The Variability of the Belgian Business Survey Indicator

Analysis and 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

i

Acknowledgement

I'm grateful to my co-supervisor, Laurent Van Belle, for his guidance and his precious help.

I'm thankful to ..... that where always available to answer my questions and whose imputs during meeting where crucial to give the best direction to my work.

the National Bank of Belgium

Jean Palate - David ... and all the members of the Department of Research and Development who where always there to answer any of my questions.

Isabelle - Marc Boumon

Vanessa Baugnet - Rudi Acx

for their time and advise during the writing of master thesis.

I'm grateful to my co-promotor Geert Molenberghs for the guidance and feedback. Geert Loosveldt and Stephan Moens for their input during the Mid-term presentation.

Last but not least, I would like to thank all the professors of the Master of Statistics that made it possible for me to acquire the knowledge and passion of statistics, needed to write this thesis.

"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

Objectives

Methodology

Plan of this Paper

Chapter 2

   3

   4

   5

CHAPTER 2

The Business Survey Indicator

This first chapter is a presentation of the Belgian business survey and the business survey indicator, also referred to as the business survey barometer or business confidence barometer/indicator. We will first go thought the evolution of the barometer since it started. Then present the recruitment method, the objectives and methodology of they survey. In the last part of the chapter, we will present the calculation method of the business survey indicator (BSI).

2.1   History

The Belgian Business Survey celebrates this year it's 65$^{\text{th}}$ anniversary, since the Survey was launched by the National Bank of Belgium in 1954. Belgium was part of the pioneers since only the United States (1930) and West Germany (1949) had a business survey.

In 1972, the results where first synthesised in an indicator. The Business Survey Barometer started by including only the industrial sector. It was then small by small enlarged to other sectors: construction, trade and services.

At the European level, it was in 1961 that the European Commission launched an harmonisation program of the Business Survey in the manufacturing industry. Since then, the sector coverage of the programme has widened considerably. The Business Survey programme was extended to the construction sector and the 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. The harmonisation only concerns the type of questions, the possible answers and the method to calculate the indicator. ... More information ca be found in the "The Joint Harmonised EU Programme of Business and Consumer Surveys User Guide" European Commission (2016).

Over the time, the business survey barometer has been well known for been a very informative and useful indicator. It was noticed in a paper by the Wall Street Journal titled "Euroland Discovers A Surprise Indicator: Belgian Confidence" (Rhoads, 1999) that acknowledged its importance and accuracy in understanding the evolution of the Belgian economy.

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

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.

Several improvement of the business survey where proposed and applied over time (1983, 1990 and 2009). The last ?improvements will be discussed in detail in section 2.4.

2.2   Sampling Method

The Belgian business survey barometer (as most of the business survey around the world), has the particularity of not using random sampling. The selection of participants is quite complex and a lot of decision are human, never is a statistical program or a random sampling system used to select new participants.

Procedure of selection

The selections of new participants is done by "waves". When the department responsible for the business survey at the NBB decides that there aren't enough participants in a specific sector anymore, the recruitment of new participants is launched.

To find the new participants, the first step is to decided for an optimal amount of new participants needed, regarding the different stratification of the sector. Each sector is composed of a quite advanced trees of sub-sectors, sub-sub-sectors, ... for example the industry sector is divided into more than 300 sub-sectors / branches over 6 different levels. A representation of the classification ca be seen in Figure 3.

As will be seen in section 2.4.2 and section 2.4.3 more in details, companies are weighted by there size (profit, number of employees, ...) and the size of the sector/branch they are part of. Based on those information, it's looked at which branches, size of companies, new respondents should be looked for. This makes the procedure quite complex and therefore contains a lot of human decisions.

Out of this process comes a list of potential new participants. This list is then send to the communication department that takes contact with those potential new participants. Not always, but it happens quite often, a representative of the National Bank visits the new participant to explain the survey and have a contact. As a reward for participating at the survey, the companies receive privileged information. Each month they receive access to sub-sector indicators information that aren't publicly distributed. This can give them economical information regarding there specific economical activity.

At the National Bank of Belgium, this procedure is usually referred to as "prospecting" rather than selecting or sampling. Since it is mostly based on recruiting new companies that will work with them, by looking at the Belgian economy and find participants willing to collaborate. For example, it's possible that the responsible person for the Horeca business survey, goes to restaurant et get acquainted with the owner, sees the interest of including him in the business survey and recruits him this way. This is a less common way, the main of the recruitment happens to the process explained above.

Leaving the Survey

An important side of business survey is that people are staying as long as possible in the survey. Ones companies are part of the survey, they stay in it until they decide to leave. There are no participants removed from the survey by the national Bank, what can happen is that if participants don't answer for three months, contact will be taken with the company to see if they want to continue to participate.

There is a real work done by the National Bank to make sure companies answer to the survey, and stay in the survey. This means that some companies are part of the survey for a very long time. To have an idea, if we look at the Survey for the Industry, trace respondents back to 1988, we can see that today, approximately one third of the respondents where already in the survey in 1988.

Why not a problem that Random Sampling isn't used

From a statistical point of view, it can be rather problematic to draw general conclusions over a populations when Random Sampling is not use. Without undermining one of the most important pillars of statistics, there are two main reasons why in this case, having a non-random selection of participants is not too problematic.

The first reason is that the method used tries to represent as good as possible the population that it wants to represent. Therefore stratification is used at a quite advanced level as explained before. We could call this recruitment method: non-random stratification, as opposed to random stratification. Since it's not using sampling but takes into account the stratification of the population it's studying.

The second reason ,the most import one, is that the business survey Indicator has no interest on its own. Indeed having a BSI equal to 0.5 or 0.1 don't mean anything, what's important is the evolution of the indicator. If it was equal to 0.3 last month and this month it's equal to 0.5, it means that the economy is most probably growing and that we can forecast an increase of GDP over the month. On the other hand if it's now equal to 0.1, it means a decrease in economical confidence among businesses and we can anticipate a deceleration/decrease of GDP over that period.

2.3 Objectives of the Business Survey

predictive power for the future months - high correlation with Economical Indicators

The main objective of the business survey Barometer is to have a feeling of how the economy is now and how it will evolve in the short term. Therefore we will first speak about nowcasting, its explanation and its implications.

average weekly work hours in manufacturing, factory orders for goods, housing permits and stock prices index of consumer expectations, average weekly claims for unemployment insurance and the interest rate

fluctuation that's mostly explained by the conjuncture

The second objective of the Business Confidence Barometer, is to make prevision as less as possible influenced by short terms variation, but the capture the more long run evolution

of the economy, more precisely, business cycles. Therefor we will here discuss business cycles.

In chapter 6, we will discuss most of the small, short-term variation that can influence the BSI and that aren't relevant. Those can be due to the responses to the survey (non-response, dropout and Attrition) but also to "real" effects (seasonal effects) that shouldn't be captured by the BSI.

2.3.1 Belgian economy and sector specific short term evolution

The business survey indicators are published at the end of each month (around the 21-25 of the month) and give a fast capture of the evolution of the Belgian economy over the past months. The data is available on stat.nbb.be and a press release is published on nbb.be. It contains a short summary and interpretation of the BSI followed by graphs to show the evolution of the BSI over the two last years for the sector specific and overall indicator. The public can use those information to have a snapshot of the economy, while other indicators, as GDP or unemployment, can take a very long time before been obtained and published.

2.3.2 Nowcasting

Also called "flash" estimation, nowcasting has increasingly gain importance in the last decade (see Foroni and Marcellino (2014), de Antonio Liedo (2014), ...). it consists in the short term estimation of the economy, mostly the GDP. It is a fundamental approach since the business survey indicator is published monthly while other indicator's like GDP are published quarterly.

As can be seen in Figure 2.1, the lag between the observation and publication is even greater since the business survey Indicator is published at the end of each month (around the 24-25$^{\text{th}}$ of each month, will the GDP is published with a lag of 3-4 weeks ????
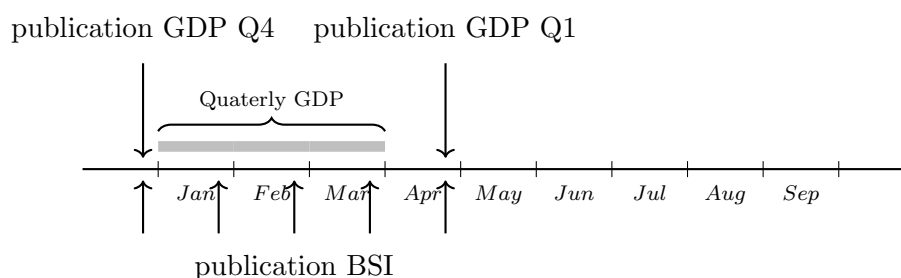
Figure 2.1: Timeline of the publication of the business survey Indicator (BSI) and the Gross Domestic Product (GDP)

Nowcasting got by the time a catch all word and include predictive models going from Linear models to space-space models and also including ARIMA or MIDAS and much more different models.

2.3.3 Business Cycles

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.

Already in 1946, Mitchell and Burns published the book "*Measuring Business Cycles*". They Business cycles where described as recursive fluctuations, affecting macroeconomics variables.

Business cycles are putting aside other effects like seasonal variation and other small variations. It is taking into account a larger trend in the economy.

What's GDP and Year on Year GDP

As seen, business cycles are defined by several macroeconomic variables. Gross domestic income (GDI), unemployment, ...

But the main indicator, is the GDP

Later on we will also use YoY GDP, for x reasons

$$\text{YoY GDP} = \frac{\text{GDP}_t - \text{GDP}_{t-12}}{\text{GDP}_{t-12}} \tag{2.1}$$

Figure 2.2 shows a simplified version of the Business Cycles theory.
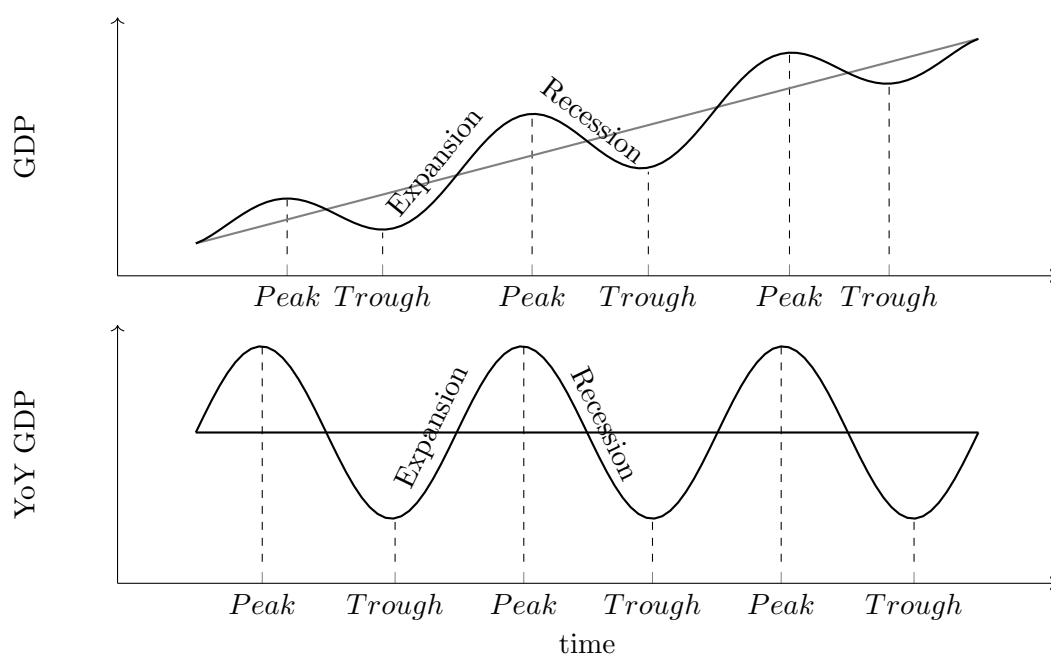
Figure 2.2: The Business Cycle theory of GDP and Year on Year GDP

Duration of a Business Cycle

Figure 2.2 can give the false impression that business cycles are of the same length, this is, in real life, not so simple.

Mitchell and Burns (1946) stated a minimum duration of 16-22 months and a maximum duration of 100-106 months. This is a quite large range and since then several researches worked on this topic. There is no clear

As a lot of research argues, business cycles are mainly empirically defined. If we look for example at the United States, a big work was done by the National Bureau of Economic Research (NBER) to date turning points for the American economy. The empirical evidence that comes out of this work, is that the time from one economic peak to the next is on average 5 and a halve years (for the period 1945-2009)

We can see from Figure 2.3, which represent the different though and peak of business cycles identified by the NBER from 1975 to 2009, that their is no symmetry of the Business Cycle. Some business cycles are very short while other last more than ten years. We can notice that periods of economical growth, usually last longer than economical decrease. ....
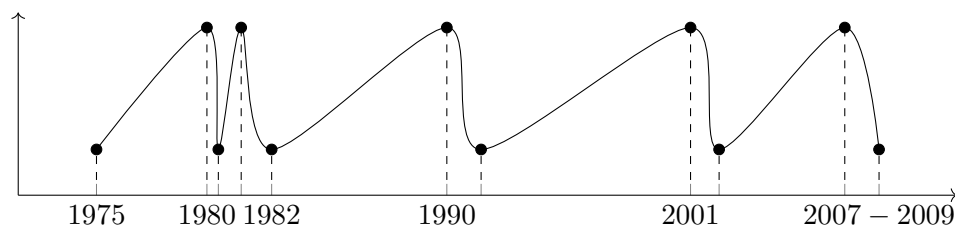
Figure 2.3: Business cycles from 1975 to 2009 of the American Economy according to the NBER

2.4 Methodology of the Business Survey

1954 begin 1972 results are synthesised in an indicator 1983 & 1990 methodological changes

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

- 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

2.4.1 Questionnaire

The questionnaire exists in two languages, french and dutch. And can be answered by mail, email, over the phone or by fax.

The questionnaire has XX questions that can be grouped in two type of questions. (1) questions concerning there current production and level of activity. (2) questions concerning there prediction, how they expect there level of activity to evolve over the next three months. It's therefore that the business survey indicator, is also called the confidence indicator, since its a measure of how confidence companies are in the Belgian economy.

Changes over time

In this paper, the answers and results of the business survey barometer will be used since 1988. Therefore it's very important to see if the answers changed over time.

A questionnaire from 1990 was found (See Appendix page 50) and can be compared to a more recent version (See Appendix page 49).

It can be seen that the layout was modified, and that the questions phrasing changed over this long period. It was before asked in the first person while it's now phrased in the third person. Aside from those small changes, the survey kept the same questions

It would be interesting to have a closer look in the potential consequences of those changes over time. The layout, the phrasing and the method of answering can potentially have an influence on the answers. Nevertheless, since it's not the subject of this paper, we will leave this study to future research. What we can say is that the influence of those changes are limited since the respondents where mostly the same when the changes happened, so the interpretation they made from one version to the other one should have stayed quite similar.

Questions taken into account for the NS975

The first question has to be interpreted the opposite way as of the three other questions. ...

see appendix page 49 for the full questionnaire

2.4.2   Weighting procedure

Each company that is part of the panel of respondents, has a weight according to it's size, the profit the company is making, the capital it's owning and other characteristics. The calculation is quite complex since it's specific to each sector. For example in the industry the ...

cite working papers

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

Adapted over time, with a smoothing effect over time

2.4.3   Globalisation procedure

Based on the size of the sector in which the company is.

The National Bank of Belgium developed a quite elaborate division of the Belgian economical activity. This means that for example, the Industry is subdivided into different sub-sectors, that them self contain sub-sectors that contain sub-sectors and so on six times. All those divisions have a percentage according to the size of those subdivision inside the division.

2.5 Calculation of the Indicator

This section presents the method for calculating the business survey 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. We first present the calculation taking into account one question unweighted and then weighted. After we will present how the indicator of different questions are combined together.

2.5.1 Unweighted Indicator

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

$$E(X) = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2.2}$$

where x_i is the answer of the respondent i and can take value -1 (negative answer), 0 (neutral answer) and 1 (positive answer). n is the number of respondents.

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

$$E(X) = \frac{\sum_{i=1}^{n} x_{+i} + \sum_{i=1}^{n} x_{0i} + \sum_{i=1}^{n} x_{-i}}{n} \tag{2.3}$$

$$E(X) = \frac{\sum_{i=1}^{n_+} x_{+i} + \sum_{i=1}^{n_0} x_{0i} + \sum_{i=1}^{n_-} x_{-i}}{n} \tag{2.4}$$

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

Since we know that $\sum_{i=1}^{n} x_{0i} = 0$, we can write

$$E(X) = \frac{\sum_{i=1}^{n} x_{+i}}{n} + \frac{\sum_{i=1}^{n} x_{-i}}{n} \tag{2.5}$$

$\sum_{i=1}^{n} x_{+i}/n$ is the proportion of positive answers and $\sum_{i=1}^{n} x_{-i}/n$ is the negative proportion of negative answer. We can write, for simplicity

$$E(X) = \pi_+ - \pi_- \tag{2.6}$$

where π_+ and π_- are the proportion of respondents answering positive and negative to the specific question. π was chosen as symbol here, since it can be interpreted as a probability: if we assume that all the respondents have the same probability giving a certain answer, π is the probability that a respondent answers positive, negative or neutral to the question.

2.5.2 Weighted Indicator

As described in section 2.4.2 and section 2.4.3, each respondent has two different weights: one according to its size, one according to the size of the sector it's part of. Those weight

are then combined and we end up with a specific weight. We have now the following equation for the indicator

$$E(X) = \sum_{i=1}^{n} (\omega_i x_i) \quad \text{where} \quad \sum_{i=1}^{n} \omega_i = 1 \tag{2.7}$$

where x_i is the answer of the respondent i and can take values -1, 0 and 1. ω_i is the weight of respondent i. The weights are here standardised so there sum is equal to one.

As for the unweighted indicator, we can decompose the equation by the three possible answers with, in this case, their according weights.

$$E(X) = \sum_{i=1}^{n} \omega_{+i} x_{+i} + \sum_{i=1}^{n} \omega_{0i} x_{0i} + \sum_{i=1}^{n} \omega_{-i} x_{-i} \tag{2.8}$$

and again we know that $\sum_{i=1}^{n} \omega_{0i} x_{0i} = 0$, so we can write

$$E(X) = \sum_{i=1}^{n} \omega_{+i} x_{+i} + \sum_{i=1}^{n} \omega_{-i} x_{-i} \tag{2.9}$$

We also know that $x_{+i} = 1$ and $x_{-i} = -1$

$$E(X) = \sum_{i=1}^{n} \omega_{+i} - \sum_{i=1}^{n} \omega_{-i} \tag{2.10}$$

That will be written as follow

$$E(X) = \Omega\pi_+ - \Omega\pi_- \tag{2.11}$$

where $\Omega\pi_+$ and $\Omega\pi_-$ are the weighted proportion of respondents answering positive and negative. this equation will be used for the same reasons as for the unweighted indicator. Same explanation also works here

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 $\pi_+ + \pi_0 + \pi_- = 1$.

From Equation 2.6 and Equation 2.11 it can be seen that the weighted and unweighted indicators are bounded between -1 and 1. In the two cases, the indicator is the smallest if everyone has a negative answer, and is the largest when every answer is positive.

2.5.3 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:

$$\text{Industry business indicator} = \frac{E(X_{Q1}) + E(X_{Q2}) + E(X_{Q3}) + E(X_{Q4})}{4} \tag{2.12}$$

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

$$\text{Industry business survey indicator } = \frac{\sum_{i=1}^{n} \left(x_{iQ1} + x_{iQ2} + x_{iQ3} + x_{iQ4} \right)}{4n} \tag{2.13}$$

$$\text{Industry BSI } = \frac{\pi_{Q1+} + \pi_{Q2+} + \pi_{Q3+} + \pi_{Q4+} - \pi_{Q1-} - \pi_{Q2-} - \pi_{Q3-} - \pi_{Q4-}}{4} \tag{2.14}$$

$$\text{Weighted Industry BSI } = 1/4(\pi_{Q1+} + \Omega\pi_{Q2+} + \Omega\pi_{Q3+} + \Omega\pi_{Q4+} \tag{2.15}$$
$$-\Omega\pi_{Q1-} - \Omega\pi_{Q2-} - \Omega\pi_{Q3-} - \Omega\pi_{Q4-}) \tag{2.16}$$

Generalisation !

The Variance of the Indicator

The variance is, with the mean, one of the first tool for Statisticians to study a certain variable. Next to the mean, that is the average value of a certain variable, the variance is the measure of the dispersion. In the context of the business survey, the variance haven't been used much, while it can be seen as "how much people agree", an important information that we can extract from the survey.

The variance will be interpreted as to what level respondents agree about the state of the Belgian economy.

This chapter will present the calculation of the variance of the unweighted and weighted indicator. It will then look into the properties and specificities of the variance of the business survey indicator.

3.1 Variance of the Unweighted Indicator

The formula of the variance can be written as

$$\mathrm{Var}(X) \;=\; E\left[(X - E(X))^2\right] = E\left(X^2\right) - E\left(X\right)^2 \tag{3.1}$$

In the case of one question of the business survey indicator, we can write and develop the equation as follow

$$\mathrm{Var}(X) \;=\; E\left(X^2\right) - E\left(X\right)^2$$

$$= \left(\frac{\sum_{i=1}^{n} x_{-i}^2}{n}\right) + \left(\frac{\sum_{i=1}^{n} x_{0i}^2}{n}\right) + \left(\frac{\sum_{i=1}^{n} x_{+i}^2}{n}\right) - E(X)^2 \tag{3.2}$$

The positive and negative answers take value -1 and 1, which means that $x_{+i}^2 = x_{+i}$, $x_{-i}^2 = |x_{-i}|$. We also know that $\left(\frac{\sum_{i=1}^{n} x_{Ni}^2}{n}\right) = 0$. With this information, we can further simplify the equation;

$$\text{Var}(X) \;\; = \;\; \pi_+ + \pi_- - E(X)^2 \tag{3.3}$$

In other words, the variance of the BSI is equal to the sum of the proportion of positive and negative answers, minus the squared indicator.

We can also replace E(X) by $\pi_+ - \pi_-$, and/or $\pi_+ + \pi_-$ by $1 - \pi_0$ (since $\pi_+\pi_0 + \pi_- = 1$). Which means that we have several different ways to write the previous equation;

$$
\begin{aligned}
\text{Var}(X) \;\; &= \;\; \pi_+ + \pi_- - E(X)^2 \\
&= \;\; \pi_+ + \pi_- - (\pi_+ - \pi_-)^2 \tag{3.4} \\
&= \;\; 1 - \pi_0 - E(X)^2 \tag{3.5}
\end{aligned}
$$

3.2 Variance of the Weighted Indicator

We can now do the same for the weighted indicator. The equation is very similar as for the variance of the unweighted variance;

$$
\begin{aligned}
\text{Var}(X) \;\; &= \;\; E\left(X^2\right) - E\left(X\right)^2 \\[2mm]
&= \;\; \sum_{i=1}^{n} \omega_i x_{+i}^2 + \sum_{i=1}^{n} \omega_i x_{0i}^2 + \sum_{i=1}^{n} \omega_i x_{-i}^2 - E(X)^2
\end{aligned}
\tag{3.6}
$$

We can here, as we did for the indicator, take into account weighted proportion and write the equation the same way as for the unweighted variance but in this case, by taking the weighted proportions. Again the variance can be written in different ways, depending on the interpretation that fits best;

$$
\begin{aligned}
\text{Var}(X) \;\; &= \;\; \Omega\pi_+ + \Omega\pi_- - (\Omega\pi_+ - \Omega\pi_-)^2 \tag{3.7} \\
&= \;\; \Omega\pi_+ + \Omega\pi_- - E(X)^2 \tag{3.8} \\
&= \;\; 1 - \Omega\pi_0 - E(X)^2 \tag{3.9}
\end{aligned}
$$

3.3 Properties

Based on the previous development of the equation of the variance of the indicator, we can make some observations.

First of all, the variance is bounded between 0 and 1. A variance can't be negative since it's a sum of squares, so the lower bound shouldn't surprise anyone. On the other hand, the upper bound is more unusual. The most interesting approach is to take Equation 3.5 and see that π_0 and E(X)^2 can only take positive values. Since both the variables have
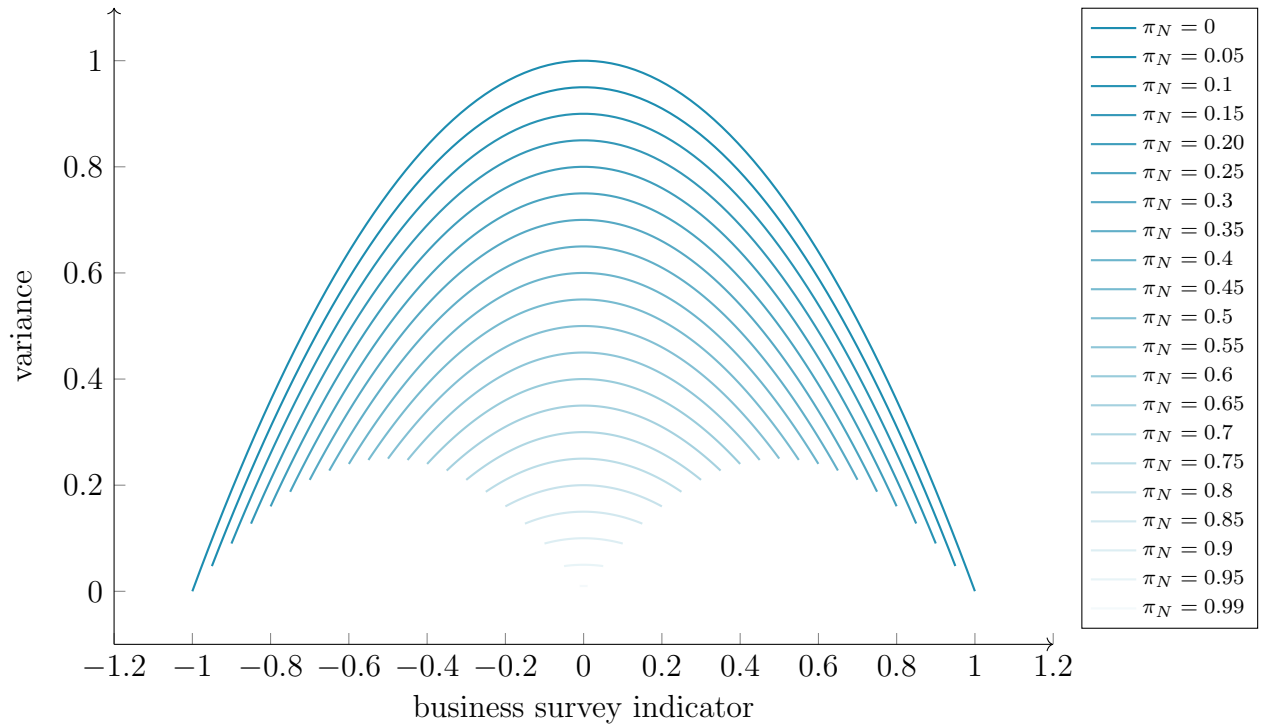
Figure 3.1: Plot of the possible values of the business survey indicator (X axis) and variance (Y axis) for different values of π_N

a minus sign in front of them, to have the highest result they should both be equal to zero. This means that the highest variance, $\mathrm{Var}(X) = 1$, is obtained when no respondent answer "neutral" and the BSI is equal to 0, in other words, half of the respondents answer "positive" while the other half answer "negative". On the other hand, the variance will be equal to zero if all the participant answer the same, that's "negative", "neutral" or "positive".

Another approach to better understand the variance of the business survey barometer, is to plot the different possible values of $\mathrm{Var}(X)$, π_0 and $\mathrm{E}(X)$ from Equation 3.5. The results can be seen in Figure 3.1.

Their are different observations that can be done from the plot; (1) there is a specific upper and lower bound for each BSI

upper bound :

$$\mathrm{Var}(X) = 1 - 0 - \mathrm{E}(X)^2 \tag{3.10}$$

lower bound :

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

$$
\begin{aligned}
\mathrm{Var}(X) &= -X - X^2 \quad \text{where } \mathrm{Var}(X) \text{ is between 0 and 1} & (3.13) \\
&= -(\mathrm{E}(X) - 0,5)^2 + 0,25 & (3.14)
\end{aligned}
$$

14

3.4 Take different questions into account

As already seen, the published indicator take different questions into account. The combination of the variance of different questions is slightly more complex than the combination of different indicators since the questions are correlated, which means that covariance has to be taken into account.

The formula to combine different variances is the following

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\text{Cov}\left(X_i, X_j\right) = \sum_{i=1}^{n}\text{Var}\left(X_i\right) + 2\sum_{1\leq i<j\leq n}\text{Cov}\left(X_i, X_j\right) \quad (3.15)$$

In the case of combining the variances of four different questions of the business survey, we have the following equation

$$
\begin{aligned}
\text{Var}\left(\frac{X_{Q1}+X_{Q2}+X_{Q3}+X_{Q4}}{4}\right) &= \frac{1}{16}[\text{Var}(X_{Q1})+\text{Var}(X_{Q2})+\text{Var}(X_{Q3})+\text{Var}(X_{Q4})\\
&\quad +2\,\text{Cov}(X_{Q1},X_{Q2})+2\,\text{Cov}(X_{Q1},X_{Q3})+2\,\text{Cov}(X_{Q1},X_{Q4})\\
&\quad +2\,\text{Cov}(X_{Q2},X_{Q3})+2\,\text{Cov}(X_{Q2},X_{Q4})+2\,\text{Cov}(X_{Q3},X_{Q4})]
\end{aligned}
$$

The complexity of the formula encourage to rather calculate the indicator (taking all the questions into account), and then calculate the variance of that indicator, we can write it as follow

$$
\begin{aligned}
\text{Var}\left(\frac{X_{Q1}+X_{Q2}+X_{Q3}+X_{Q4}}{4}\right) &= \frac{1}{16}\text{Var}\left(\frac{\sum_{i=1}^{n}(x_{iQ1}+x_{iQ2}+x_{iQ3}+x_{iQ4})}{n}\right)\\
&= \frac{1}{16}\text{Var}\left(\pi_{1+}+\pi_{2+}+\pi_{3+}+\pi_{4+}-\pi_{1-}-\pi_{2-}-\pi_{3-}-\pi_{4-}\right)
\end{aligned}
$$
$$(3.16)$$

The generalisation of the previous equation can be written as follow

$$
\begin{aligned}
\text{Var}\left(\frac{X_{Q1}+X_{Q2}+\ldots+X_{Qn}}{n_Q}\right) &= \frac{1}{n_Q^2}\text{Var}\left(\frac{\sum_{i=1}^{n}(x_{iQ1}+x_{iQ2}+\ldots+x_{iQn})}{n}\right)\\
&= \frac{1}{n_Q^2}\text{Var}\left(\pi_{1+}+\pi_{2+}+\ldots+\pi_{n+}-\pi_{1-}-\pi_{2-}-\ldots-\pi_{n-}\right)
\end{aligned}
$$
$$(3.17)$$

The Indicator of the Evolution of Individual Responses

In the same logic as for the variance, a proposition is done here of a way to extract more information out of the business survey. As explained in the first chapter, the business survey is answered by the same companies over time, there are some new recruits and companies leaving the survey, but the survey can be referred to as a panel survey. Is it possible to have more information by taking taking the evolution of the individual respondents into account ?

This chapter will address this question by applying a method proposed by Caron et al. (1996) that, by taking all the individual evolution of responses into account, offers a method to calculate an indicator that will be called the indicator of the evolution of individual responses (EIR).

First the calculation will be described for the and the intuition will be developed .

We will here describe and develop this indicator, that we will call Evolution of individual Responses (EIR). It can be understood as the indicator of the changes in individual answers between t-1 and t.

As for the variance, The EIR has no ambition of replacing the BSI but rather to propose an addition information.

Explanation

To understand this new indicator, it's important to see that if we take only one period into account, there are three different possibilities of answers; "negative", "neutral" and "positive". When we take two periods into account, a month (t) and the previous one (t-1) for example, there are nine possible situations as represented in Table 4.1. We can again see that π is used since we will speak about proportions and probabilities of been in a certain group, as done before.

		t		
		-	**0**	**+**
	-	z_{i--}	z_{i-0}	z_{i-+}
$t-1$	**0**	z_{i0-}	z_{i00}	z_{i0+}
	+	z_{i+-}	z_{i+0}	z_{i++}

Table 4.1: Possible observation when taking t and t-1 into account

The same as for the BSI, the EIR take different values. In the case of the BSI, if the answer was negative, it would take value "-1", neutral it would take value "0" and positive it would take value "1".

The EIR defers in the sens that it's a measure of change, so if the answer of a certain respondent is the same at a certain time and at the previous survey, $x = 0$. On the other hand, if a certain participant changes his answer for a more positive answer it will take value 1, except if it's a radical change from a "negative" to a "positive" answer, then it will take value 2. Same the other way around, if it decreases it will take value -1, except for a radical change from "positive" to "negative". We will here use z rather than x to make a clear distinction between the BSI and the EIR.

A summary to better understand this new indicator can be seen in Table 4.2.

		t		
		-	**0**	**+**
	-	$z_{i--} = 0$	$z_{i-0} = 1$	$z_{i-+} = 2$
$t-1$	**0**	$z_{i0-} = -1$	$z_{i00} = 0$	$z_{i0+} = 1$
	+	$z_{i+-} = -2$	$z_{i+0} = -1$	$z_{i++} = 0$

Table 4.2: Value given at each type of z_i

4.1 Indicator of the Unweighted Evolution of Individual Responses

The Indicator of the evolution of the individual responses can be obtained by taking the mean of the values, as defined in Table 4.2, the formula is then

$$
\mathrm{E}(Z) \quad = \quad \frac{\sum_{i=1}^{n} z_i}{n} \tag{4.1}
$$

$$
= \quad \left(\frac{\sum_{i=1}^{n} z_{--i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{-0i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{-+i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{0-i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{00i}}{n} \right)
$$

$$
+ \left(\frac{\sum_{i=1}^{n} z_{0+i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{+-i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{+0i}}{n} \right) + \left(\frac{\sum_{i=1}^{n} z_{+-i}}{n} \right) \tag{4.2}
$$

As for the indicator, we will now move from individual responses, to proportions, that can be seen in Table 4.3.

		t	
	-	**0**	**+**

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

Table 4.3: Possible observation when taking t and t-1 into account

The EIR can now be calculated with the following expression after taking out π_{--}, π_{00} and π_{++} since z_{--}, z_{00} and z_{++} have value zero.

$$\mathrm{E}(Z) = \pi_{0+} + \pi_{-0} - \pi_{+0} - \pi_{0-} + 2\pi_{-+} - 2\pi_{+-} \tag{4.3}$$

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

?Interpretation

4.2 Indicator of the Weighted Evolution of Individual Responses

bla bla bla

4.3 Take different questions into account

bla bla bla

4.4 Generalisation for n time frames

Changes in the economy are usually taking several months to influence all the companies. There can be some lag of effects on for example larger companies, of a very specific sector. The idea of only taking a certain month and the previous month can seems non-sufficient, it's relevant to take a larger period into account.

Different methods where explored to take more than two periods into account, but it was found as been flawed and very complex to interpret. We rather propose here a simple generalisation; use t-n rather than t-1. You compare then that period t-n with the the answer at t, without taking the between answers into account.

t-n vs t

come up to the same except that you use a larger period into account

---

The Variance of the Evolution of Individual Responses

The new indicator, the evolution of individual responses, also has a variance that has some interpretation interest. While the EIR is a measure of in which direction companies change there answer, the variance of the EIR can be understood as the measure of changes in answers. Therefore the Var(EIR) can be seen as the volatility of the indicator, in the sens that the variance of the EIR account for the dispersion of the difference in answers over a two times period.

difference var(BSI) & var(EIR)

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.

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

5.1 Variance of the Unweighted Evolution of Individual Responses

$$
\begin{aligned}
Var(Z) &= \pi_{0+} + \pi_{-0} + \pi_{+0} + \pi_{0-} + 4\pi_{-+} + 4\pi_{+-} \\
&\quad -(\pi_{0+} + \pi_{-0} - \pi_{+0} - \pi_{0-} + 2\pi_{-+} - 2\pi_{+-})^2 \\
&= \pi_{0+} + \pi_{-0} + \pi_{+0} + \pi_{0-} + 4\pi_{-+} + 4\pi_{+-} - E(Z)^2 \\
&= 1 - \pi_{++} - \pi_{00} - \pi_{--} + 3\pi_{+-} + 3\pi_{-+} - E(Z)^2
\end{aligned} \tag{5.1}
$$

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

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

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

5.2 Variance of the Weighted EIR

5.3 Properties

quite similar to Var(BSI)

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

Property 2:

5.4 Take different questions into account

20

Non-Response, Dropout, Attrition and Seasonal Effects

Before we model the data, it's important to look at the different effects that could mislead the outcome of the analysis. After exploring the data and doing a literature review, there are four effects that could mislead the results. Three are due to the survey and one to external effects, seasonal effects. We will first look into the three survey based issues; non-response, dropout and attrition, and will then explain the seasonal correction applied to the data.

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

6.1 Non-Response

Non-response is one of the most studied issues in Surveying. Indeed it is a large issue two ways - non participation to the survey - participating but not answering

- non participation to the survey has already been mentioned before, the stratified procedure should avoid as much as possible it, and in the main time, since the importance is the evolution of the Indicator.

- participating but not answering solved by the National Bank by filling the non-answering by the previous response of the participant. The response rate is usually around 95% for the business survey. We could from there say that non-response is not very important. Before concluding it, since the mean study is the evolution rather than the indicator in it self, we plotted the non-response with the evolution of the Year on Year GDP to see if there seems to be a relation between the two. For example we could expect more or less non-response during crises or in certain economy situations. Since the main interest of the BSI is to explain Business cycles, this could be an issue.

plot non-response

6.2 Dropout and Attrition

Those two issues are related to the structure and organisation of the survey. As explain in section 2.2, ones participants are recruited to participate in the business survey, they stay as long as they want. This brings to major potential issues: dropout and attrition

6.2.1 Dropout

Participants leaving the survey, are they different from

The National Bank doesn't keep track from reasons why participants leave the survey. From discussions with the employees at the NBB, the two mean reasons are (1) the company going bankrupt, acquired or merged and (2) the responsible person at the company leaves his job there and the new responsible person doesn't see the interest in participating anymore. This is an issue, since it means that it's a certain type of profiles that leave the survey. If this type of profile have a different opinion or respond differently than the remaining companies this will create bias.

This is the case

It could be argued that the bias is very diffused due to the small amount of companies leaving the survey each month. Again the fact that the evolution is important, means that we have a very small bias for each month (if we take a long period into account then the bias become larger) we are only comparing month to month evolution, and when the business survey is published, the 3-4 last year are showed.

It's not the subject of this work, so will not dive more into this bias, but we would recommend to have a closer look into this.

6.2.2 Attrition

Attrition, also called Panel Conditioning, is present when participants change there behaviour between different rounds of sampling. A very interesting master thesis was done about the Belgian Labour Force Survey, where attrition was found to be significant Priyana Hardjawidjaksana (2019). The Belgian Labour Force Survey was convenient to test for attrition, since it's a survey where respondents reply four times to the same survey with a lag of six months.

In the case of the Business Industry Survey it's a harder to test for it since we have only two major periods of recruitment for the period at interest (1988 - 2018); in the early 1900 and between XXX and XXXX with some company

plot of the amount of recruited over time

Two possible indicators to show the presence of Attrition, is the correlation between time and variance(BSI) and the correlation between time and variance(EIR).

corr(time,var(BSI))

corr(time,var(EIR))

Non parametric test Das et al. (2011)

6.3 Seasonal Effects and Correction

The National Bank, before publishing the business survey indicator, applies a X11 seasonal correction.

The literature about seasonal effects is very rich and variate

JDemetra+

The department of Research and development of the NBB developed JDemetra+ and has since been recommended by the European Central Bank (ECB) and the European Statistical Office (Eurostat) for all National Statistics Institutes (NSI) of the European Union.

Test for seasonality where done with JDemetra (see appendix page 47) and it was concluded that there was seasonal effect for each of the variables at hand.

Seasonality Tests

bla bla

Seasonality Test	BSI	Var(BSI)	EIR	Var(EIR)	EIR2	Var(EIR2)	EIR3	Var(EIR3)
Auto-correlations at seasonal lags	YES	YES	YES	YES	YES	YES	YES	YES
Friedman (non parametric)	YES	YES	YES	YES	YES	YES	YES	YES
Kruskall-Wallis (non parametric)	YES	YES	YES	YES	YES	YES	YES	YES
Spectral peaks	YES	YES	YES	?	YES	?	YES	YES
Periodogram	YES	YES	YES	YES	YES	YES	YES	YES
Seasonal dummies	YES	YES	YES	YES	YES	YES	YES	YES
Seasonal dummies (AMI)	YES	YES	YES	YES	YES	YES	YES	YES

Table 6.2: Seasonality Tests

More details in Appendix

Seasonal Correction

The Seasonal Correction is done with RJDemetra, the R package based on JDemetra+.
The results can be seen in
add var after JDemetra
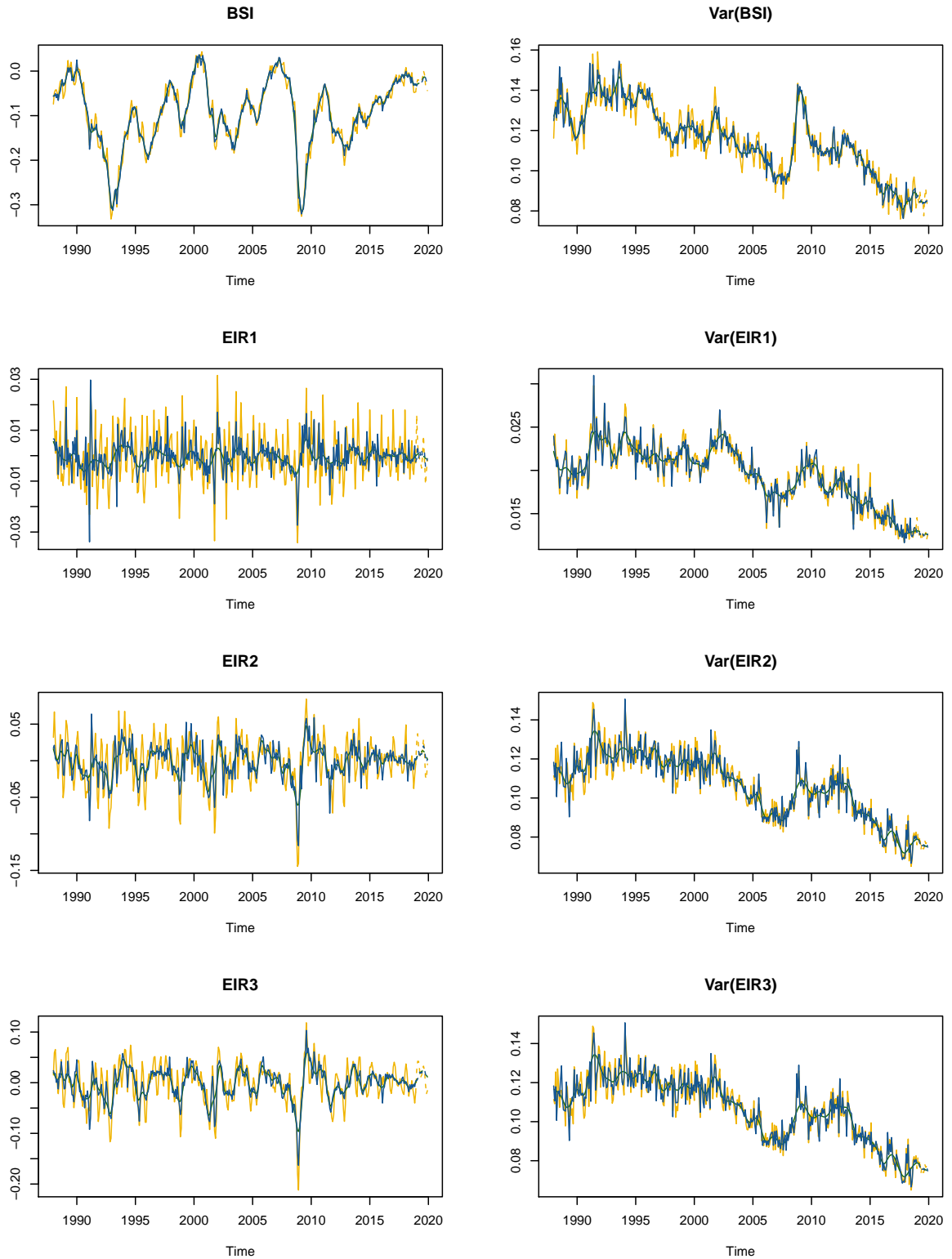YoY GDP is already, by nature, seasonally corrected.
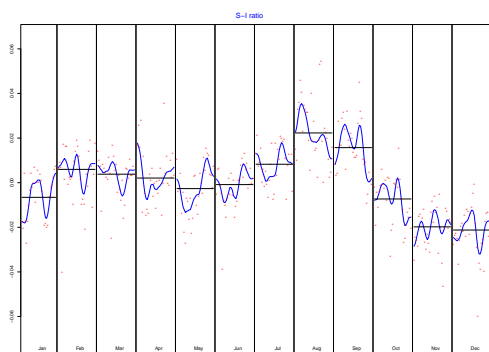
Figure 6.1: Plot of the industry business survey indicator (NS975), the indicator of the evolution of individual responses with the previous month (EIR1), two months (EIR2) and three months earlier (EIR3), with for each of them, their variance. The yellow lines are the raw data, the blue lines the seasonally corrected data and the green line is the trend of the variable.

24

(a) BSI
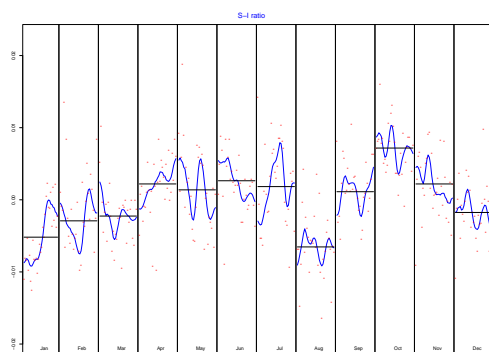
(b) Var(BSI)
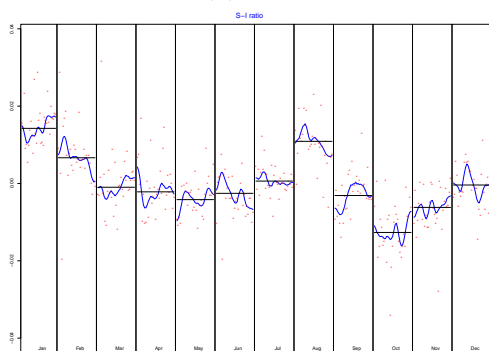
(c) EIR1
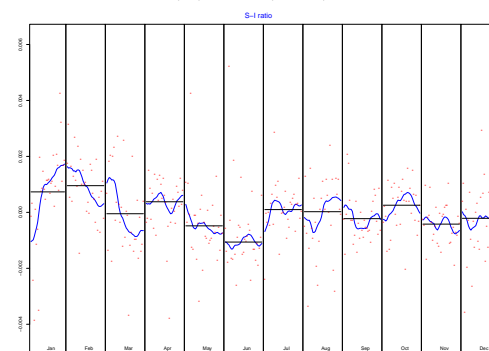
(d) Var(EIR1)
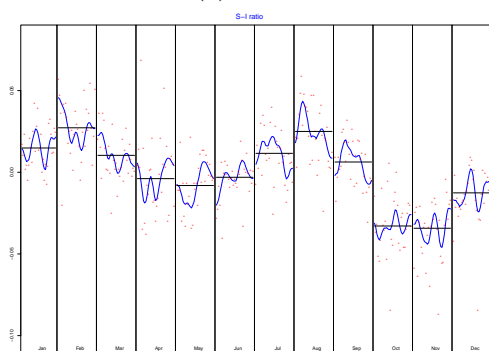
(e) EIR2
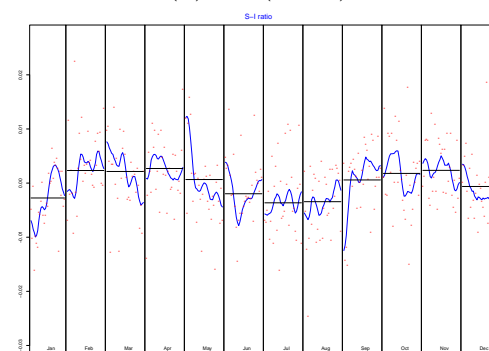
(f) Var(EIR2)
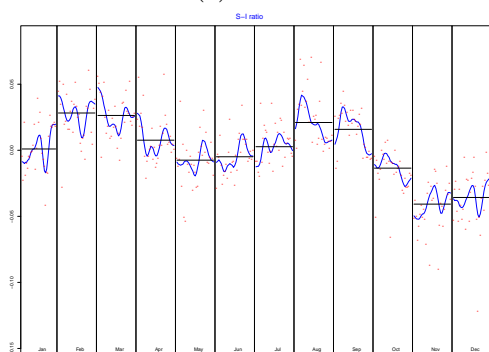
(g) EIR3
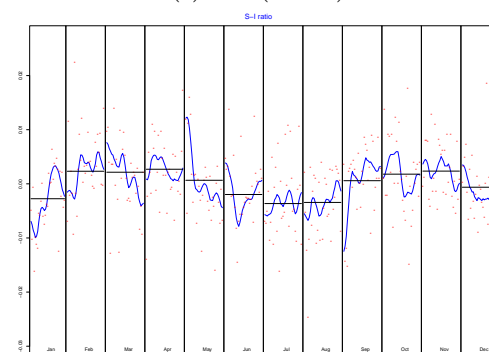
(h) Var(EIR3)

Figure 6.2: plots of....

25

Exploratory Analysis

The previous chapter already did some of the exploratory analysis.

We will here have some further observations based on the time trends of the different variables at hand. In a second time w

Data At hand

Small remember - Four questions summarized together from the industry survey - 1988 - 2018 - BSI, var BSI, EIR and EIR2

Some things have to be looked at before starting the modelling process

NON Weighted ! to complicated

7.1 The industry business survey indicator, the individual evolution in responses and their variances

plot of the different variables

7.2 Correlations Analysis

There are three different correlations that need to be looked at

7.2.1 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

Table 7.2: Correlation Matrix

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

7.2.2 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

Table 7.3: 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
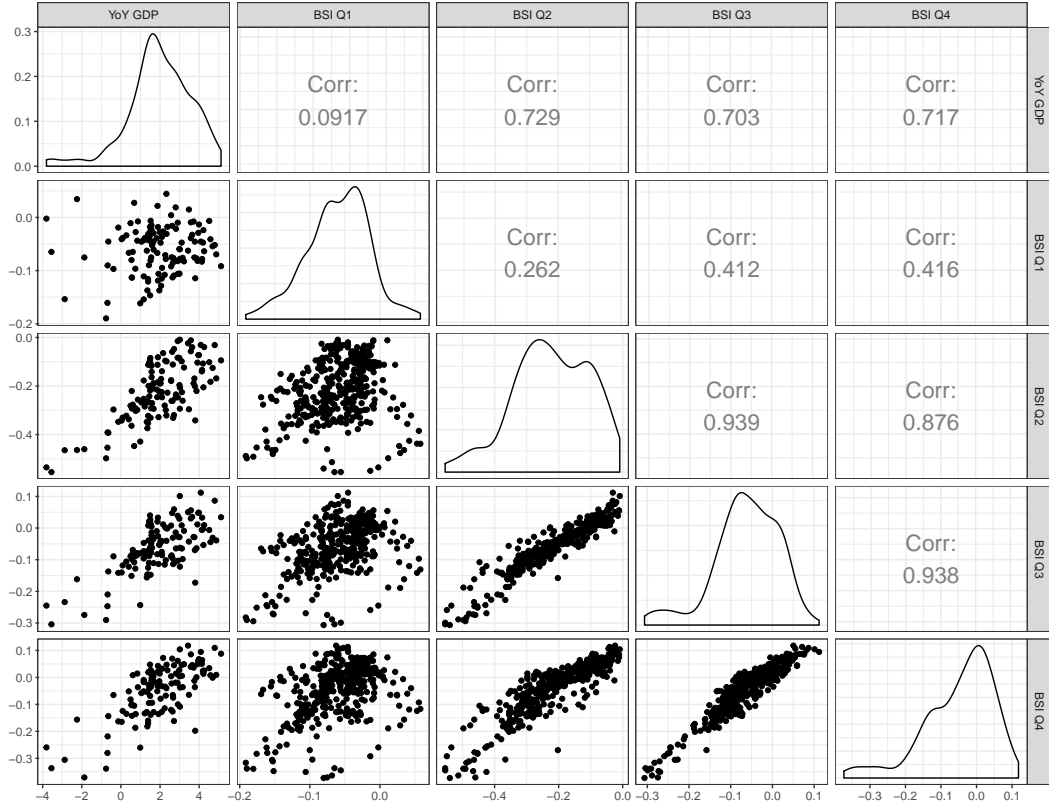
Figure 7.1: Plot

Table 7.4: 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
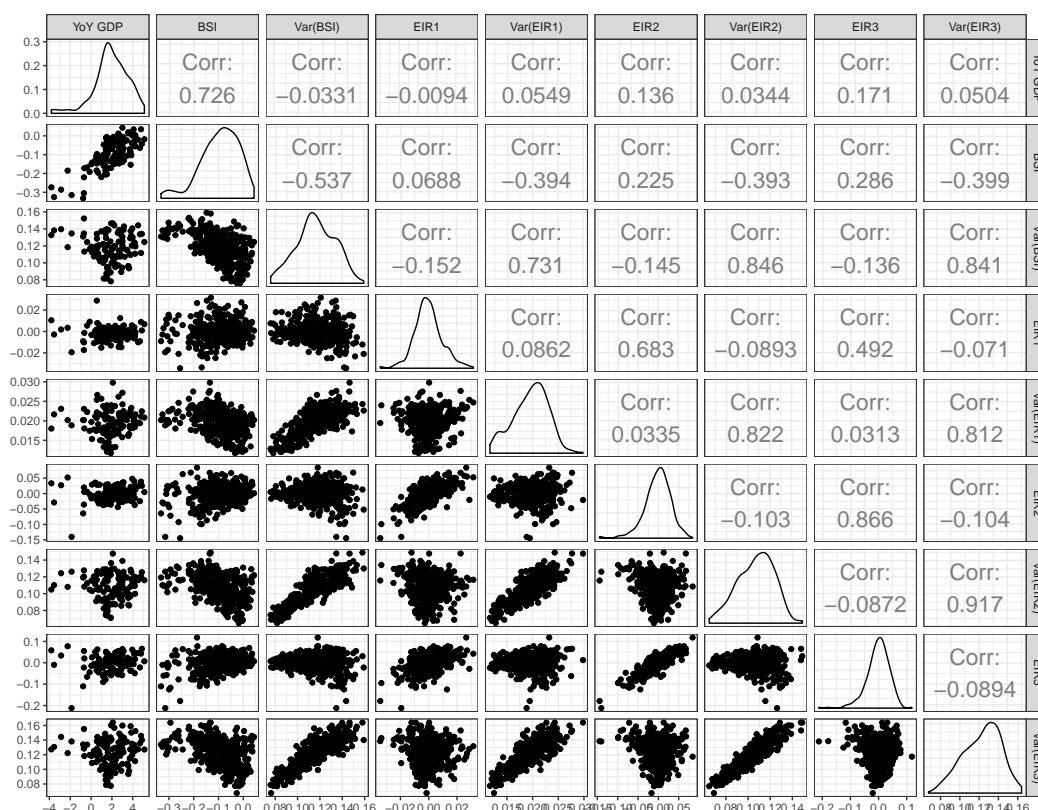
Figure 7.2: Plot

GDP vs Indicator

GDP vs Var

GDP vs Z

GDP vs Var Z

7.3   Auto-Correlation

Test stationarity of the time series (ADF) In order to test the stationarity of the time series, let's run the Augmented Dickey-Fuller Test using the adf.test() function from the tseries R package.

First set the hypothesis test:

The null hypothesis H0 : that the time series is non stationary The alternative hypothesis HA : that the time series is stationary

adf.test(candyts)

Figure 7.3: Plot

Figure 7.4: Plot

7.4 Correlation between Z and Z2 and var(Z) and var(Z2)

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

Table 7.5: Correlation Matrix

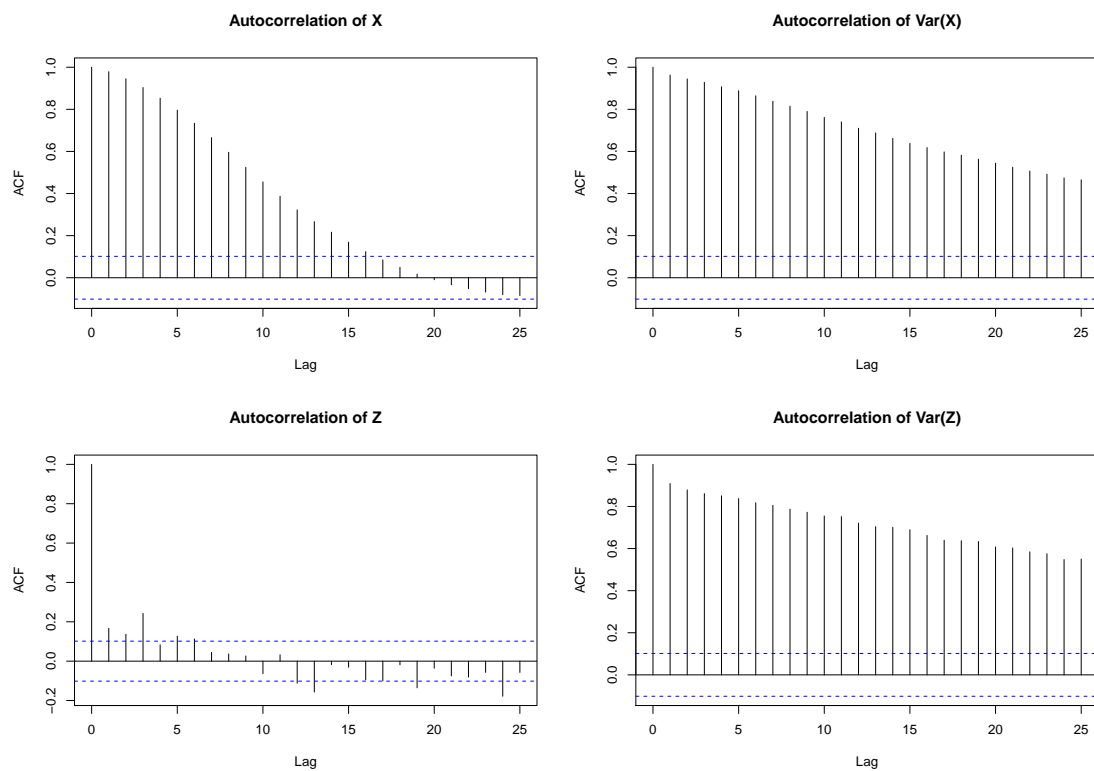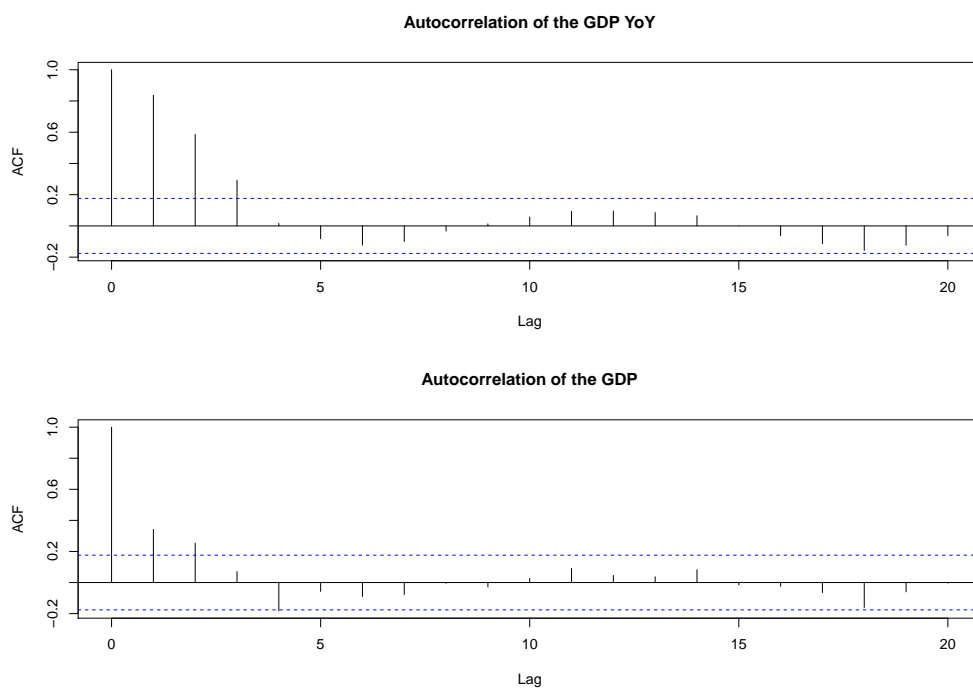|	GDP	GDP_year	E_3	E_3_lag1	E_3_lag2	E_3_lag3	E_3_lag4
GDP	1	0.628	0.477	0.520	0.545	0.546	0.531
GDP_year	0.628	1	0.707	0.679	0.673	0.628	0.560
E_3	0.477	0.707	1	0.969	0.948	0.906	0.846
E_3_lag1	0.520	0.679	0.969	1	0.975	0.940	0.892
E_3_lag2	0.545	0.673	0.948	0.975	1	0.974	0.933
E_3_lag3	0.546	0.628	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 7.6: 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

Linear Models

In the process of deciding which modelling technique to use, a

The interest of this paper is to study the variance of the business survey indicator and explore the possible interest of the indicator of the evolution in individual responses. To achieve this objective, it's important to chose a model that account easily for the interest of each model

Better model could be used; Linear Autoregressive, ARIMA, State Space Models, ... but here interest is to ...

8.1 Method
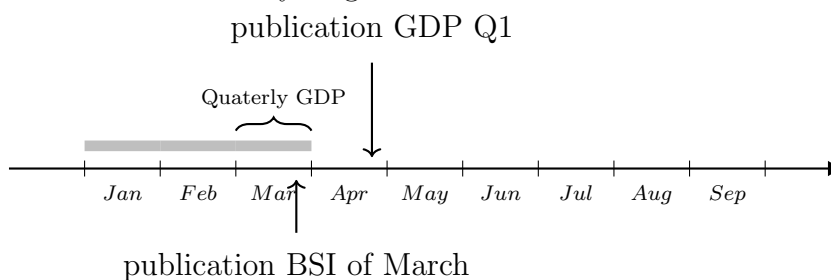
Why we are doing this

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 the last month of the quarter. This is the common way to go since

publication GDP Q1

Quaterly GDP

| Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep |

publication BSI of March

8.2 Linear Model

$$YoY\ GDP_t = \beta_0 + \sum_{i=1}^{n} \beta_i X_t + \epsilon_t \tag{8.1}$$

GDP_t GDP growth over the last semester
X_t monthly predictors
β_0 constant
β_i regression coefficients
bla bla bla

Table 8.1

	Linear Regression				
	Year on Year GDP (in %)				
	(1)	(2)	(3)	(4)	(5)
Constant	3.429***	−1.740***	−1.821***	−1.756***	−1.761***
	(0.160)	(0.631)	(0.615)	(0.635)	(0.634)
BSI	15.773***	21.443***	21.548***	21.310***	21.116***
	(1.317)	(1.252)	(1.224)	(1.306)	(1.310)
Var(BSI)		49.102***	36.477***	40.631***	39.874***
		(5.869)	(8.089)	(11.895)	(11.844)
EIR1			−28.718**		
			(13.792)		
Var(EIR1)			78.555**		
			(35.187)		
EIR2				−0.709	
				(3.878)	
Var(EIR2)				9.213	
				(11.207)	
EIR3					1.248
					(2.615)
Var(EIR3)					9.944
					(11.141)
Observations	124	124	124	124	124
R^2	0.540	0.709	0.728	0.711	0.711
Adjusted R^2	0.537	0.704	0.719	0.701	0.701
Residual Std. Error	1.112	0.889	0.866	0.894	0.893
F Statistic	143.463***	147.283***	79.773***	73.079***	73.247***

Note: *p<0.1; **p<0.05; ***p<0.01

8.3 Evaluation / Model selection

8.3.1 R-square

8.3.2 AIC and BIC

8.3.3 Mean Square Prediction Error

8.3.4 Diebold-Mariano Test

8.3.5 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.

8.4 Relative Importance

Table 8.2

Z3_sa	Z2_sa	Z_sa	Var_Z2_sa	Var_Z3_sa	Var_Z_sa	Var_sa	E_sa
0.010	0.018	0.027	0.029	0.039	0.094	0.095	1.006

Table 8.3

Z2_sa	Z3_sa	Z_sa	Var_Z2_sa	Var_Z3_sa	Var_sa	Var_Z_sa	E_sa
−0.003	−0.002	0.0003	0.002	0.035	0.038	0.046	0.882

8.5 Variance(X) VS Variance(Z) VS Variance(Z2)

8.6 E(Z) VS E(Z2)

8.7 Take Question 1 out of the calculation of the Indicator

8.8 until 2012 to see if variance still significant, not attrition creating effect

8.9 Model with before 2000 data

Table 8.4

	Linear Regression					
	Year on Year GDP (in %)					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.429***	−1.740***	−1.821***	4.437***	0.660	−0.519
	(0.160)	(0.631)	(0.615)	(0.211)	(1.899)	(2.041)
BSI	15.773***	21.443***	21.548***	17.374***	19.212***	20.327***
	(1.317)	(1.252)	(1.224)	(1.547)	(1.758)	(1.813)
Var(BSI)		49.102***	36.477***		30.545*	31.371**
		(5.869)	(8.089)		(15.269)	(14.992)
EIR1			−28.718**			−37.120**
			(13.792)			(17.899)
Var(EIR1)			78.555**			53.535
			(35.187)			(52.433)
Observations	124	124	124	48	48	48
R^2	0.540	0.709	0.728	0.733	0.754	0.779
Adjusted R^2	0.537	0.704	0.719	0.727	0.744	0.758
Res. Std. Error	1.112	0.889	0.866	0.851	0.825	0.801
F Statistic	143.463***	147.283***	79.773***	126.060***	69.144***	37.845***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8.5

	Linear Regression					
	Year on Year GDP (in %)					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.429***	−1.740***	−1.821***	3.903***	−1.068	−1.764
	(0.160)	(0.631)	(0.615)	(0.180)	(1.102)	(1.163)
BSI	15.773***	21.443***	21.548***	17.494***	21.469***	22.103***
	(1.317)	(1.252)	(1.224)	(1.379)	(1.518)	(1.505)
Var(BSI)		49.102***	36.477***		43.809***	38.891***
		(5.869)	(8.089)		(9.608)	(10.307)
EIR1			−28.718**			−38.430**
			(13.792)			(17.020)
Var(EIR1)			78.555**			63.393
			(35.187)			(46.225)
Observations	124	124	124	88	88	88
R^2	0.540	0.709	0.728	0.652	0.720	0.740
Adjusted R^2	0.537	0.704	0.719	0.648	0.714	0.727
Residual Std. Error	1.112	0.889	0.866	1.083	0.977	0.954
F Statistic	143.463***	147.283***	79.773***	161.027***	109.438***	58.910***

Note: *p<0.1; **p<0.05; ***p<0.01

CHAPTER 9

Conclusion

A large

It was seen that

High correlation between var(BSI) and var(EIR) -¿ people seems to change in the same direction

---

Discussion

10.1   Recruitment procedure and panel data

not real sampling theory

10.2   Z that takes more periods into account

10.3   Limitations

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

10.4   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

leave question 18 out of NS975

10.5   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

Alcaniz, M., Costa, A., Luna, C., Rovira, C., and Guillén, M. (2006). Calculation of the Variance in Surveys of the Economic Climate. *SSRN Electronic Journal*.

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

Caron, N., Ravalet, P., and Sautory, O. (1996). Estimation de la précision d'un solde dans les enquêtes de conjoncture auprès des entreprises. Technical report.

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 Antonio Liedo, D. (2014). Nowcasting Belgium. *SSRN Electronic Journal*.

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.

European Commission, T. (2016). The Joint Harmonised EU Programme of Business and Consumer Surveys User Guide. page 56.

Foroni, C. and Marcellino, M. (2014). A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates. *International Journal of Forecasting*, 30(3):554–568.

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

Mitchell, W. C. and Burns, A. F. (1946). *Measuring Business Cycles*. NBER.

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
ECB	European Central Bank
EIR	Evolution in individual Responses
Eurostat	The European Statistical Office
GDP	Gross Domestic Product
NBB	The National Bank of Belgium
NBER	The National Bureau of Economic Research (US)
NSI	National Statistics Institutes

List of Figures

List of Tables

0.1 Questions NS975

Questions taken into account for NS975: ....
originally question Q18, 27, 32 and 33, for simplicity numbered here as 1, 2, 3 and 4.

Verloop en beoordeling

1. Uw huidige voorraad van dit product beschouwt u, voor het seizoen, als:
   □ hoger dan normaal (te hoog) □ normaal (voldoende) □ lager dan normaal (te laag)

2. Uw huidige gezamenlijke orderpositie voor dit product beschouwt u als:
   □ hoger dan normaal □ normaal □ lager dan normaal

Vooruitzichten voor de volgende drie maanden

3. Het personeel (arbeiders en technici) tewerkgesteld voor de fabricatie van dit product zal volgens u:
   □ worden uitgebreid □ onveranderd blijven □ worden verminderd

4. De vraag van uw klanten naar dit product zal volgens u:
   □ belangrijker □ even belangrijk □ minder belangrijk
   zijn dan gewoonlijk tijdens die periode van het jaar.

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

Test for seasonality

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 in 2018

NATIONALE BANK VAN BELGIE

Departement Studiën

CONJUNCTUURONDERZOEKINGEN

de Berlaimontlaan 5 - 1000 BRUSSEL
TEL. (02) 221 49 97
TELEFAX (02) 221 31 07

> Kruis het vakje aan dat overeenstemt met uw antwoord en stuur **één exemplaar** van de vragenlijst terug **binnen de 10 dagen.**
> **Het andere exemplaar is bestemd voor uw dossier.**

NIJVERHEID

De geheimhouding van de antwoorden is gewaarborgd

Het gedeelte onder de stippellijn terugsturen

Beschouwde maand
←————————

Produkt :
● 4100 ————————————

(left margin, vertical): Nationale Bank van België - 1990 - Alle rechten tot gebr.·'· van de vragenlijst voorbehouden.

Tijdens de beschouwde maand...

15 ... is ons **produktietempo** voor dat produkt

versneld   **15** ☐ 1
onveranderd gebleven ☐ 5
vertraagd ☐ 9

t.o.v. de vorige maand.

(Geen rekening houden met de schommelingen die alleen voortspruiten uit het van maand tot maand veranderlijk aantal werkdagen of die te wijten zijn aan betaald verlof.)

17 ... zijn de **verkoopprijzen** van dat produkt

gestegen   **17** ☐ 1
onveranderd gebleven ☐ 5
gedaald ☐ 9

t.o.v. de vorige maand.

(Geef de tendens van uw prijzen aan op basis van uw contracten of uw aanbiedingen; eventueel rekening houden met prijsherzieningen.)

18 Onze **huidige voorraad** van dat produkt moet worden beschouwd als

hoger dan normaal   **18** ☐ 1
normaal ☐ 5
lager dan normaal ☐ 9

Wij hebben nooit een voorraad van dat produkt ☐

(Bij het beantwoorden van de volgende twee vragen (22 en 26), mag U geen melding maken van de zuivere seizoenstijgingen of -dalingen die het verloop van de bestellingen gedurende de beschouwde maand kunnen beïnvloed hebben. Men moet dus de tendens van de bestellingen doen uitkomen.)

22 ... zijn de **bestellingen**, vanwege de binnenlandse markt, voor dat produkt

vermeerderd   **22** ☐ 1
onveranderd gebleven ☐ 5
verminderd ☐ 9

t.o.v. de vorige maand.

Wij leveren dat produkt nooit op de binnenlandse markt ☐

(Rekening houden met de bestellingen, ontvangen van andere afdelingen van uw firma, en met maakloonwerk.)

26 ... zijn de **bestellingen**, die wij voor dat produkt inschreven **bij de uitvoer**

vermeerderd   **26** ☐ 1
onveranderd gebleven ☐ 5
verminderd ☐ 9

t.o.v. de vorige maand.

Wij voeren dat produkt nooit uit ☐

(Rekening houden met maakloonwerk.)

27 Onze **huidige gezamenlijke orderpositie** voor dat produkt mag worden beschouwd als

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

(Indien uw onderneming uitsluitend uit voorraad levert, dient U « orderpositie » op te vatten als « het peil van de vraag » naar dat produkt.)

28 Indien wij het huidige fabricatietempo voor dat produkt handhaven is onze **activiteit** nog **verzekerd** voor ongeveer ☐ maand(en) of gedeelte van maanden

(Te ramen op basis van uw orderpositie of, bij gebrek hiervan, op basis van uw produktieplannen.)

31 Onze **huidige positie inzake bestellingen uit het buitenland**, voor dat produkt, mag worden beschouwd als

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

(Geef, bij gebrek aan orderpositie, uw oordeel over het peil van de buitenlandse vraag naar dat produkt.)

Vooruitzichten.

Tijdens de volgende drie maanden...

32 ... zal, naar wij voorzien, het **personeel** (arbeiders en technici) tewerkgesteld aan de fabricatie van dat produkt

vermeerderen   **32** ☐ 1
onveranderd blijven ☐ 5
verminderen ☐ 9

(Het instellen van gedeeltelijke werkloosheid dient als een vermindering van het personeel te worden beschouwd.)

33 ... zal, volgens onze inlichtingen, **de vraag van onze klanten** naar dat produkt

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

zijn dan gewoonlijk gedurende die periode van het jaar

(Dus de tendens van de vraag der klanten weergeven met uitsluiting van de zuivere seizoenstijgingen of -dalingen, die haar verloop gedurende de volgende drie maanden zouden kunnen beïnvloeden.)

34 ... zullen onze **verkoopprijzen** van dat produkt waarschijnlijk

stijgen   **34** ☐ 1
onveranderd blijven ☐ 5
dalen ☐ 9

● Vermeld alle schommelingen (van uw produktietempo, bestellingen, enz.) zelfs indien ze van geringe omvang zijn.

● Antwoord elke maand op alle vragen. Indien U **gedurende een maand, in tegenstelling met de voorgaande maand**, het bestudeerde produkt niet heeft geproduceerd (of geen bestellingen heeft ontvangen, enz.) dient U toch deze vraag te beantwoorden en vermeldt U « verminderd ». Antwoord « onveranderd » gedurende de maand(en) dat deze toestand voortduurt totdat de produktie (of de bestellingen, enz.) hernemen; op dat ogenblik, vermeldt U « vermeerderd ».

Op de vraag betreffende de prijzen slechts antwoorden indien U gedurende de beschouwde maand werkelijk contracten afsloot of aanbiedingen deed.

● **Stuur uw antwoord terug vóór de 10ᵉ van de maand die volgt op de bestudeerde maand, zoniet brengt U de snelle mededeling van de resultaten in het gedrang.**

Figure 2: The Business Survey Questionnaire in Dutch for the Industrial Sector in 1990

50
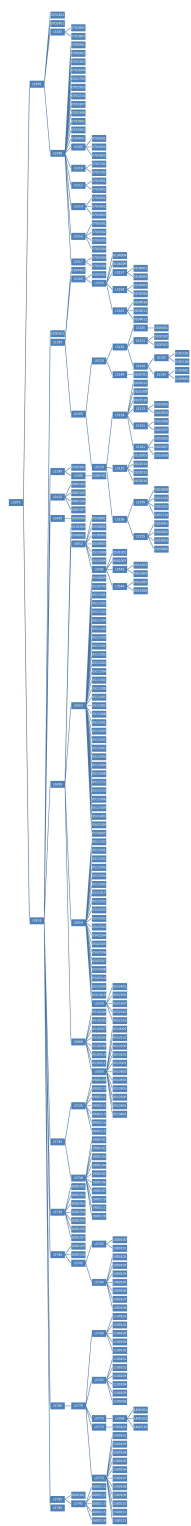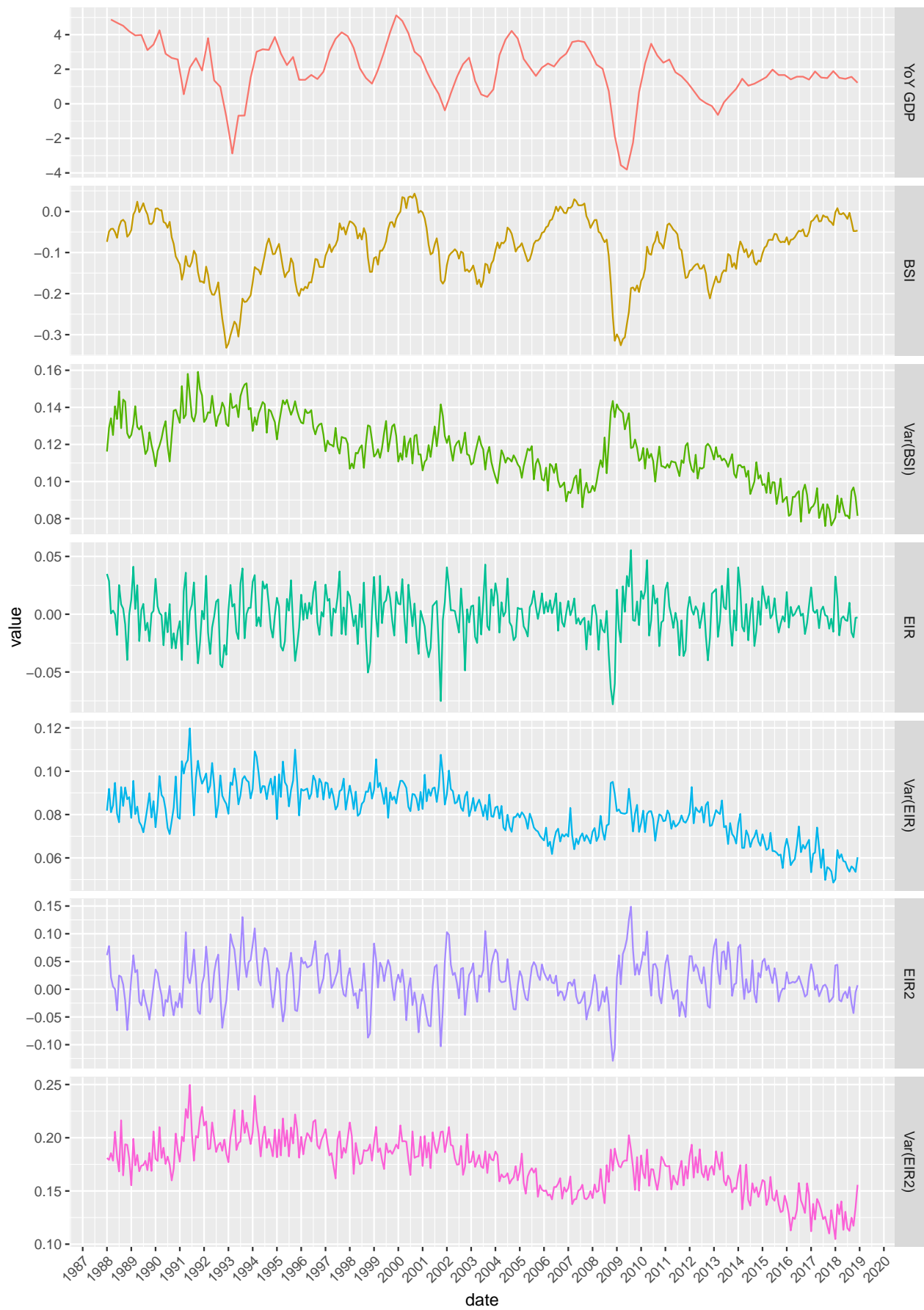
Figure 3: XXXX

Figure 4: Plot

52

Code

R code for Seasonal Adjustment

R code for Creating Lags

R code for Linear (Auto-Regressive) Models

https://github.com/fabricevb/Master-Thesis/blob/master/R%20Code/Linear%20Regression.R

R code for Markov Switching Models