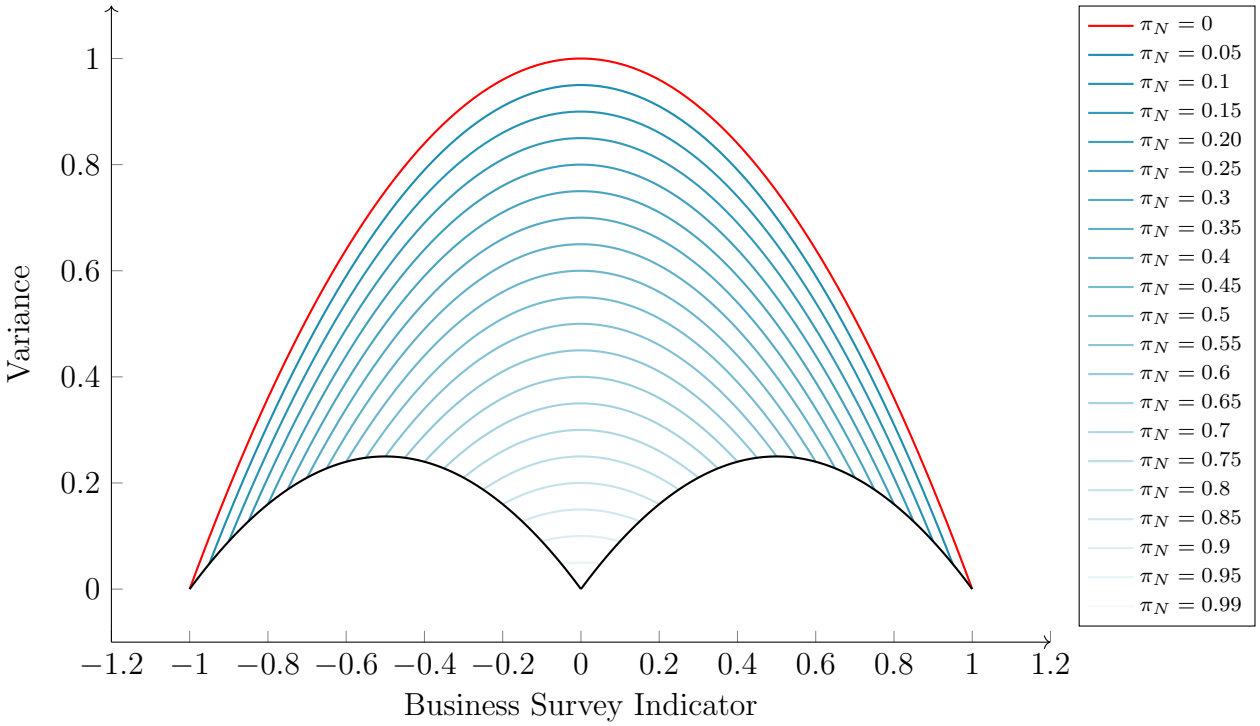


Figure 3.1: Plot of the BSI on it's variance for different values of π_N ($\pi_N = 1 - \pi_+ - \pi_-$)

Property 1: The variance of X is bounded between -1 and 1

Property 2: The variance = A5 and E(X)

Property 3: .....

upper bound :

$$\mathrm{Var}(X) = 1 - 0 - \mathrm{E}(X)^2 \tag{3.8}$$

lower bound :

$$\begin{aligned} \mathrm{Var}(X) &= -(\mathrm{E}(X) + 0,5)^2 + 0,25 \quad \text{where } \mathrm{Var}(X) \text{ is between -1 and 0} && (3.9)\\ &= X - X^2 && (3.10) \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= -(\mathrm{E}(X) - 0,5)^2 + 0,25 \quad \text{where } \mathrm{Var}(X) \text{ is between 0 and 1} && (3.11)\\ &= -X - X^2 && (3.12) \end{aligned}$$

3.3 Take different questions into account

Since

$$\mathrm{Var}\left(\frac{X_{Q1}+X_{Q2}+X_{Q3}+X_{Q4}}{4}\right) \neq \frac{\mathrm{Var}(X_{Q1})+\mathrm{Var}(X_{Q2})+\mathrm{Var}(X_{Q3})+\mathrm{Var}(X_{Q4})}{16}$$

$$\mathrm{Var}\left(\frac{X+Y}{2}\right) = \frac{1}{4}\left(\mathrm{Var}(X)+\mathrm{Var}(Y)+\mathrm{Cov}(X,Y)\right) \tag{3.13}$$

$$\begin{aligned}\mathrm{Var}\left(\frac{X_{Q1}+X_{Q2}+X_{Q3}+X_{Q4}}{4}\right) &= \frac{1}{16}[\mathrm{Var}(X_{Q1})+\mathrm{Var}(X_{Q2})+\mathrm{Var}(X_{Q3})+\mathrm{Var}(X_{Q4})\\ &\quad +\mathrm{Cov}(X_{Q1},X_{Q2})+\mathrm{Cov}(X_{Q1},X_{Q3})+\mathrm{Cov}(X_{Q1},X_{Q4})\\ &\quad +\mathrm{Cov}(X_{Q2},X_{Q3})+\mathrm{Cov}(X_{Q2},X_{Q4})+\mathrm{Cov}(X_{Q3},X_{Q4})]\end{aligned}$$

$$\begin{aligned}\mathrm{Var}\left(\frac{X_{Q1}+X_{Q2}+X_{Q3}+X_{Q4}}{4}\right) &= \mathrm{Var}\left(\frac{\sum_{i=1}^{n}\left(x_{iQ1}+x_{iQ2}+x_{iQ3}+x_{iQ4}\right)}{4n}\right)\\ &= \mathrm{Var}\left(\frac{\pi_{1+}+\pi_{2+}+\pi_{3+}+\pi_{4+}-\pi_{1-}-\pi_{2-}-\pi_{3-}-\pi_{4-}}{4}\right)\\ &= \frac{1}{16}\mathrm{Var}\left(\pi_{1+}+\pi_{2+}+\pi_{3+}+\pi_{4+}-\pi_{1-}-\pi_{2-}-\pi_{3-}-\pi_{4-}\right)\end{aligned} \tag{3.14}$$

$$\begin{aligned}\mathrm{Var}\left(\frac{X_{Q1}+X_{Q2}+\ldots+X_{Qn}}{n}\right) &= \mathrm{Var}\left(\frac{\pi_{1+}+\pi_{2+}+\ldots+\pi_{n+}-\pi_{1-}-\pi_{2-}-\ldots-\pi_{n-}}{n}\right)\\ &= \frac{1}{n^2}\mathrm{Var}\left(\pi_{1+}+\pi_{2+}+\ldots+\pi_{n+}-\pi_{1-}-\pi_{2-}-\ldots-\pi_{n-}\right)\end{aligned} \tag{3.15}$$

3.4 Discussion regarding the 'true variance'

There is another way to calculate the variance that have been ignored for, that is calculating the variance for each lowest group of globalisation, and then combine those calculated variances.

Interestingly, it have been calculated for several Questions of the business barometer, and it is approximately 10 times smaller than the variance based on all the answer a ones.

The reasons why it will not be used here - losing information - weight of globalisation taken into account in the weighted variance

CHAPTER 4

Indicator of the Evolution of Individual Responses

Why this new indicator

Also Called Z indicator

Can be better understood as the indicator of the Changes in individual answers between t-1 and t

cite the paper by INSEE

An issue for this indicator was to find an optimal name for it so that it would be easily understand by the largest number.

The Indicator of the Evolution of Individual Responses

Explain the new indicator

		t		
		-	**0**	+
$t-1$	**-**	π_{--}	π_{-0}	π_{-+}
	0	π_{0-}	π_{00}	π_{0+}
	+	π_{+-}	π_{+0}	π_{++}

where

$$\pi_{++} + \pi_{+0} + \pi_{+-} + \pi_{0+} + \pi_{00} + \pi_{0-} + \pi_{-+} + \pi_{-0} + \pi_{--} = 1 \tag{4.1}$$

The Indicator of the evolution of the individual responses can be obtained by

$$E(Z) = \pi_{0+} + \pi_{-0} - \pi_{+0} - \pi_{0-} + 2\pi_{-+} - 2\pi_{+-} \tag{4.2}$$

where π is the proportion/probability of respondent answering (-,0,+) at $t-1$ and t

$$E(Z) = \pi_{0+} + \pi_{-0} - \pi_{+0} - \pi_{0-} + 2\pi_{-+} - 2\pi_{+-} = \begin{array}{c|ccc} & \textbf{-} & \textbf{0} & + \\ \hline \textbf{-} & 0 & +1 & +2 \\ \textbf{0} & -1 & 0 & +1 \\ + & -2 & -1 & 0 \end{array} \tag{4.3}$$

Interpretation

CHAPTER 5

Variance of the Evolution of Individual Responses (Volatility of Responses)

That we will also call the **volatility of the indicator**, in the sens that the variance of the evolution of the indicator account for the dispersion of the difference in answers over a two times period.

In this case, the highers variance of Z, will be obtained when half of the companies went from a negative answer to a positive answer and the other half did the opposite and changed from a positive answer at t-1 to a negative answer at t.

waza see Chapter 5

The idea is that this variance of Z is complementary to the estimation of Z since they have two very interesting but different interpretations. Further interpretation will be

5.1 Presentation

$$\pi_{++} + \pi_{+0} + \pi_{+-} + \pi_{0+} + \pi_{00} + \pi_{0-} + \pi_{-+} + \pi_{-0} + \pi_{--} = 1 \tag{5.1}$$

$$\begin{aligned} Var(Z) &= \pi_{0+} + \pi_{-0} + \pi_{+0} + \pi_{0-} + 4\pi_{-+} + 4\pi_{+-} \\ &\quad -(\pi_{0+} + \pi_{-0} - \pi_{+0} - \pi_{0-} + 2\pi_{-+} - 2\pi_{+-})^2 \\ &= \pi_{0+} + \pi_{-0} + \pi_{+0} + \pi_{0-} + 4\pi_{-+} + 4\pi_{+-} - E(Z)^2 \\ &= 1 - \pi_{++} - \pi_{00} - \pi_{--} + 3\pi_{+-} + 3\pi_{-+} - E(Z)^2 \end{aligned} \tag{5.2}$$

$$
\begin{aligned}
Var(Z) &= \left(\begin{array}{c|ccc} & \textbf{-} & \textbf{0} & + \\ \hline \textbf{-} & 0 & +1 & +4 \\ \textbf{0} & +1 & 0 & +1 \\ + & +4 & +1 & 0 \end{array}\right) - \left(\begin{array}{c|ccc} & \textbf{-} & \textbf{0} & + \\ \hline \textbf{-} & 0 & +1 & +2 \\ \textbf{0} & -1 & 0 & +1 \\ + & -2 & -1 & 0 \end{array}\right)^2 \\
&= \left(\begin{array}{c|ccc} & \textbf{-} & \textbf{0} & + \\ \hline \textbf{-} & 0 & +1 & +4 \\ \textbf{0} & +1 & 0 & +1 \\ + & +4 & +1 & 0 \end{array}\right) - (E(Z))^2 \\
&= 1 + \left(\begin{array}{c|ccc} & \textbf{-} & \textbf{0} & + \\ \hline \textbf{-} & -1 & 0 & +3 \\ \textbf{0} & 0 & -1 & 0 \\ + & +3 & 0 & -1 \end{array}\right) - (E(Z))^2
\end{aligned}
$$

5.1.1 Properties

Property 1: the variance of Z is bounded between -1 and 1

Property 2:

CHAPTER 6

Non-Response, Dropout, Attrition and Seasonal Effects

Table 6.1: Correlation Between Time and different variables

	GDP	YoY GDP	BSI	Var	Z_I	Var_Z_I
Time	−0.223	−0.352	0.148	−0.775	0.060	−0.728

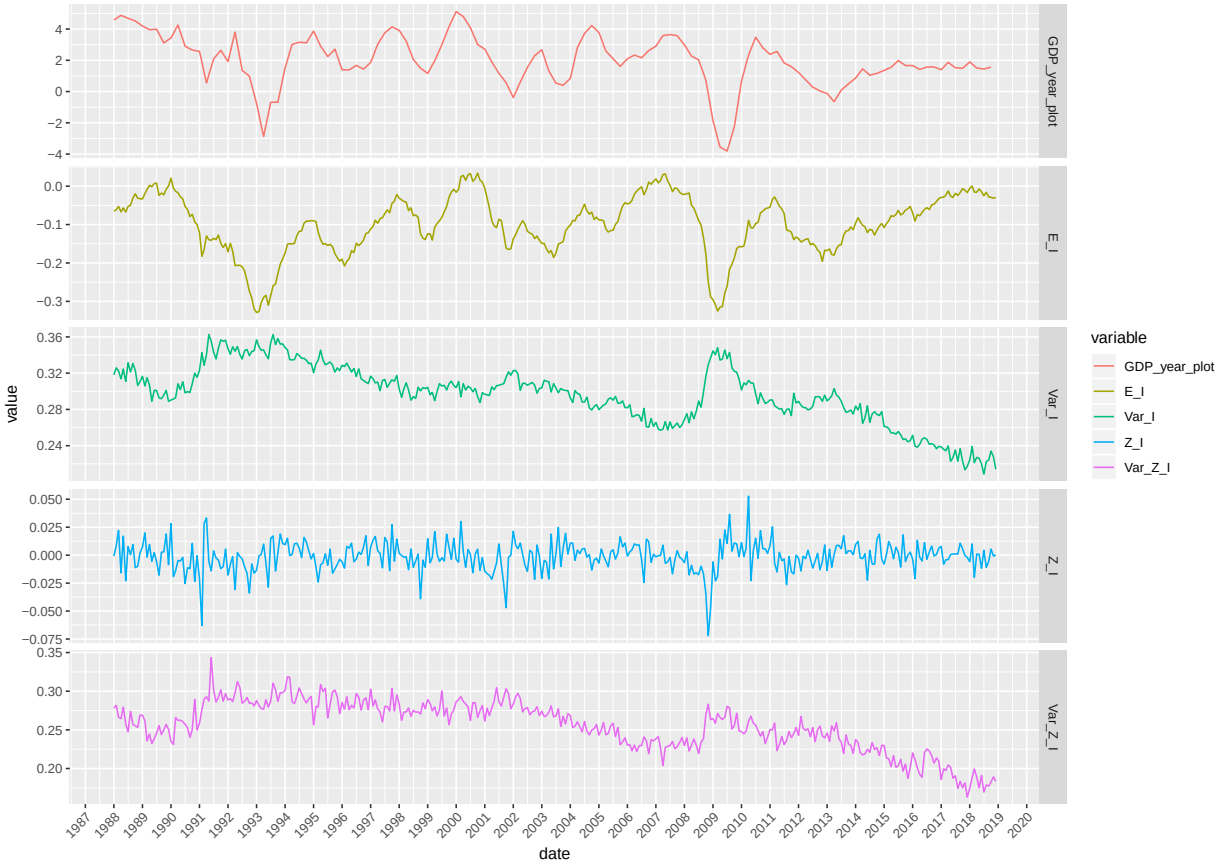


Figure 6.1: Plot

6.1 Non-Response

plot non-response

6.2 Dropout

Non parametric test Das et al. (2011)

6.3 Attrition

Attrition / Panel Conditioning Master Thesis done about the Belgian Labor Force Survey, where attrition was studied Priyana Hardjawidjaksana (2019)

Non parametric test Das et al. (2011)

limitation: only some periods of

6.4 Seasonal Effects

The National Bank, before publishing the Business Survey Indicator, applies a X11 seasonal correction

The literature about seasonal effects is very rich and variate

- NBB developed JDemetra+ and has since been recommended by the ECB and Eurostat for all NSI in Europe.

at the same time the department of Business Survey uses as a X11 adapted method to correct for seasonality because don't want to correct for previous publications.

Methodology

- test for seasonality
- run the analysis without corrections
- apply corrections and see if more accurate

Seasonal correction of the Indicator

Seasonal correction of the Variance

Seasonal correction of the Indicator of the Evolution

Seasonal correction of the Variance of Z

Seasonal correction of the Proportions

6.5 Limitations

explain the issue of seasonal correction on "future data"

CHAPTER 7

Exploratory Analysis

7.1 Data At hand

Four questions
1988 - 2018

7.2 Small vs Large

7.3 By Sector

7.4 Correlations

There are three different correlations that need to be looked at

7.5 Correlation between questions

Table 7.1: Correlation Matrix

	E_1	E_2	E_3	E_4
E_1	1	0.262	0.412	0.416
E_2	0.262	1	0.939	0.876
E_3	0.412	0.939	1	0.938
E_4	0.416	0.876	0.938	1

7.6 Auto-Correlation

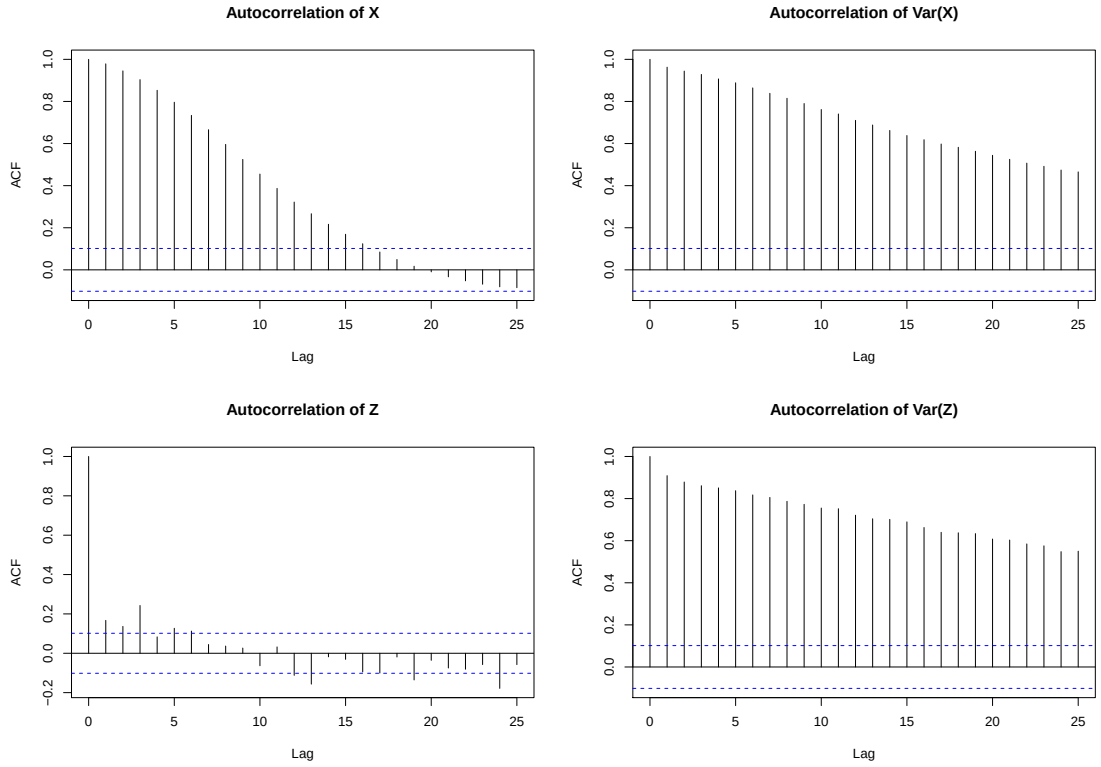


Figure 7.1: Plot

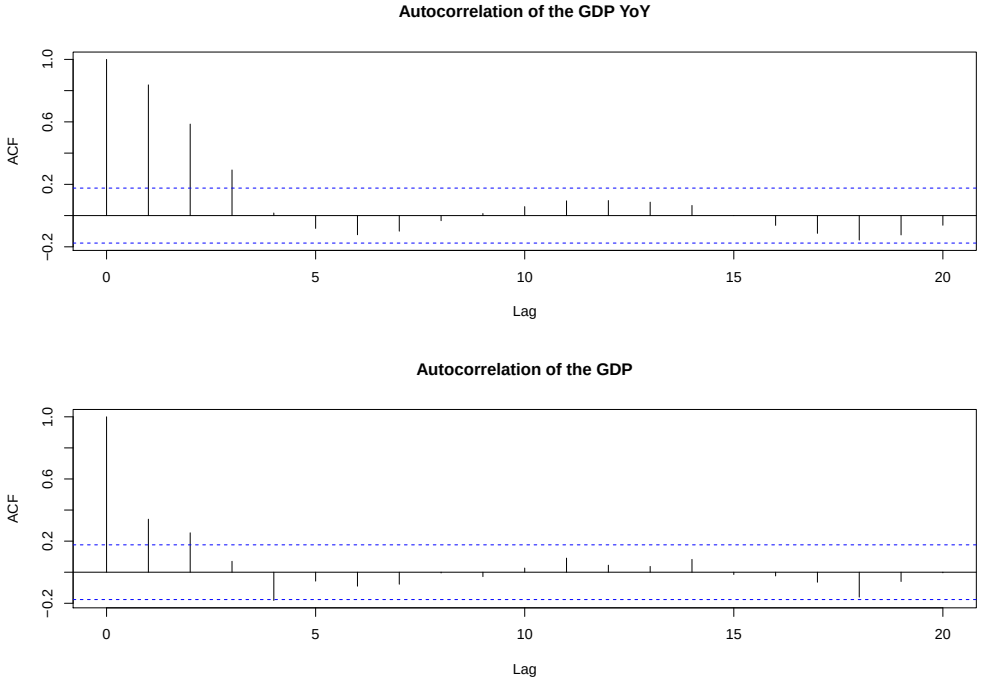


Figure 7.2: Plot

7.7 Correlation with GDP

Belgian industry claims 25% of the labour force in Belgium and have been shown as been the best indicator to predict the year to year GDP citeAlain Quartier and Isabelle

GDP vs GDP YoY

What's Year on Year GDP

$$YoYGDP = \frac{GDP_t - GDP_{t-12}}{GDP_{t-12}} \tag{7.1}$$

Correlation

Table 7.2: Correlation Matrix

	GDP	GDP_year	E_I	E_1	E_2	E_3	E_4
GDP	1	0.628	0.502	0.222	0.439	0.473	0.556
GDP_year	0.628	1	0.707	0.092	0.729	0.703	0.717
E_I	0.502	0.707	1	0.483	0.952	0.982	0.963
E_1	0.222	0.092	0.483	1	0.266	0.414	0.406
E_2	0.439	0.729	0.952	0.266	1	0.942	0.886
E_3	0.473	0.703	0.982	0.414	0.942	1	0.938
E_4	0.556	0.717	0.963	0.406	0.886	0.938	1

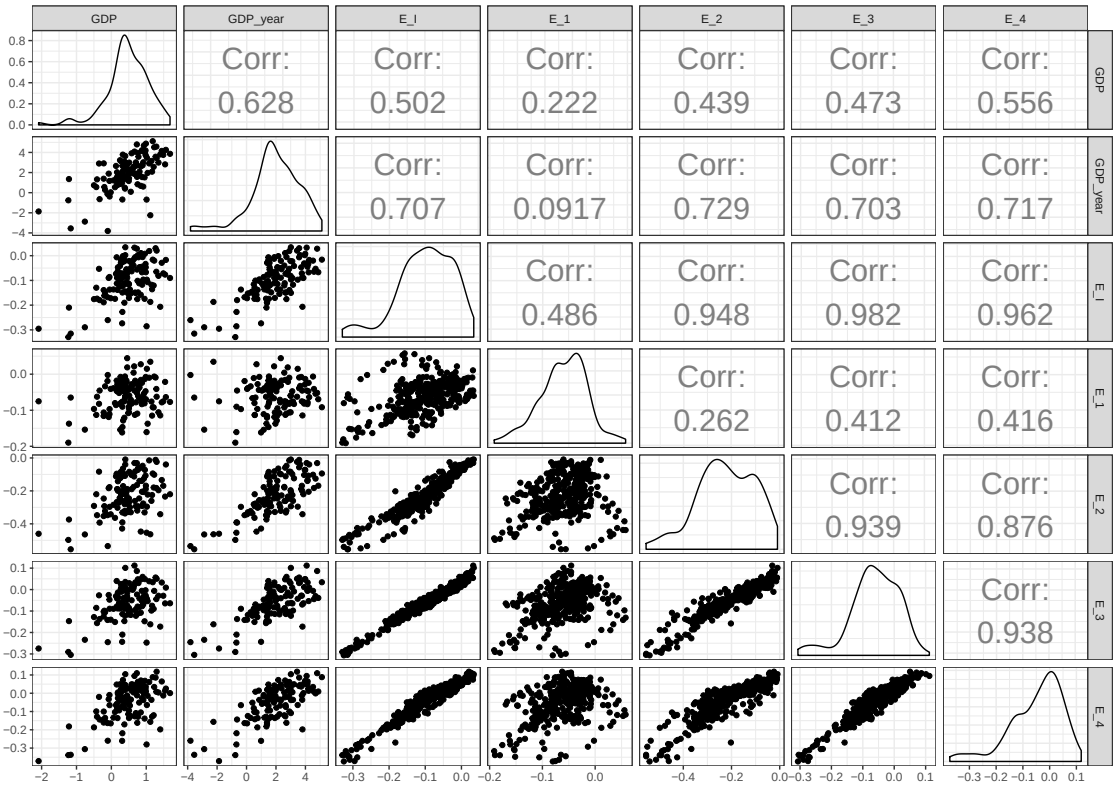


Figure 7.3: Plot

7.8 GDP vs Explanatory Indicators

Table 7.3: Correlation Matrix

	GDP	GDP_year	E_I	Var_I	Z_I	Var_Z_I
GDP	1	0.628	0.502	−0.074	0.151	0.021
GDP_year	0.628	1	0.707	−0.101	−0.058	0.011
E_I	0.502	0.707	1	−0.615	0.166	−0.484
Var_I	−0.074	−0.101	−0.615	1	−0.045	0.887
Z_I	0.151	−0.058	0.166	−0.045	1	−0.072
Var_Z_I	0.021	0.011	−0.484	0.887	−0.072	1

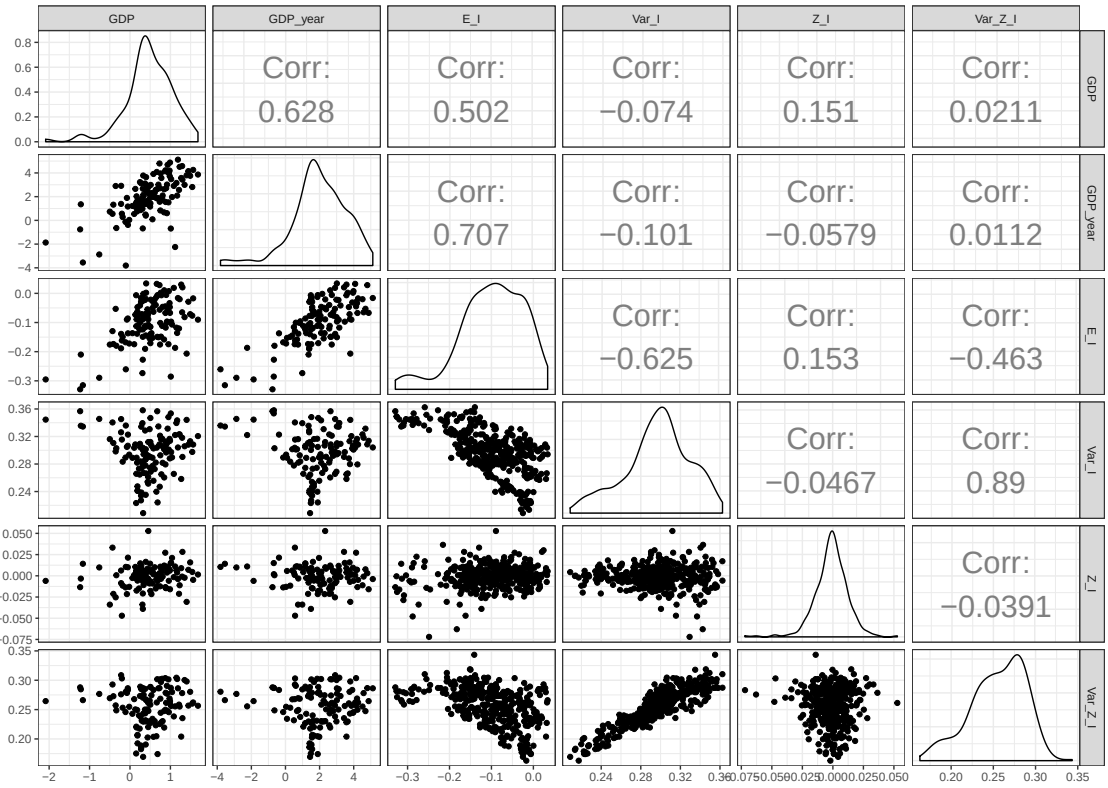


Figure 7.4: Plot

GDP vs Indicator

GDP vs Var

GDP vs Z

GDP vs Var Z

7.9 Specificity of question 3 and 4, are peoples predictions correct ?

Table 7.4: Correlation Matrix

	GDP	GDP_year	E_3	E_3_lag1	E_3_lag2	E_3_lag3	E_3_lag4
GDP	1	0.628	0.477	0.520	0.545	0.546	0.531
GDP_year	0.628	1	0.707	0.679	0.673	0.628	0.560
E_3	0.477	0.707	1	0.969	0.948	0.906	0.846
E_3_lag1	0.520	0.679	0.969	1	0.975	0.940	0.892
E_3_lag2	0.545	0.673	0.948	0.975	1	0.974	0.933
E_3_lag3	0.546	0.628	0.906	0.940	0.974	1	0.969
E_3_lag4	0.531	0.560	0.846	0.892	0.933	0.969	1

Table 7.5: Correlation Matrix

	GDP	GDP_year	E_4	E_4_lag1	E_4_lag2	E_4_lag3	E_4_lag4
GDP	1	0.628	0.558	0.555	0.591	0.566	0.536
GDP_year	0.628	1	0.719	0.650	0.647	0.593	0.501
E_4	0.558	0.719	1	0.959	0.941	0.890	0.804
E_4_lag1	0.555	0.650	0.959	1	0.970	0.928	0.863
E_4_lag2	0.591	0.647	0.941	0.970	1	0.970	0.917
E_4_lag3	0.566	0.593	0.890	0.928	0.970	1	0.959
E_4_lag4	0.536	0.501	0.804	0.863	0.917	0.959	1

CHAPTER 8

Linear (Auto-Regressive) Models

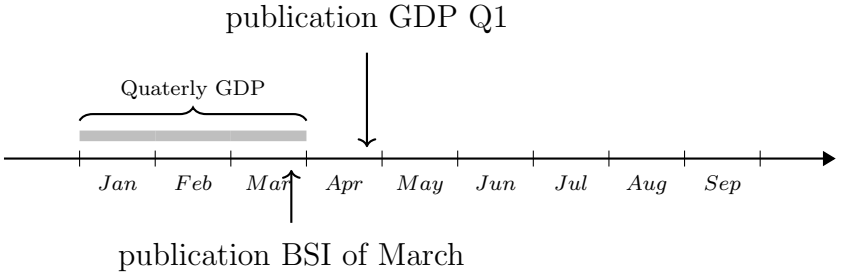
8.1 Method

8.1.1 Month vs Quarterly data

Error to aggregate everything to quarterly - lost of information

8.1.2 Timing of the Data

The quaterly GDP and the Quaterly YoY GDP is set at t. This is the common way to go..



8.2 Linear Model

$$GDP_t = \mu + \sum_{i=1}^{n} \sum_{j=0}^{q} \beta_{1,j} X_{i,t-j} + \epsilon_t \tag{8.1}$$

Auto-Regressive model

$$GDP_t = \mu + \sum_{j=1}^{p} \phi_j GDP_{t-j} + \sum_{i=1}^{n} \sum_{j=0}^{q} \beta_{1,j} X_{i,t-j} + \epsilon_t \tag{8.2}$$

where

GDP_t	GDP growth over the last semester
$X_{i,t}$	monthly predictors
μ	constant
ϕ_j	auto-regressive coefficients
$\beta_{i,j}$	regression coefficients

8.3 Model

	Linear Regression			
	Year on Year GDP (in %)			
	(1)	(2)	(3)	(4)
Constant	3.434***	1.114***	−1.829**	−4.052***
	(0.170)	(0.255)	(0.867)	(0.922)
3 months lag of YoY GDP		0.690***	0.525***	
		(0.066)	(0.078)	
BSI	15.053***	4.957***	10.621***	22.080***
	(1.364)	(1.387)	(2.079)	(1.392)
Variance			12.831***	27.506***
			(3.628)	(3.349)
Observations	124	123	123	124
R^2	0.499	0.743	0.767	0.679
Adjusted R^2	0.495	0.738	0.761	0.673
Residual Std. Error	1.172	0.847	0.809	0.943
F Statistic	121.753***	173.266***	130.755***	127.781***

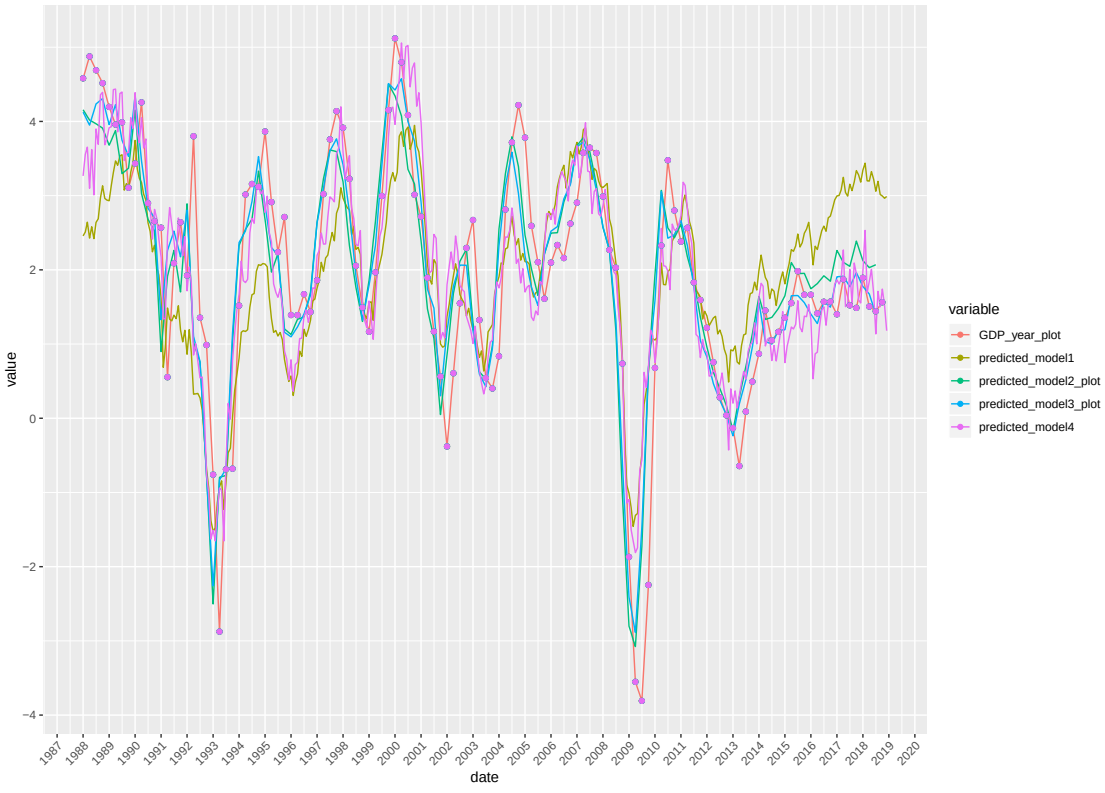


Figure 8.1: Plot

	GDP_year			
	(1)	(2)	(3)	(4)
Constant	3.364***	−4.052***	−3.993***	−0.985
	(0.173)	(0.922)	(0.917)	(0.797)
E_I	14.491***	22.080***	21.579***	5.629***
	(1.390)	(1.392)	(1.421)	(2.091)
E_I_diff	11.538*		8.181	24.063***
	(6.560)		(5.309)	(4.488)
Var_I		27.506***	27.105***	7.179**
		(3.349)	(3.340)	(3.433)
GDP_year_lag1				0.687***
				(0.076)
R^2	0.512	0.679	0.685	0.813
Adjusted R^2	0.504	0.673	0.677	0.806
Residual Std. Error	1.162	0.943	0.938	0.729
F Statistic	63.468***	127.781***	86.947***	128.113***

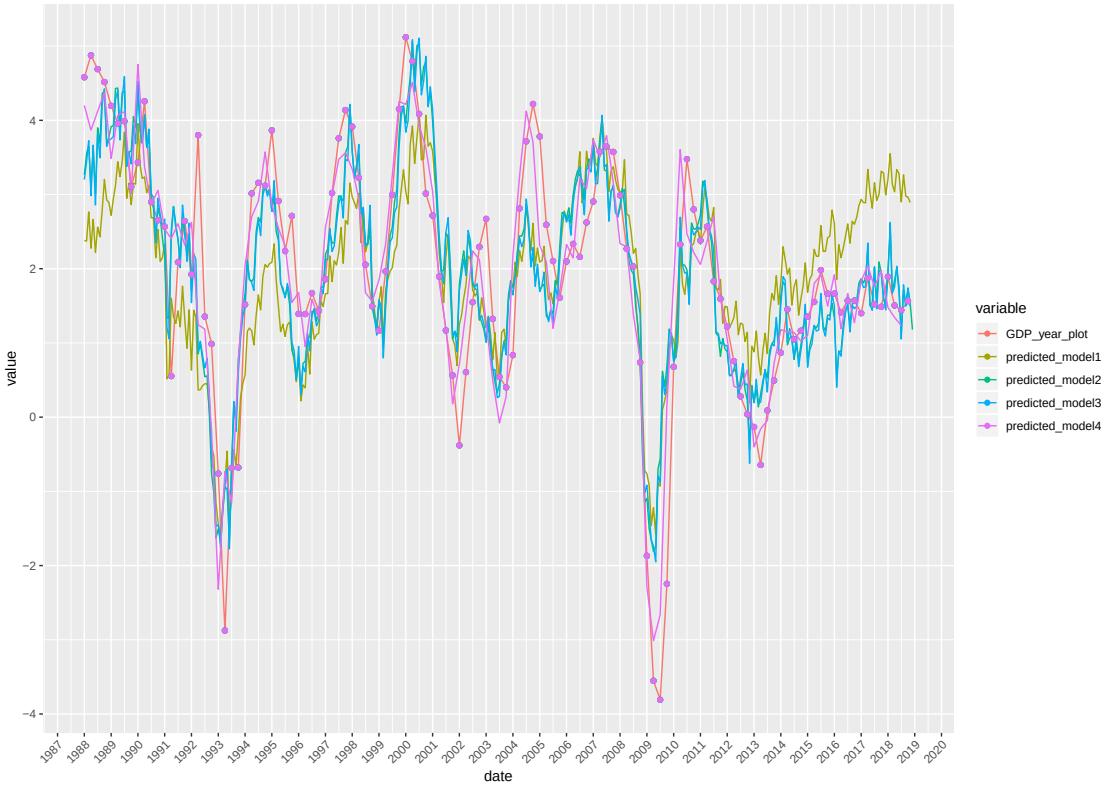


Figure 8.2: Plot

	GDP_year	
	(1)	(2)
E_I	22.783***	23.088***
	(1.340)	(1.319)
Var_I	22.151***	28.508***
	(6.050)	(3.133)
Z_I	−24.279***	−25.008***
	(5.765)	(5.746)
Var_Z_I	6.609	
	(5.383)	
Constant	−4.090***	−4.265***
	(0.871)	(0.861)
R^2	0.726	0.722
Adjusted R^2	0.717	0.716
Residual Std. Error	0.878 (df = 119)	0.880 (df = 120)
F Statistic	78.806*** (df = 4; 119)	104.133*** (df = 3; 120)

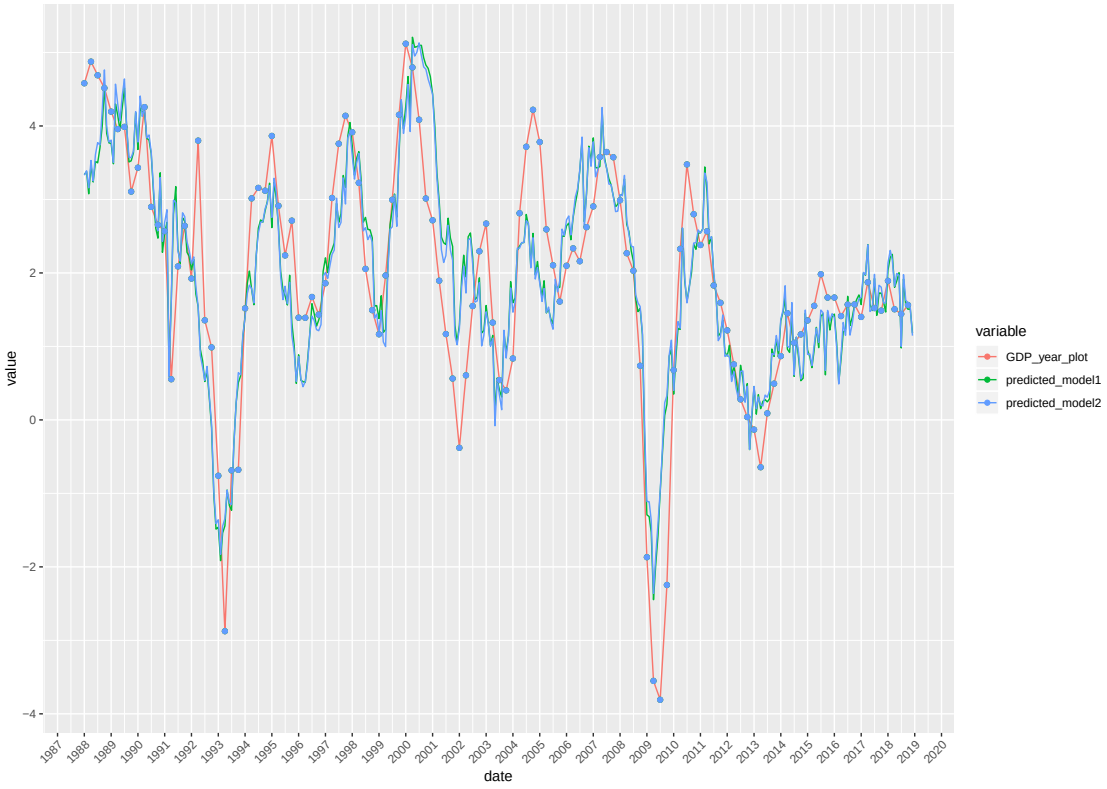


Figure 8.3: Plot

Table 8.1

	Linear Regression					
	Year on Year GDP (in %)					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.453***	3.478***	−3.862***	−4.013***	−4.033***	−3.931***
	(0.170)	(0.166)	(0.899)	(0.841)	(0.841)	(0.847)
BSI	15.186***	15.406***	22.166***	22.696***	22.930***	22.695***
	(1.365)	(1.323)	(1.387)	(1.306)	(1.324)	(1.345)
diff BSI		−17.924***		−21.927***	−13.699	−13.415
		(6.422)		(4.993)	(9.270)	(9.274)
Variance			26.936***	27.669***	27.760***	22.626***
			(3.272)	(3.072)	(3.071)	(5.985)
Z					−11.584	−11.014
					(10.999)	(11.014)
Volatility						5.489
						(5.492)
Observations	124	123	124	123	123	123
R^2	0.504	0.536	0.682	0.724	0.727	0.729
Adjusted R^2	0.500	0.528	0.677	0.717	0.717	0.717
Residual Std. Error	1.167	1.126	0.938	0.872	0.872	0.872
F Statistic	123.760***	69.317***	129.635***	104.119***	78.438***	62.950***
Note:						*p<0.1; **p<0.05; ***p<0.01

8.4 Evaluation

8.4.1 R-square

8.4.2 AIC and BIC

8.4.3 Mean Square Prediction Error

8.4.4 Diebold-Mariano Test

8.4.5 Out-of-Sample performances

ME: Mean Error
RMSE: Root Mean Squared Error
MAE: Mean Absolute Error
MPE: Mean Percentage Error
MAPE: Mean Absolute Percentage Error
MASE: Mean Absolute Scaled Error
ACF1: Autocorrelation of errors at lag 1.

8.5 log(GDP)

8.6 Variance(X) VS Variance(Z)

8.7 Take Question 1 out of the calculation of the Indicator

CHAPTER 9

Markov Switching Models

Small Introduction + why are we using it

Since the pioneer work by Hamilton (1989), Markov Switching models have been largely used to model business cycles.

Markov Switching models have been very popular since Hamilton (1989) to model business cycles and predict Turning points (see Duprey and Klaus (2017), ...).

Able to predict the 2008 financial crisis if used MS-VAR model Gadea Rivas and Perez-Quiros (2015)

9.1 Model(s) Specification

9.1.1 Notation

$S_t = \{0, 1\}$	states
$N = 2$	number of states (2)
$T = 372$	number of observations
$x_{t=1...T}$	(hidden) state at time t
$y_{t=1...T}$	Change of the Industrial production indices at time t
$p_{i=1...n,j=1...n}$	probability of transition from state i to state j
$F(y\|\theta)$	probability distribution of an observation, parametrized on θ

9.1.2 Model

Model

$$y_t = \begin{cases} \mu_0 + \sum_{j=1}^{p} \phi_j GDP_{t-j} + \epsilon_t & \text{if } S_t = 0 \\ \mu_1 + \sum_{j=1}^{p} \phi_j GDP_{t-j} + \epsilon_t & \text{if } S_t = 1 \end{cases} \tag{9.1}$$

where ϵ is $N(0, \sigma_s)$

$\mu_s = \beta_0 = c_s = \alpha_s$	regime-specific mean
$\beta_s = \phi_s$	regime-specific auto-regressive parameter
σ_s	regime-specific variance

Transition equation/probability

$$P = P(S_t = s_t \mid S_{t-1} = s_{t-1}) = \begin{bmatrix} 1-p_t & p_t \\ q_t & 1-q_t \end{bmatrix} \tag{9.2}$$

We have then,

$P(S_t = 1 \mid S_{t-1} = 1) = p_t$
$P(S_t = 0 \mid S_{t-1} = 1) = 1 - p_t$
$P(S_t = 0 \mid S_{t-1} = 0) = q_t$
$P(S_t = 1 \mid S_{t-1} = 0) = 1 - q_t$

or

$$P = P(S_t = s_t \mid S_{t-1} = s_{t-1}) = \begin{bmatrix} 1-p_t & p_t..... \\ q_t..... & 1-q_t \end{bmatrix} \tag{9.3}$$

CHAPTER 10

Conclusion

A large
    It was seen that

CHAPTER 11

Discussion

11.1 Recruitment procedure and panel data

not real sampling theory

11.2 Z that takes more periods into account

11.3 Limitations

Variance influence by drop-out, attrition, ...

11.4 Improve the Business Survey

Change participants

Mo

From a statisticians point of view, a more sampling theory Including SRS or else would be more optimal

11.5 Further Research

More complex Nowcasting model with Space space models / MIDAS ....

Combine mixed models and Markov Chain for Panel Data (de Haan-Rietdijk et al., 2017)

State Space Model

Bayesian estimation Bialowolski et al.

Bibliography

Bialowolski, P., Kuszewski, T., and Witkowski, B. Bayesian averaging of classical estimates in forecasting macroeconomic indicators with appliction of business survey data.

Commission, E. The Joint Harmonised EU Programme of Business and Consumer Surveys User Guide. page 56.

Das, M., Toepoel, V., and van Soest, A. (2011). Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys. *Sociological Methods & Research*, 40(1):32–56.

De Greef, I. and Van Nieuwenhuyze, C. (2009). The National Bank of Belgium's new business survey indicator.

de Haan-Rietdijk, S., Kuppens, P., Bergeman, C. S., Sheeber, L. B., Allen, N. B., and Hamaker, E. L. (2017). On the Use of Mixed Markov Models for Intensive Longitudinal Data. *Multivariate Behavioral Research*, 52(6):747–767.

Duprey, T. and Klaus, B. (2017). How to predict financial stress? An assessment of Markov switching models. Technical report.

Gadea Rivas, M. D. and Perez-Quiros, G. (2015). The Failure to Predict the Great Recession - A view through the role of credit. *Journal of the European Economic Association*, 13(3):534–559.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57(2):357.

Hlavac, M. (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.

Priyana Hardjawidjaksana, I. (2019). *Investigating Attrition in Belgian Labor Force Survey*. PhD thesis, KU Leuven.

Rhoads, C. (1999). Euroland Discovers A Surprise Indicator: Belgian Confidence. *Wall Street Journal, Europe; Brussels*, page 1.

Vanhaelen, J.-J., Dresse, L., and De Mulder, J. (2000). The Belgian Industrial Confidence Indicator: Leading Indicator of Economic Activity in the Euro Area? *SSRN Electronic Journal*.

List of Abbreviations

BSI Business Survey Indicator
GDP

List of Figures

List of Tables

Appendix

Further Explanation of the Evolution of the responses ...

Notation	x_{t-1}	x_t	z_t
π_{--}	-1	-1	0
π_{-0}	-1	0	1
π_{-+}	-1	1	2
π_{0-}	0	-1	-1
π_{00}	0	0	0
π_{0+}	0	1	1
π_{+-}	1	-1	-2
π_{+0}	1	0	-1
π_{++}	1	1	0

Further Explanation of the Markov Switching model ...

CONJUNCTUURONDERZOEKINGEN - NIJVERHEID

TEL. + 32 (0)2 221 49 97 E-mail: nationalbanksurveys@nbb.be

Kruis per vraag, het vakje aan dat overeenstemt met uw antwoord en stuur per product uw antwoorden binnen de 10 dagen terug. U kunt gebruik maken van ons gratis faxnummer 0800 95 969 (enkel geldig binnen België) of van het faxnummer + 32 (0)2 221 31 07 (vanuit het buitenland).

Product: **OKTOBER 2018**

Gelieve enkel voor het hierboven vermelde product te antwoorden. Vermeld alle schommelingen, zelfs indien ze van geringe omvang zijn. Antwoord elke maand op alle vragen. Indien u in de beschouwde maand het product niet heeft geproduceerd (of geen bestellingen heeft ontvangen), vermeldt u "verminderd". Antwoord "onveranderd" gedurende de maand(en) dat deze toestand voortduurt totdat de productie herneemt. Uw antwoorden worden strikt vertrouwelijk behandeld.

Verloop en beoordeling

- Uw productietempo voor dit product is in september 2018 t.o.v. augustus 2018:

  15 ☐ 1 versneld ☐ 5 onveranderd ☐ 9 vertraagd

  Hou geen rekening met schommelingen als gevolg van het maandelijks veranderlijke aantal werkdagen of betaald verlof.

  Tijdens de afgelopen 3 maanden was de trend van uw productie voor dit product:

  16 ☐ 1 stijgend ☐ 5 onveranderd ☐ 9 dalend

  Hou evenwel geen rekening met louter seizoengebonden schommelingen.

- Uw verkoopprijzen voor dit product zijn in september 2018 t.o.v. augustus 2018:

  17 ☐ 1 gestegen ☐ 5 onveranderd ☐ 9 gedaald

  Geef de tendens van uw prijzen aan op basis van uw contracten of uw aanbiedingen.

- Uw huidige voorraad van dit product beschouwt u, voor het seizoen, als:

  18 ☐ 1 hoger dan normaal (te hoog) ☐ 5 normaal (voldoende) ☐ 9 lager dan normaal (te laag)

  Kruis "niet van toepassing" aan indien u nooit voorraad hebt van dit product. ☐ niet van toepassing

  Bij het beantwoorden van de volgende twee vragen (22 en 26), mag u geen melding maken van de zuivere seizoenschommelingen die het verloop van de bestellingen gedurende de maand september 2018 kunnen hebben beïnvloed. De werkelijke tendens van de bestellingen moet dus tot uiting komen.

- Uw ontvangen bestellingen voor dit product vanwege de binnenlandse markt zijn in september 2018 t.o.v. augustus 2018:

  22 ☐ 1 vermeerderd ☐ 5 onveranderd ☐ 9 verminderd

  Hou eveneens rekening met de van andere afdelingen van uw firma ontvangen bestellingen en met loonwerk.

  Kruis "niet van toepassing" aan indien u dit product nooit op de binnenlandse markt levert. ☐ niet van toepassing

- Uw ontvangen bestellingen voor dit product vanwege de buitenlandse markt zijn in september 2018 t.o.v. augustus 2018:

  26 ☐ 1 vermeerderd ☐ 5 onveranderd ☐ 9 verminderd

  Hou eveneens rekening met loonwerk.

  Kruis "niet van toepassing" aan indien u dit product nooit op de buitenlandse markt levert. ☐ niet van toepassing

- Uw huidige gezamenlijke orderpositie voor dit product beschouwt u als:

  27 ☐ 1 hoger dan normaal ☐ 5 normaal ☐ 9 lager dan normaal

  Indien uw onderneming uitsluitend uit voorraad levert, dient u "orderpositie" op te vatten als "het peil van de vraag" naar dit product.

- Indien u het huidige fabricatietempo voor dit product handhaaft, is uw activiteit nog verzekerd voor ongeveer:

  28 ☐☐ , ☐ maand(en) en/of gedeelten van een maand.

  Te ramen op basis van uw orderpositie of, bij gebrek hieraan, op basis van uw productieplannen.

- Uw huidige positie inzake bestellingen uit het buitenland voor dit product beschouwt u als:

  31 ☐ 1 hoger dan normaal ☐ 5 normaal ☐ 9 lager dan normaal

  Indien uw onderneming uitsluitend uit voorraad levert, dient u uw buitenlandse orderpositie op te vatten als "het peil van de buitenlandse vraag" naar dit product.

  Kruis "niet van toepassing" aan indien u dit product nooit uitvoert. ☐ niet van toepassing

Vooruitzichten voor de volgende drie maanden

- Het personeel (arbeiders en technici) tewerkgesteld voor de fabricatie van dit product zal volgens u:

  32 ☐ 1 worden uitgebreid ☐ 5 onveranderd blijven ☐ 9 worden verminderd

  Het invoeren van gedeeltelijke werkloosheid dient als een vermindering van het personeel te worden beschouwd.

- De vraag van uw klanten naar dit product zal volgens u:

  33 ☐ 1 belangrijker ☐ 5 even belangrijk ☐ 9 minder belangrijk

  zijn dan gewoonlijk tijdens die periode van het jaar.

  Geef enkel de tendens van de vraag van de klanten weer en laat derhalve de zuivere seizoenschommelingen buiten beschouwing.

- Uw productie zal voor dit product volgens u:

  36 ☐ 1 toenemen ☐ 5 gelijk blijven ☐ 9 afnemen

- Uw verkoopprijzen van dit product zullen volgens u:

  34 ☐ 1 stijgen ☐ 5 onveranderd blijven ☐ 9 dalen

4100N

REFERENTIE:
Enquête:

Figure 1: The Business Survey Questionnaire in Dutch for the Industrial Sector

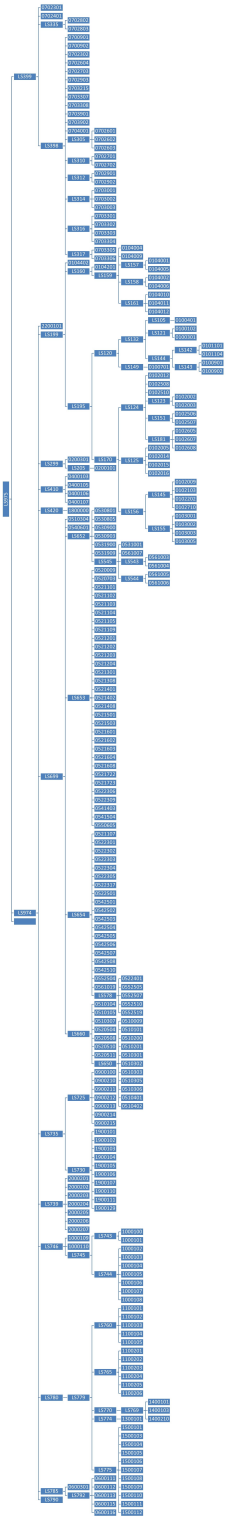


Figure 2

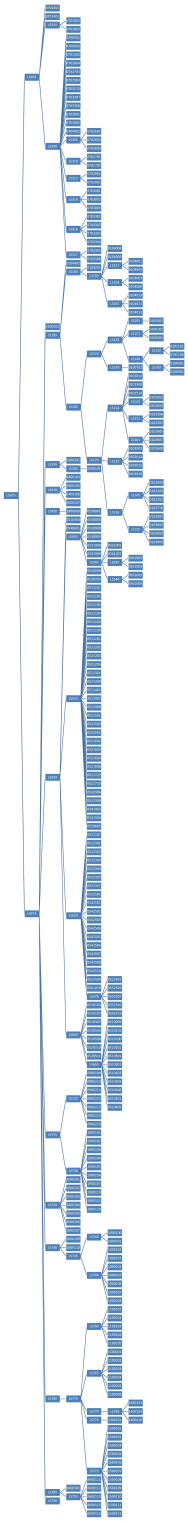


Figure 3

Code

R code for Seasonal Adjustment

R code for Creating Lags

R code for Linear (Auto-Regressive) Models

R code for Markov Switching Models

AFDELING
Straat nr bus 0000
3000 LEUVEN, BELGIË
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
www.kuleuven.be