BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

A Comparison of Norwegian Business Cycle Indicators

Navn: Erik Hulleberg, Emil Sildnes

Start: 15.01.2020 09.00

Finish: 01.09.2020 12.00

TABLE OF CONTENTS

ABSTRACT

This thesis investigates how different economic indicators can classify economic recession and forecast GDP in Norway. The study focuses particularly on the Financial News Index (FNI) and the financial crisis. In-sample evaluation is done through Receiver Operating Characteristics (ROC) and analyzed with the help of AUROC score. Out-of-sample evaluation is done through simple linear regression and analyzed with the help of Mean Squared Error (MSE) and Root Mean Square Error (RMSE). With both in-sample and out-of-sample analyzes, our results indicate that FNI should be considered as a highly valuable economic indicator in the future.

ACKNOWLEDGEMENTS

We wish to thank our supervisor Leif Anders Thorsrud at BI Norwegian Business School for contributing with an inspiring topic and literature, and for the guidance during our research. His contribution with access to FNI data and analytical methods has been invaluable to the final result. We would also like to thank Finans Norge/Kantar TNS for allowing us to use dataset regarding the Consumer Confidence Index.

1. INTRODUCTION

We started this thesis by focusing on economic crises and especially the financial crisis. In March, the whole world included Norway ended up in a current economic crisis. The economy changes rapidly, and it is, therefore, vital for policy makers to get an accurate and up-to-date view of the economy as soon as possible. The Financial crisis and Covid-19 are costly economic events, and the outcome may depend on how early the information is obtained.
Policy makers have to make a judgement of the state of the economy, which is a fact about millions of transactions and activities across a wide geographical area. There is no formal definition of what economic activity is, and the true state of the economic activity (expansion/recession) is unobservable, even retrospectively (Travis and Jordà, 2011).

Early signals of the crisis will provide an opportunity for better countercyclical measures and dampen the negative effect of the crisis. The information is crucial to determine the interest rate-setting for policy makers. In Norway, the policy rate dropped to zero in May (zero lower bound), which makes government spending much more effective when monetary policy is ineffective. The financial support packages are assessed at a time when critical financial variables have not been published. One of the most significant problems during economic crises is the lack of data on essential variables. GDP is a measure designed specifically to give a unifying picture of the overall economy. Hence, GDP is the gold standard by which recessions are a valid measure. It is not observable as it is compiled on a quarterly frequency and published with a considerable lag at the end of the corresponding reference period (Thorsrud, 2018). Accordingly, various economic indicators must be used to provide a comprehensive picture of the economy.

Several researchers have tried to solve this problem using indicators with higher frequency than GDP. Estrella and Mishkin concluded that the yield curve spread could have a useful role in macroeconomic prediction. Rundebusch and Williams (2009) documented that the term spread is an essential tool in predicting recessions. Travis and Jorda (2011) compared various leading indicators. Some indicators are considered leading, such as stock prices (Estrella and Mishkin, 1998), oil prices (Hamilton, 1996) and survey data (Hansson et al., 2005). High frequent indicators are a valid measure for capturing the changes in the economy. Nevertheless, few have used alternative daily data sources that capture what the media writes about various topics.

Our contribution to this literature is to look at Norway and a new type of leading indicator. The indicator applies a different type of data that has not been used before. This is the Financial News Index (FNI). In addition to FNI, we use other indicators that collect data information more traditionally. The reason for this is to make a comparison of the indicators specifically for Norway. For this thesis, we will have a particular focus on the financial crisis and analyze how different critical leading indicators for the Norwegian economy performs in terms of catching the Norwegian business cycle and predicting sharp turns. We are using in-sample estimates to evaluate the accuracy and classification of each indicator. Furthermore, we use out-of-sample estimates to evaluate each indicator´s ability to predict GDP growth. The in-sample evaluation is performed by the Receiver Operating Characteristic (ROC) method, while the out-of-sample prediction is performed by simple linear regression. We use AAstveit et al. (2016) recessions details and use these as the “truth” for Norwegian business cycles.

Our in-sample estimates are evaluated by AUROC score, and out-of-sample estimates are evaluated using Mean Squared Error (MSE) and Root Mean Square Error (RMSE). Out in-sample results state that the Consumer Confidence Index (CCI), the Business Tendency Survey (BTS) and FNI are the best indicators to classify business cycles. FNI maintains the AUROC score better than BTS at different sample-sizes. Our out-of-sample results conclude that FNI outperforms the other indicators, especially on the financial crisis.
Howrey (2001) found that consumer confidence index is a significant predictor for the rate of GDP growth in the future and the possibility of a recession. These results are consistent with our findings, where CCI has the highest AUROC score and second-best MSE and RMSE score. The predictive power of business surveys is addressed in numerous studies. Hansson et al. (2003) study the forecasting performance of business survey data in Sweden. They use a DF model and find good results for forecasting GDP-growth. Existing literature concludes that both CCI and BTS are good indicators of macroeconomic considerations, which our

results verify. There is little available documented literature about FNI, which is the best indicator in our thesis, and thus require more consideration in the future.

2. METHODOLOGY

In our thesis, we will apply a quantitative research method by studying five macroeconomic indicators. We will use both in-sample and out-of-sample predictions. MSE and RMSE are crucial tools to evaluate out-of-sample forecast results, while AUROC evaluates in-sample results.

2.1 LEADING, LAGGING AND COINCIDENT INDICATORS

We will compare observations for these indicators with their respective historical values in terms of classifying recessions in the Norwegian economy. Which means that we must set a threshold that the indicator always has been below when signalling a recession.

There are several challenges regarding these indicators, and the relationship between timely indicators and GDP are unstable (Stock & Watson, 2003). To distinguish between indicators, we can separate into leading, lagging and coincident indicators. (Stock & Watson, 1988). Leading is an economic factor that changes before the rest of the economy begins to go in a particular direction. In this way, such indicators will be beneficial in predicting the economy in the short term. The consumer confidence index is an example of such an indicator, and these indicators get the most attention.

Lagging is something that occurs after a significant shift in the target variable occurred. These are economic statistics that generally have a delayed reaction to a change in the business cycle. The lag usually comes a few quarters of a year. An example of a lagging indicator is unemployment. It takes time for companies to respond to a decline in production by laying off employees. Coincident change approximately the same time as the whole economy. Further, we can separate through "hard" and "soft" information. "Hard" information whose value has been verified and "soft" information disclosures such as forecast, unaudited statements and press releases (Peterson, 2004). High-frequent financial data are abundant, and the information from coincident indicators are difficult to state whether that drive or reflect economic fluctuations (Thorsrud, 2018). Coincident indicators are made because of the uncertainty regarding the GDP, and therefore valuable to present the current economic situation (Stock & Watson, 1989, Evans 2005, Banbura et al. 2011).

The importance of different indicators with different frequency has increased over recent years. The relative importance of changes in the trend and cyclical swings in explaining the quarterly movements in economic aggregates (Stock and Watson, 1988). Stock and Watson also examine indexes of leading and coincident economic indicators using the tools of modern time series econometrics (Stock and Watson, 1989). The past fifteen years have seen considerable research on forecasting economic activity and inflation using asset prices (Stock and Watson, 2003).
Kelley (2019) examined the accuracy of different indicators in the past recession, and summarize what these indicators suggest about the future. Indexes that combine several macroeconomic measures have historically done better than other indicators at signalling recessions (Kelley, 2019).

2.2 RECEIVER OPERATING CHARACTERISTICS

The Receiver Operating Characteristic (ROC) curve method was already introduced by Peterson and Birdsall (1954). Berge and Jorda (2011) have recently used the ROC to evaluate the recession classification ability of various leading indicators. The ROC curve is a useful measure because it precisely captures the ability of each model to precisely categorize recessions and expansions. Two types of mistakes can be made; the model fails to predict a recession, or the model gives a false recession signal. It maps the true positive rate and false positive. In particular, using the area under the ROC curve (AUROC), we can evaluate the classification ability of the model over a full range of different cuts to determine a recession, instead of evaluating predictive power at any arbitrary threshold. To assess the recession classification ability of different indicators, we follow Berge and Jorda (2011), and Liu and Moench (2016) and calculate the AUROC that takes into account each point on the ROC curve.

As in Travis and Jorda's (2011), we categorize aggregate economic activity into phases of expansions and contractions and evaluate the indicators ability to classify such phases using ROC curves and area under the curve (AUROC) statistics.

We can define the ROC curve as a probability curve, usually shown graphically, which represents the performance of a classification model at all classification thresholds. The curve plots two parameters: true positive rate and false positive rate. Furthermore, we have the terms true positive, true negative, false positive and false negative.

True positive rate (TPR), also called sensitivity, is defined as:

$$TPR = \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

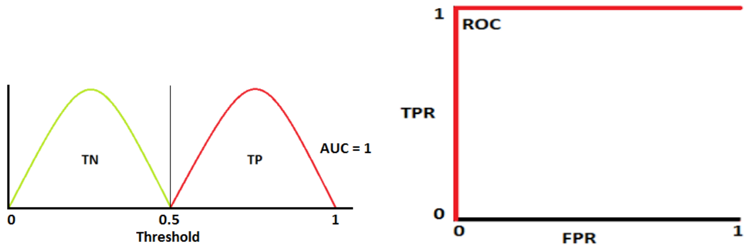
TPR is calculated based on how many correct positive results have been produced in relation to the total number of positive samples available in the test set.

False Positive Rate (FPR) is defined:

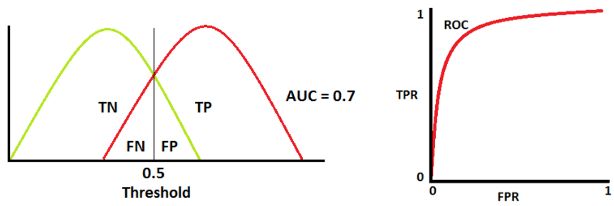
$$FPR = \frac{\text{False Positive}}{\text{False Positive + True Negative}}$$

FPR is calculated based on how many errors produced positive results are predicted, relative to all the negative samples in the test set.

The ideal for a ROC curve would be that it predicts 100% correct results, and this will be equivalent to a singular point at the coordinate (0.1).
A standard measure of the general classification ability is the Area Under ROC (AUROC) curve. For a perfect classifier of the business cycle, AUROC will be 1. This means that it has a good measure of separability. If AUROC is 0.5, then it means that the model does not have class separation capability in that ROC is a curve of probability. We plan the distribution of these probabilities. We can define the ROC curve as a



This is a perfect classifier. These two curves don't overlap, so it is able to separate between positive and negative class.



When two distributions overlap, we get type 1 and type 2 errors. Depending upon our threshold, we can minimize or maximize them. Here we see that AUROC is 0,7, meaning there is a 70 % chance that the model will be able to separate between positive and negative class.

2.3 FORECAST

Simple linear regression is a statistical model with a single explanatory variable. It concerns two-dimensional sample points with one independent variable and one dependent variable. The dependent variable is known as y, and the independent variable is known as x.

(1) $$y_t = \alpha + \beta x_t + \varepsilon_t$$

The whole point of simple linear regression is to find the parameters alpha and beta with the lowest possible error term. The model minimizes squared errors to prevent positive errors from compensating for negative errors, and vice versa.

(2) $$\hat{\alpha} = \min_{\alpha} \sum_{t=1}^{T} (y_t - \alpha - \beta x_t)^2 = \min_{\alpha} \sum_{t=1}^{n} \varepsilon_t^2$$

(3) $$\widehat{\beta} = \min_{\beta} \sum_{t=1}^{T} (y_t - \alpha - \beta x_t)^2 = \min_{\beta} \sum_{t=1}^{n} \varepsilon_t^2$$

This procedure is called Ordinary Least Squared errors (OLS).

The ε_t represents the error term and, α, β are the true unobserved parameters of the regression. In-sample statistics cover t = 1…,T for which we have observations, the out-of-sample statistics cover T+h where h = 1,…,H, for which we do not know the true values (Bjørnland and Thorsrud, 2015).

(4) $$y_{t+h} = \hat{\alpha} + \hat{\beta} x_{t+h} + \varepsilon_{t+h}$$

The estimated beta multiplied with x_{t+h} (which represent the indicator) together with the estimated alpha will give a forecast of y_{t+h} (which represent GDP-growth). Where the error term represents the difference between estimated $\hat{y}_{t+h}$and y_{t+h}. We assume that all the indicators are either at a higher frequency or leading; thus, we have the indicators' observation at x_{t+h}.

The accuracy of the forecast is evaluated using Mean Squared Error (MSE) and Root Mean Square Error (RMSE). The purpose of both measuring devices is to achieve results as close to zero as possible. The MSE is divided into a variance, covariance and a bias. The variance tells how different the predicted values are from the actual values. At the same time, the covariance captures the remaining unsystematic part of the errors. We want most of the forecast error to consist of the covariance so that the error is a result of random events and not systematic features of the data. Finally, the bias section tells us how significant the deviation is between the predicted mean and the actual mean, meaning a higher value indicating a higher degree of systematic error. Root mean square error (RMSE) is the standard deviation of the prediction errors. Mathematically, RMSE is the square root of MSE.

(5) $$\frac{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}{T} \quad \text{(MSE)}$$

(6) $$\sqrt{\frac{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}{T}} \quad \text{(RMSE)}$$

When researchers measure a model´s ability to predict, they are typically more interested and concerned by looking at the out-of-sample forecast rather than the in-sample forecast. We have chosen to focus on the out-of-sample forecast when predicting FNI, CCI, unemployment, rate spread and business tendency survey.

3. DATA

We evaluate a list of economic indicators from a variety of sources. There are two disadvantages when choosing indicators when assessing the state of the Norwegian economy using high-frequency indicators. The first is the selection of appropriate indicators from a wide range, and the second problem is related to signal extraction from the selected indicators. The problem is whether the indicators may reflect short-term idiosyncrasy rather than a business trend. We choose appropriate indicators based on correlations between indicators and GDP and the explanatory power of each indicator around GDP turning points (Bhadury et al., 2019).

Based on these criteria, we have selected five indicators. These five indicators are unemployment, financial news index, consumer confidence index, business trend survey and interest rate spread. The reason these indicators are chosen is their promising results as a macro-economic indicator throughout the years. Hansson et al. (2003) concluded that the business tendency survey served as a good indicator for the Swedish economy in the two last decades (before 2003). Mazurek and Mielcová (2017) proved that CCI was a suitable predictor of GDP for the USA. Unemployment is an essential business cycle indicator for the Norwegian economy (Sparrman, 2012). Interest rate is widely used as an economic indicator around the world. Typically, this applies to 10 years – 3 months, due to the lack of data availability for three months interest rate, we have in this thesis chosen three years instead. There are three reasons why we have not chosen more indicators: Too low frequency, limited data availability for Norway and better focus on those we already have chosen. We are choosing these due to their promising results as macro-economic indicators throughout the years.

3.1 THE FINANCIAL NEWS INDEX

The Financial News Index (FNI) is created at BI Norwegian business school by retriever and CAMP (FNI - Retriever / CAMP BI). The indicator is high-frequent and designed to track Norwegian GDP growth and the business cycle. Its underlying indicators are daily time series representing how much the media writes about various topics. The average value of the index is zero. Progressively bigger positive values indicate progressively better business cycle conditions than average. The construction of FNI is based on the work of Leif Anders Thorsrud (2018), «Words are the new numbers». It is originally on a daily frequency; classification is obtained by constructing quarterly averages of the daily series. Monthly frequency is also used for several analyzes. The same procedure, as mentioned above, has been used for changing the frequency to monthly. FNI has observations from the first day in 1990, and we have used until mid-February 2020.

3.2 CONSUMER CONFIDENCE INDEX

Consumer Confidence Index (CCI) (Finans Norge/Kantar TNS), is based on the premise that if consumers are optimistic, they tend to purchase more goods and services, which should stimulate the whole economy. The CCI is carried out by Kantar TNS on behalf of Finance Norway. The survey is a gauge of Norwegian households' economic expectations and is composed of five categories combined into one main indicator. Combining consumer and business survey data could help improve the forecast of GDP fluctuations (Dées and Brinca, 2011).
It is originally on quarterly frequency, starting 1992Q2 and ending 2020Q1.

3.3 LABOUR FORCE SURVEY

The unemployment rate (SSB, 2020) is an economic indicator based on the labour force survey (LFS) provides information on the proportion of the population who are employed and unemployed. The statistics shed light on the connection to the labour market for different population groups. The statistics are published quarterly, typically four weeks after the end of the quarter. The figures from the

LFS include unemployed who do not register with NAV. Some of those who take part in labour market measures and some of the disabled.
The dataset starts in 1972Q1, but we only use the observations made from 1978Q1. The reasoning behind this is that we do not have available GDP observation until 1978Q1. There are both observations in per cent and by unemployed persons (1000 persons). Both in-sample and out-of-sample analyzes is done with percentage data.

3.4 INTEREST RATE SPREAD

We also evaluate the difference between long- and short-run interest rates (Norges Bank, 2020). The interested rate spread corresponds to the difference between the three-year bond yield and the ten-year bond yield set by Norges Bank. We have chosen to use the three-year interest rate instead of the 3-month interest rate because the data set will not include interest rates during the great recession.
The data set used starts 1987Q1 and ends 2020Q2. It is initially on a monthly frequency, but the classification is achieved by assuming that the economy remains in the phase on each month within the quarterly classification periods.

3.5 BUSINESS TENDENCY SURVEY

Business Tendency Survey (SSB, 2020) is a qualitative survey that studies managers' assessments of developments for characteristics such as production, capital utilization, employment, order access by market and prices. It provides an overview of the current business situation and can predict short-term development. Information from these surveys has proven to be of particular value when it comes to predicting turning points in the business cycle and early signals. We have chosen to use a confidence indicator where the product type is total consumer goods. The figures are smoothed seasonally adjusted and dates back to 1990Q1. The sample used has originally quarterly frequency.

3.6 GROSS DOMESTIC PRODUCT

Gross domestic product (GDP) (SSB, 2020) is the total market value of all the finished goods and services produced within a country's borders in a specific time period. It provides an overview of the state and development of the national economy. We have chosen GDP mainland Norway, market values. Used constant 2017-prices, seasonally adjusted for business cycle dating analyzes and change in volume from the previous quarter, seasonally adjusted (per cent) in out-of-sample analyzes. It dates back to 1978Q1 and ends 2020Q1.

4. EXPERIMENT

In this section, we will show how the data is organized and how the analyzes are performed.

4.1 NORWEGIAN BUSINESS CYCLE DATING

Recession is a macroeconomic term that refers to a significant decline in general economic activity in a designated region. It is typically recognized after two consecutive quarters of economic decline. As a rule of thumb, a recession is defined as two consecutive quarters with economic decline.
The NBER's traditional definition emphasizes that a recession involves a significant decline in economic activity that is spread across the economy and lasts more than a few months (Hall et al., 2003)

We use Aastveit et al. (2016) recessions details and use these as the “truth” for Norwegian business cycles. The business cycle chronology is constructed using a method to extract the unobserved phases: a Markow-switching factor model (MS-FMQ). Aastveit et al. (2016) used different models and found that the peak and through dates provided by a quarterly Markow-switching factor model provided the most reasonable definition of reference Norwegian business cycles. The MS-FMQ business reference cycle are from 1978Q1-2011Q4. Details about the chronology is summarized in Table 6 in Appendix.

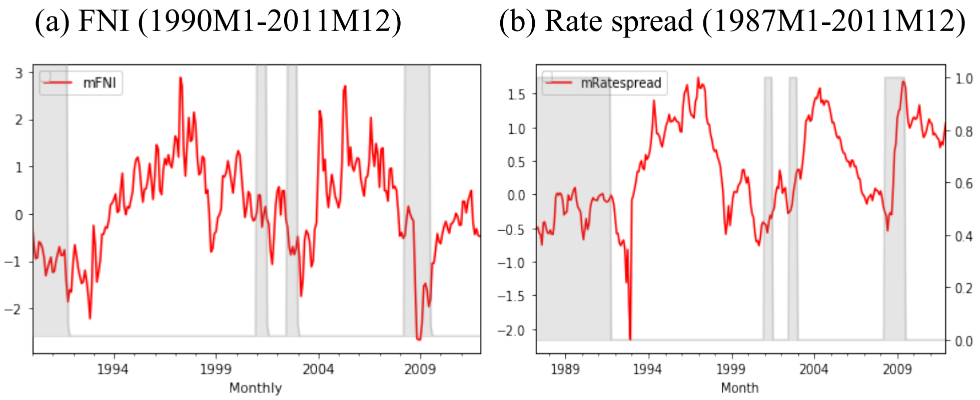


Figure 1: Financial News Index and Rate Spread (10 year-3year) monthly plotted with Aastveit et al. (2016) recession periods. Recessions are illustrated using grey colour shading.

Figure 1 shows the indicators that have available data at a higher frequency than GDP. Details regarding daily frequency are described in Appendix Figure 10. Figure 1 shows that FNI starts with a recession and reaches its peak between 96-97. It seems to classify the second and third recession in the aftermath of the recession itself. The financial crisis gets classified almost entirely, and FNI gets historical bottom halfway through that crisis.

The recessions are based on the GDP data, and thus it is logical that it decreases during the recessions, shown in Figure 2. GDP data can be challenging to decipher. It is easy to see that the two last decades have seen a substantial increase in GDP. With the constant increase, it can be challenging to identify the downturns. During the financial crisis, the GDP data had a significant decline, which confirms the observations made from the other indicators. Figure 2 illustrates that BTS also classifies the financial crisis well and gets its bottom point in the middle of this crisis. It gets a similar curve through the previous recession. LFS shows a clear picture that it is increasing during a recession. Especially the first recession, which is also the longest, this point primarily emerges from Figure 2. Although this is in line with theory, it is crucial to notice. LFS is a lagging indicator and will have the opposite effect than the other indicators. Therefore it is not easy to interpret LFS alongside the others. After recessions, the indicator tends to increase.

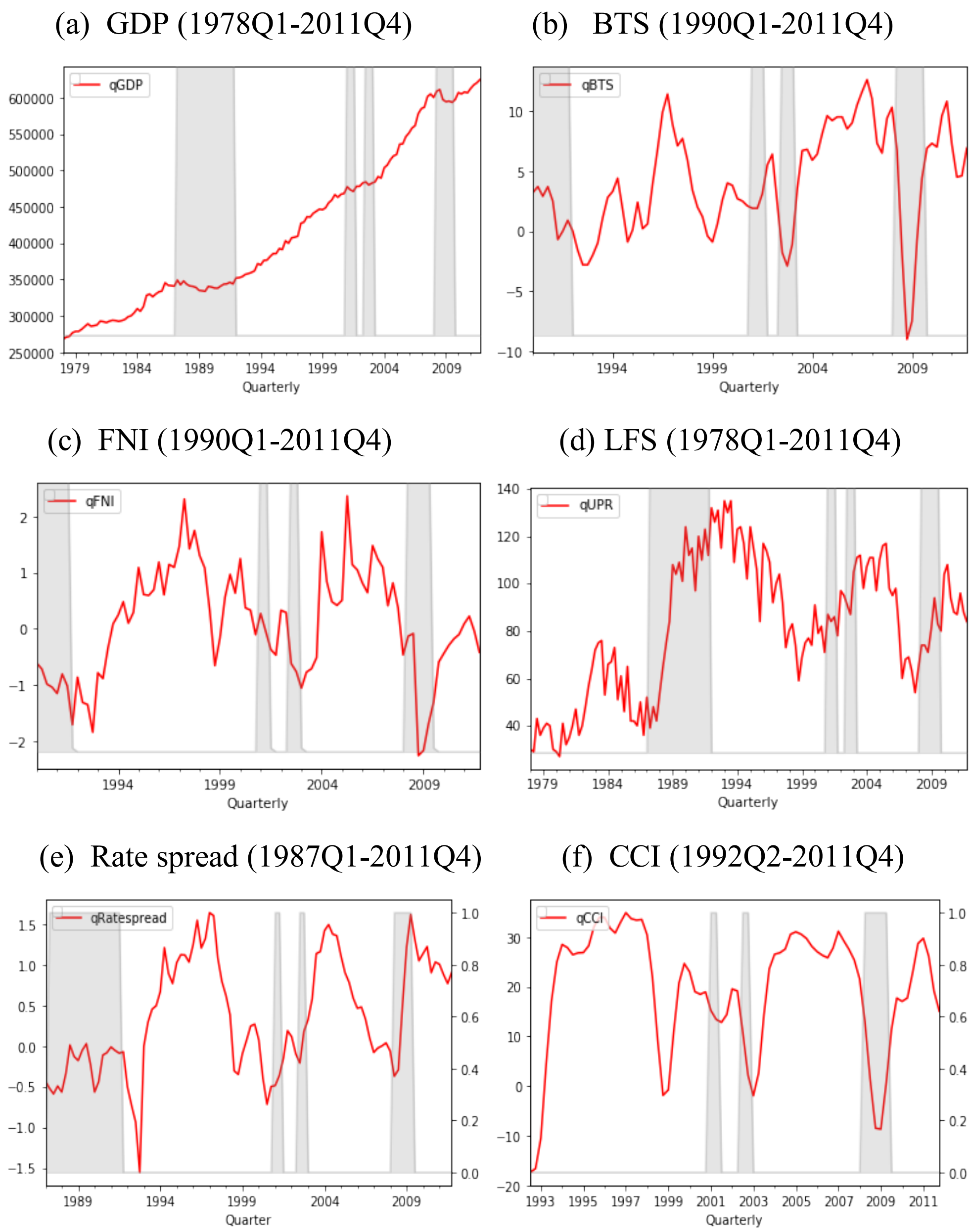


Figure 2: Every indicator and GDP quarterly plotted with Aastveit et al. (2016) recession periods. Recession are illustrated using grey colour shading. Sample size is set with each indicator´s own start date to 2011Q4.

When the CCI data starts, the economy comes straight from a recession. This means that it has its bottom point right from the start. CCI also eminently classifies the financial crisis. It was reaching a new low point midway through the crisis. For a more manageable comparison, Figure 9 in the Appendix provides a picture of all indicators plotted quarterly for the period 1992Q2-2011Q4 with Aastveit et al. (2016) recession dating.

The yield curve shows how the effective interest rate varies with maturity. Figure 3 shows that the long term interest rate is usually higher than the short term. The reason for this is that it will compensate for higher interest rate risk. The long term interest rate can also reflect the market´s expectation for economic growth. Hence, high long term rates signal belief in a bright economic future—low interest rates signal expectations of weak growth and the need for new interest rate cuts.

Three - and 10-years interest rate (1987M3-2011M12

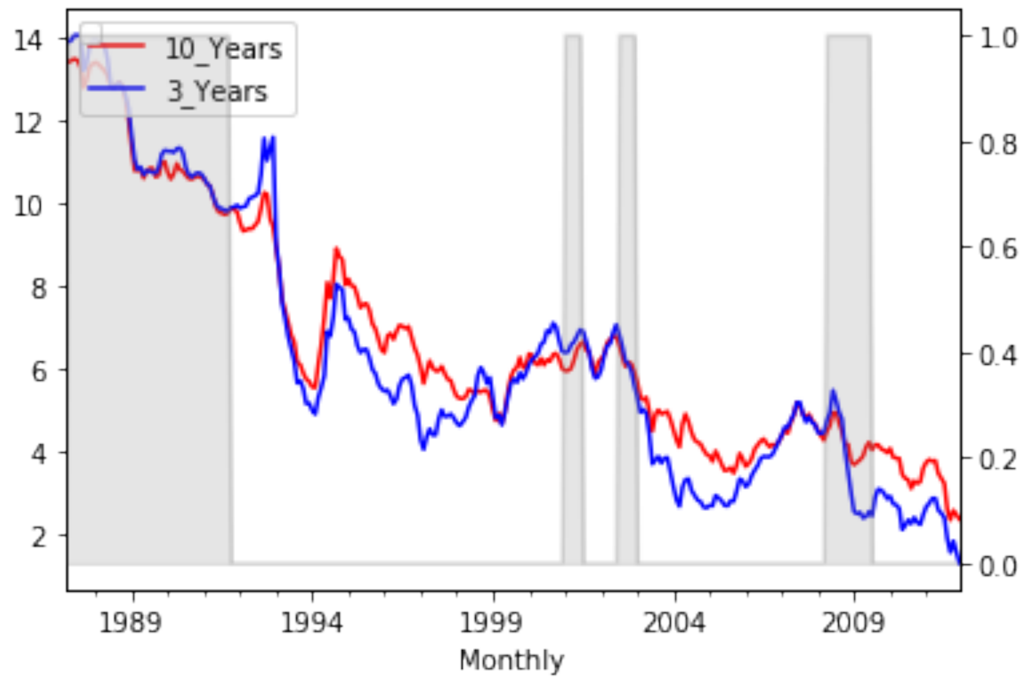


Figure 3: Interest rate 10- and 3-year monthly plotted with Aastveit et al. (2016) recession periods. Recession are illustrated using grey colour shading

Several of the indicators seem graphical to classify the financial crisis well. BTS, CCI and FNI have relatively similar curves throughout the data set and agree on both the highs and lows. The approach on rate spread and unemployment needs to be slightly different from the others. Unemployment expects to increase both during and after a crisis.

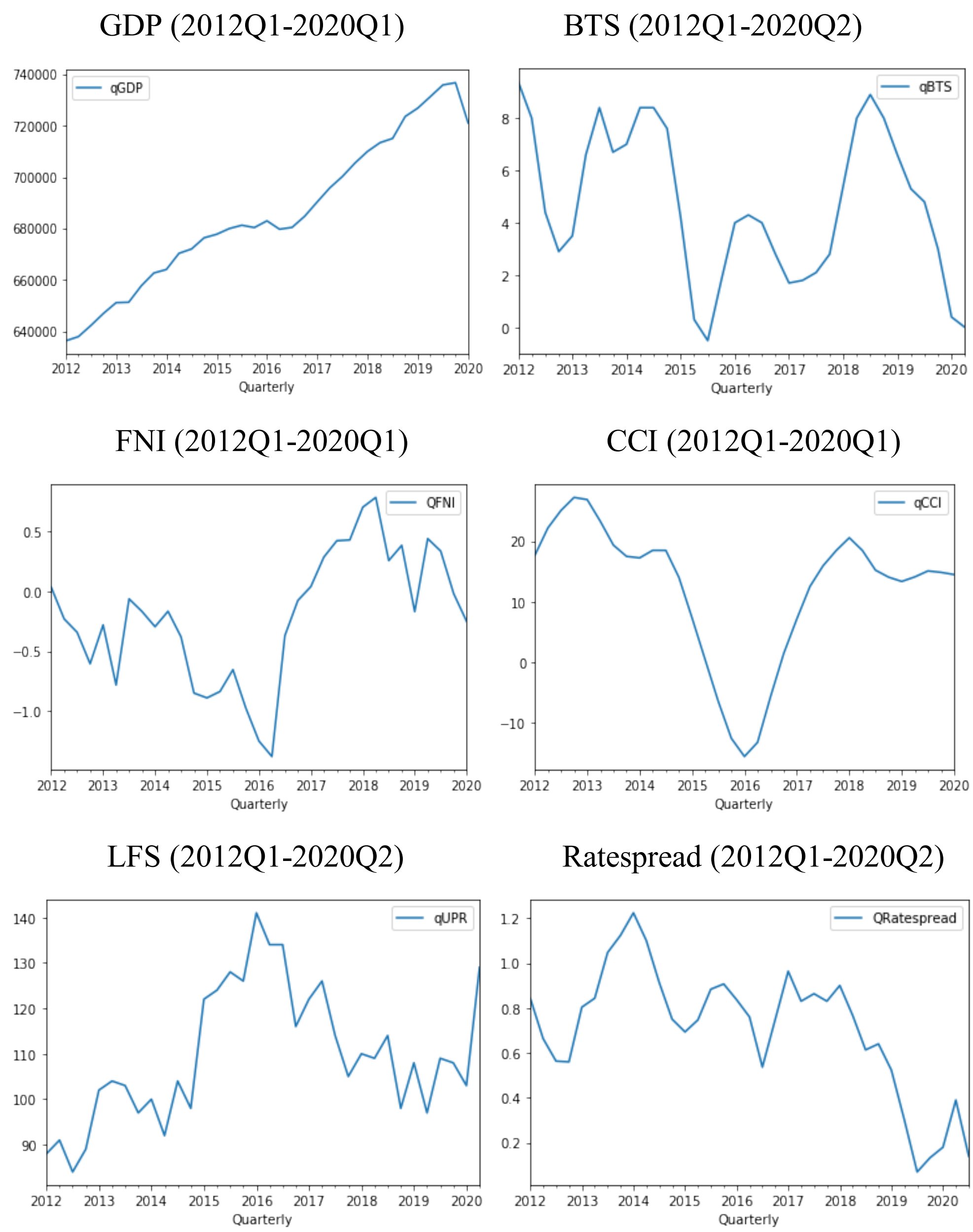


Figure 4: Every indicator and GDP quarterly plotted from 2012Q1 to 2020Q1/2020Q2.

To show the entrance and start of COVID-19, Figure 4 describes how the indicators appear to capture cyclical fluctuations in recent time. Most indicators appear to decrease, which may signal a coming recession. LFS starts to increase sharply at the turn of the year 2020, which confirms the result that many companies laid off large parts of the workforce. Note that LFS, BTS, and Rate Spread are plotted until 2020Q2 to bring out the fluctuations at the turn of the year. Figure 4 is quarterly, and Figure 11 and 12 in the Appendix provides a plot at a higher frequency.

4.2 IN-SAMPLE ANALYZES

The indicators are on different frequencies with varying start and end dates. CCI has the shortest dataset with start date 1992Q2. Due to varying frequency, all indicators have been set quarterly for the main analyzes. The data is made quarterly by assuming the months inside the quarter remains the same during the period. Since the business cycle dating of Aastveit et al. (2016) have an end date in 2011Q4, all indicators end 2011Q4 for our analyzes. To summarize, in the primary analyzes, the indicators are made quarterly and used in the period 1992Q2-2011Q4.

The reason of splitting the sample is to make an objective analyze of each indicator against each other.

For further analysis, all indicators have been estimated quarterly with their start date. Interest rate spread has data available from 1987M3 and estimated on both monthly and quarterly frequency for the entire period. FNI is originally on the daily frequency with start date 1990D1 and estimated at daily, monthly, and quarterly frequencies in addition to the estimation in the main analyzes.

The depended variable is a binary recession indicator which takes on the value of one during a recession and zeroes during an expansion. For the in-sample analyzes, our sample runs from 1978Q1-2011Q4 and covers a total of four recession, which ranges in duration from three to 19 quarters.

4.3 OUT-OF-SAMPLE ANALYZES

Since GDP is the change in volume from the previous quarter, seasonally adjusted (per cent), we aggregate the indicators on a quarterly frequency. CCI has the shortest dataset, and therefore, all the indicators have start date 1992Q2. We set all the indicators and GDP with the end date 2006Q2. Since indicators usually are leading, we assume that they were published a quarter before GDP. Thus, we know the value of the indicator at x_{t+h}, but we do not know GDP at y_{t+h}. We regress each indicator with GDP for the period 1992Q2-2006Q4. Then we use the estimated beta and alpha from the regression together with the actual value of the indicator to predict GDP for 2007Q1.

In the next round, we add one observation for actual GDP, so that we have GDP (dependent variable) for a period longer than the previous round. GDP can now be regressed with the indicator for a period longer with the same procedure. The regression gives a new estimated value for beta and alpha, so that we can predict GDP for one period longer (2007Q2).

We repeat this procedure until 2010Q4. Finally, we have estimated values for GDP for each quarter between 2007Q1-2010Q4 for each indicator. In this way, we can compare the actual numbers of GDP against the predicted.
Using MSE and RMSE, we can now evaluate the indicators ability to forecast GDP. To include observations before, during and after the financial crisis, we started prediction before and ended after the crisis. Hence, we can now analyze the behaviour of the indicators around the crisis.

We also want to assess information about the behaviour of the indicators around Covid-19 pandemic. In order to do that, we started predicting the indicators from 2018Q1 and ended 2020Q1. The same calculation predicts using the same procedure as for the financial crisis. Since the pandemic is not over, these results will only say something about before and the entrance of Covid-19.

5. RESULTS

We categorize overall economic activity in phases of expansion and contraction such as Travis and Jorda (2011). Then we make assessments of the indicators ability to classify such phases by using the Receiver Operating Characteristics (ROC) curves and area under the curve (AUROC) statistics. Then we evaluate each indicator´s ability to predict GDP, which is assessed in part two of this section. The prediction is made before, during and after the financial crisis. When it comes to Covid-19, it will be before and the beginning of the crisis.

5.1 IN-SAMPLE EVALUATION

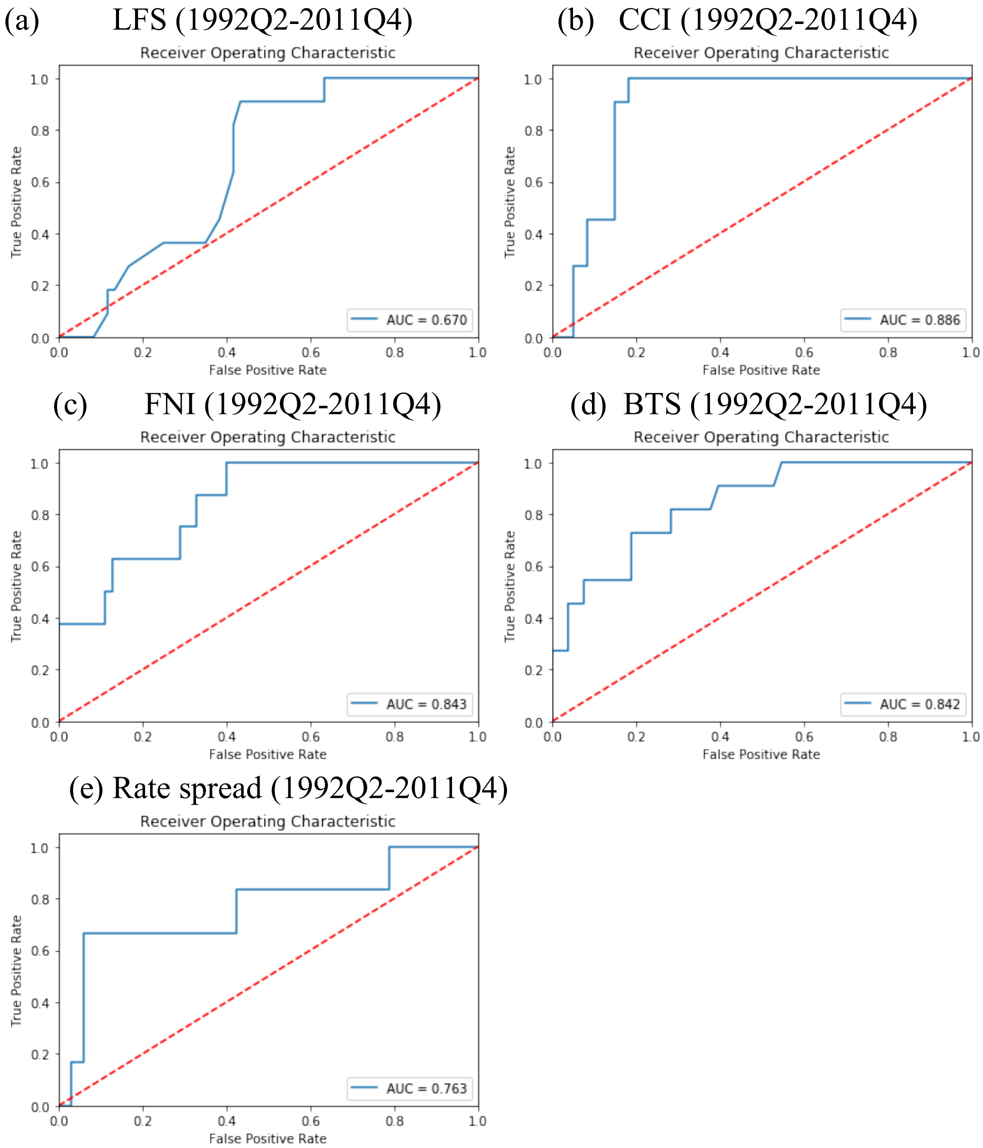


Figure 5: Receiver Operating Characteristics results for every indicator in the period 1992Q2-2011Q4. All indicators are aggregated at quarterly frequency.

In Norway, there is no official business cycle dating committee as it is in the United States (NBER). Because of this, we use the business cycle chronology of Aastveit et al. (2016), more specifically, MS-FMQ. With a focus on our contribution, we compare the performance of FNI with the other indicators to see who historically have been the best at classify recessions. We mainly focus on the AUROC statistic.

Figure 5 summarizes the in-sample classification result in the sample. By dividing the sample into a quarterly frequency with an equal start date (1992Q2), we see that FNI has an AUROC score of 0,843, which is a useful classification of the Norwegian business cycle compared to alternative indicators. CCI is the best indicator with an AUROC score on 0,886 and BTS is relatively even to FNI with an AUROC score on 0,842. These three indicators outperform both LFS and Rate Spread, which receive an AUROC score 0,670 and 0,763. From Figure 13 in the Appendix, where the AUROC score is estimated quarterly based on the indicators´ start date, FNI is more accurate than BTS. These have the same start date (1990Q1) and are a relevant comparison. The same applies to Rate Spread and LFS, where both have the same start date (1978Q1) and shows the Rate Spread has a much higher AUROC score than LFS.

(a) FNI (1990M1-2011M12) (b) Rate spread (1987M3-2011M12)

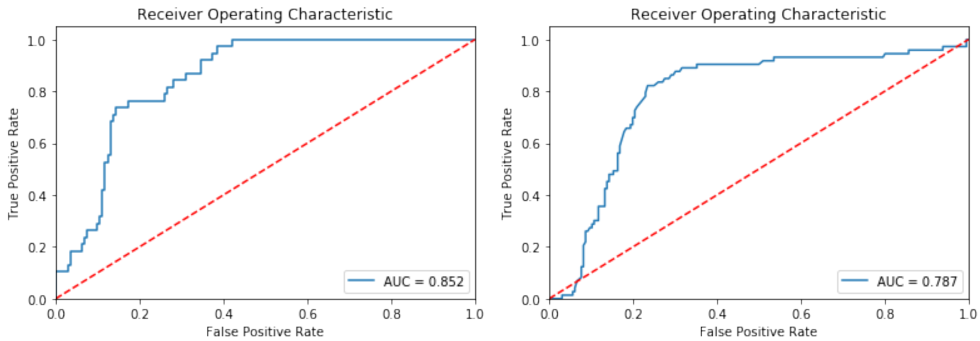


Figure 6: Receiver Operating Characteristics results for Financial News Index and the Rate Spread. From each indicator own start date to 2011Q4. Both indicators are aggregated to monthly frequency.

We want to know the state of the Norwegian economy continuously, high-frequent indicators will be more relevant. FNI and Rate Spread are both aggregated on a higher frequency. Figure 6 shows their classification ability at a monthly frequency and both manage to increase the AUROC score on a monthly basis. Important to notice that this is done with original start date, and Figure 14 in the Appendix shows that their performance decrease with start date 1992M4. Figure 15 in the Appendix shows that FNI, which is originally on a daily frequency, manages to maintain an AUROC score of over 0,8 on a daily frequency. This applies to both start 1992Q2 and 1990Q1 (original start date).

Table 1: All indicators AUROC scores on every possible frequency. Estimated at their own start date and from start date 1992Q2.

	Quarterly (1992Q2)	Quarterly (orginal start date)	Monthly (1992Q2)	Monthly (orignial start date)	Daily (1992Q2)	Daily (original start date)
FNI	0,843	0,845	0,798	0,852	0,804	0,829
Rate Spread	0,763	0,789	0,693	0,787		
CCI	0,886					
BTS	0,842	0,783				
LFS	0,670	0,639				

Table 1 summarizes the scores on different sample sizes and frequencies, and proves that FNI manages to obtain a fairly high score on different frequencies and sample sizes.

5.2 OUT-OF-SAMPLE EVALUATION

The relationship between every single indicator and GDP is evaluated through linear regression and used to predict the next quarter GDP. Several indicators are initially on different frequencies but made quarterly to make the estimates as fair as possible. To evaluate the accuracy of predicting GDP-growth, we use mean squared error and root mean square error.

All estimated values are from 2007Q1-2011Q4, and all indicators have been set to the same start date 1992Q2. The calculated value from the regression is applied to estimate the next quarter of GDP. We have multiplied this value with the following quarter result of the indicator, which was not involved in the regression.

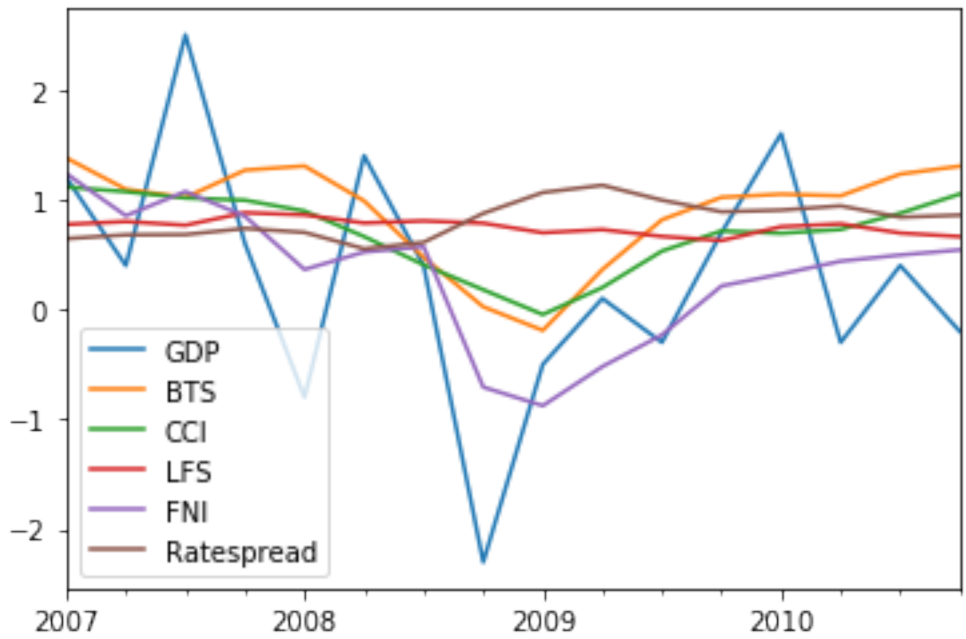


Figure 7: Predicted value of each indicator for the period 2007Q1 to 2010Q4 with actual GDP. All indicators are aggregated to quarterly frequency with start date 1992Q2.

From Figure 7, we can conclude that GDP has a more significant variation than the indicators. Only FNI has the same peak as GDP (third quarter 2007). Nevertheless, the increase in GDP is more extensive than what all the indicators predict. From Figure 7, we can also obtain that FNI is the best indicator to predict GDP, which both the MSE and RMSE results confirm (Table 2).

Table 2: Mean Squared Error (MSE) and Root Mean Square Error (RMSE) results for the predicted values of each indicator against GDP for the period 2007Q1 to 2010Q4. All indicators are aggregated to quarterly frequency with start date 1992Q2.

2007Q1-2010Q4	BTS	CCI	LFS	FNI	Rate spread
MSE	1,233	1,060	1,358	0,656	1,586
RMSE	1,110	1,030	1,165	0,810	1,259

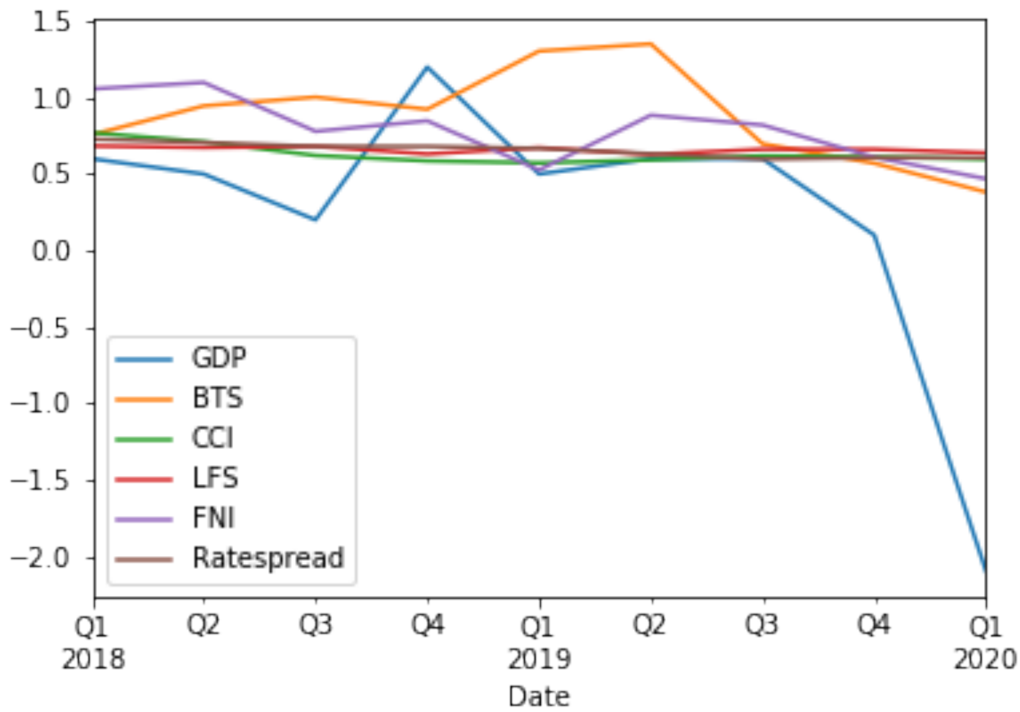


Figure 8: Predicted value of each indicator for the period 2018Q1 to 2020Q1 with actual GDP. All indicators are aggregated to quarterly frequency with start date 1992Q2.

From Figure 8 we can obtain how the indicators predicted GDP both before and at the beginning of Covid-19. We can conclude that the indicators perform reasonably well until the end of 2019. GDP falls significantly, and none of the indicators manages to predict the fall. Important to notice that the Covid-19 pandemic is a crisis the world has never seen before, and it is unfair to expect the indicators to predict the crisis in advance. The impact of Covid-19 also affects the MSE and RMSE scores.

FNI is the indicator that performs the best also here, although it is much closer than the previous crisis. Hence, FNI is the best indicator to predict GDP.

Table 3: Mean Squared Error (MSE) and Root Mean Square Error (RMSE) results for the predicted values of each indicator against GDP for the period 2018Q1 to 2020Q1. All indicators are aggregated to quarterly frequency with start date 1992Q2.

2018Q1-2020Q1	BTS	CCI	LFS	FNI	Ratespread
MSE	0,950	0,908	0,938	0,891	0,908
RMSE	0,975	0,953	0,969	0,944	0,953

6. CONCLUSION

We wanted to check previous Norwegian economic crises and whether it was possible to find early signals for these business cycle fluctuations. Economic recessions are costly events for the nation as a whole and therefore crucial to get early information to make countercyclical measures and dampen the negative effect of the crisis. Lack of data makes it very difficult to make the right decisions at an early stage. One of the most important indicators for the entire economy is GDP and this is published quarterly with a considerable lag, usually many months. Our contribution is to use FNI, which is a high-frequency indicator, to analyze its performance against indicators that are considered to be accurate indicators of the economy. There have been several recessions in Norway over the years, and by looking at different indicators, we have evaluated which of these have historically been best a classify business cycles and predict GDP during both Financial crisis and COVID-19. We have used in-sample and out-of-sample estimates to evaluate the accuracy and classification of each indicator.

In-sample results shows that CCI is the most accurate indicator, because of a small sample-size and quarterly frequency the indicator is difficult to compare with the others. FNI scores high on every sample-size and frequency and is thus a credible source for classifying previous recessions. Out-of-sample forecast have been analyzed with a particular focus on the financial crisis and also analyzed for the COVID-19 entrance. FNI outperforms the other for the financial crisis and is also best for the COVID-19 entrance. Hence, FNI is a reliable indicator of the change in GPD around crises.

Our thesis compares five different indicators, and a potential weakness is that there are several good leading indicators at a higher frequency than quarterly. For this thesis, it would have been more relevant and analyzed ten years-three months Yield spread which is historically a good indicator. There are numerous methods for forecasting the indicators with GDP; some of the models are more advanced and can provide more precise analyzes.

7. REFERENCES

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. Journal of Business and Economic Statistics 32 (1), 48–68.

Aastveit, K. A., A. S. Jore, and F. Ravazzolo (2016). Identification and real-time fore- casting of Norwegian business cycles. International Journal of Forecasting 32(2), 283 – 292.

Aastveit, K. A., Anundsen, A. K., and Herstad, E. I (2017). Residential investment and recession predictability. CAMP Working Paper Series No 8/2017.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data flow. Working Paper Series 1564, European Central Bank

Berge, T. J. and Jord`a, O (2011). Evaluating the Classification of Economic Activity into Recessions and Expansion. Department of Economics.

Claveria, O., E. Pons, and R. Ramos (2007). Business and consumer expectations and macroeconomic forecasts. International Journal of Forecasting 23 (1), 47–69.

Dées, S. and P., S. Brinca (2011). Consumer Confidence as a Predictor of Consumption Spending. Evidence For The United States and The Euro Area. Working Paper Series 1349, European Central Bank.

Estrella, A., & Mishkin, F. S. (1997). The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. European economic review, 41(7), 1375-1401.

Estrella, A. and F. S. Mishkin (1998). Predicting U.S. Recessions: Financial Variables As Leading Indicators. Review of Economics and Statistics 80 (1), 45–61.

Evans, M. D. (2005). Where are we now? real-time estimates of the macro economy. International Journal of Central Banking 1 (2), 127–175.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics 55 (4), 665–676.

Hall, Robert, Martin Feldstein, Jeffrey Frankel, Robert Gordon, Christina Romer, David Romer, and Victor Zarnowitz (2003). The NBER's Recession Dating Procedure. http://www.nber.org/cycles/recessions.pdf.

Hamilton, J. D. (1983). Oil and the Macroeconomy since World War II. Journal of Political Economy 91 (2), 228–248.

Hamilton, J. D. (1996). This is what happened to the oil price-macroeconomy relationship. Journal of Monetary Economics 38 (2), 215–220.

Hansson, J., P. Jannson, and M. Lof. 2003. Business Survey Data: Do They Help in Forecasting the Macro Economy? Sverigees Riksbank Working Paper No. 151

Howrey, E. Philip. 2001. "The Predictive Power of the Index of Consumer Sentiment." Brookings Papers on Economic Activity. 1, pp. 175–216.

Kelley, D (2019). Which Leading Indicators Have Done Better At Signaling Past Recessions? Chicago Fed Letter, No. 425, 2019.

Liu, W. and E. Moench (2014). What Predicts U.S. Recessions? Federal Reserve Bank of New York Staff Reports, no.691. Federal Reserve Bank of New York Staff Reports.

Mazurek, J., Mielcová, I. (2017). Is consumer confidence index a suitable predictor of future economic growth? an evidence from the USA, E & M Ekonomie a Management 20, P. 30-45.

Norges Bank. (2020). Rentestatistikk. Retrieved from https://www.norgesbank.no/Statistikk/Rentestatistikk/

Petersen, M., A. (2004). Information: Hard and Soft. Kellogg School of Management Northwestern University and NBER).

Ravazzolo, F. and P. Rothman (2013). Oil and U.S. GDP: A Real-Time Out-of-Sample Examination. Journal of Money, Credit and Banking 45 (2-3), 449–463.

Ravazzolo, F. and P. Rothman (2016). Oil-price density forecasts of US GDP. Studies in Nonlinear Dynamics & Econometrics 20 (4), 441–453.

Sparrman, V. (2012). Arbeidsledighet som konjunkturindikator og forklaringsfaktor i makromodeller, (in Norwegain. Unemployment as economic indicator and explanatory variable in macroeconomic models), Økonomiske analyser 5, 21–25.

SSB. (2020). StatBank Norway. Collected from: https://www.ssb.no/en/statbank/

SSB (2020). Labour force survey. Table: 08518: Unemployed persons, by age and sex 1972K1 - 2020K2

SSB (2020). Business tendency survey for manufacturing, mining and quarrying. Table: 08267: Business tendency survey. Composite indicators 1990K1 - 2020K2

SSB (2020). National accounts. Table: 09190: Final expenditure and gross domestic product. Unadjusted and seasonally adjusted figures 1978K1 - 2020K1

Stock J. H. and M. W. Watson (1988). A PROBABILITY MODEL OF THE COINCIDENT ECONOMIC INDICATORS. Working Paper No. 2772, National Bureau Of Economic Research.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In NBER Macroeconomics Annual 1989, Volume 4, NBER Chapters, pp. 351–409. National Bureau of Economic Research, Inc

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. Journal of Economic Literature 41 (3), 788–829.

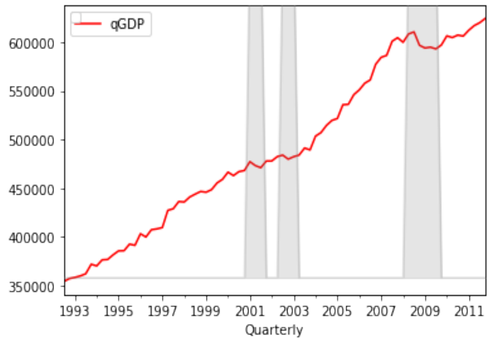
Thorsrud, L. A. (2018). Words are the new numbers: A newsy coincident index of the business cycle. Journal of Business & Economic Statistics (Forthcoming).

Travis, J. B. and s. Jordà (2011). Evaluating the Classification of Economic Activity into Recessions and Expansions. American Economic Journal: Macroeconomics 3(2), 246–77.

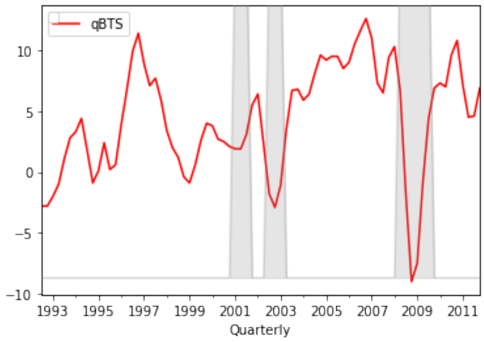
8. APPENDIX

8.1 RECESSION PLOTS

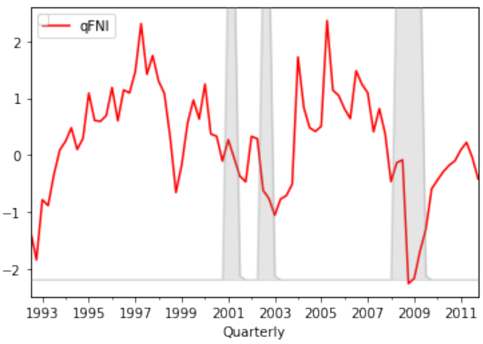
(a) GDP (1992Q2-2011Q4)



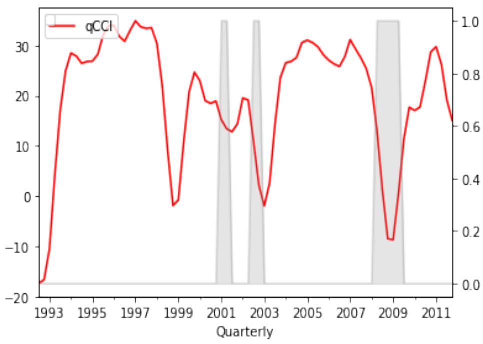
(b) BTS (1992Q2-2011Q4)



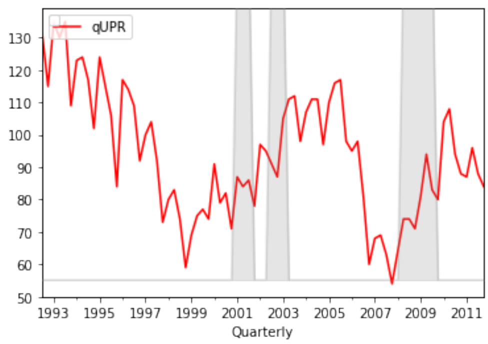
(c) FNI (1992Q2-2011Q4)



(d) CCI (1992Q2-2011Q4)



(e) LFS (1992Q2-2011Q4)



(f) Rate spread (1992Q2-2011Q4)

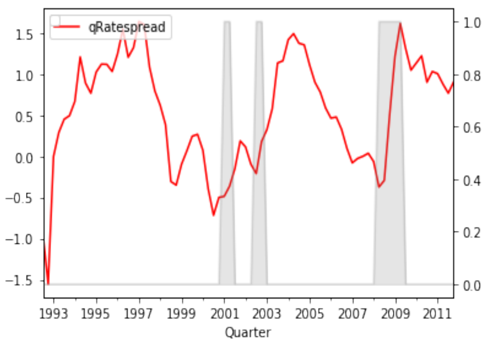


Figure 9: Every indicator and GDP quarterly plotted with Aastveit et al. (2016) recession periods. Recession are illustrated using grey colour shading. Sample sizes is equal for each variable (1992Q2-2011Q4)

FNI (1978D1-2011D365)

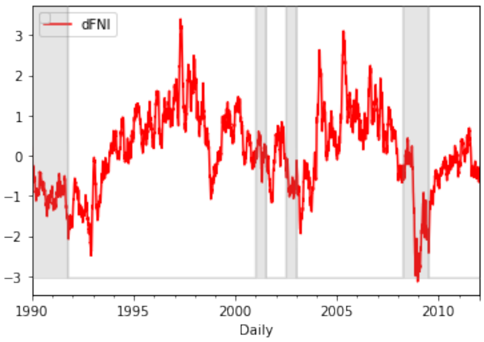


Figure 10: FNI daily plotted with Aastveit et al. (2016) recession periods. Recession are illustrated using grey colour shading. Sample size is set to 1990D1 to 2011M12.

8.2 COVID-19 CRISIS PLOTS

(a) FNI (2012M1-2020M2) (b) Rate spread (2012M1-2020M7)

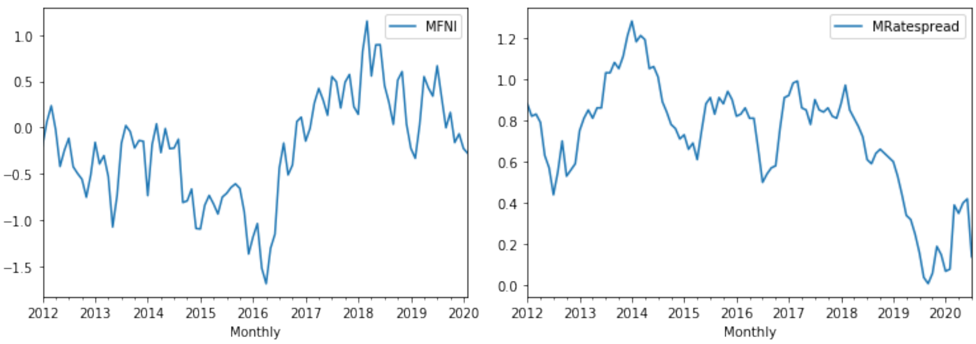


Figure 11: FNI and Rate Spread monthly plotted from 2012M1 to 2020M2/2020M7.

FNI (2012/01/01-2020/02/29)

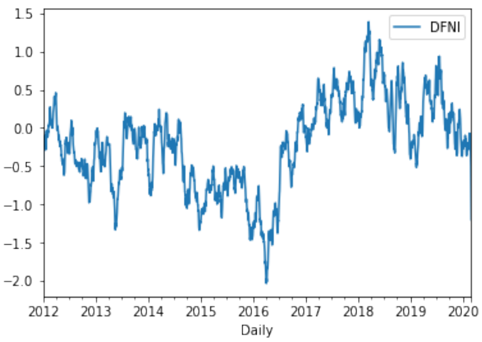


Figure 12: FNI daily plotted from 2012/01/01 to 2020/02/29.

8.3 OUT-OF-SAMPLE EVIDENCE

Table 4: Predicted values for each indicator and GDP-growth for the period 2007Q1-2010Q4. All indicators are aggregated to quarterly with start data 1992Q2.

Date	GDP	BTS	CCI	LFS	FNI	Ratespread
2007Q1	1,20	1,38	1,11	0,77	1,24	0,64
2007Q2	0,40	1,09	1,07	0,80	0,85	0,68
2007Q3	2,50	1,02	1,01	0,77	1,07	0,68
2007Q4	0,60	1,27	0,99	0,88	0,85	0,74
2008Q1	-0,80	1,31	0,90	0,86	0,36	0,70
2008Q2	1,40	0,99	0,66	0,79	0,52	0,55
2008Q3	0,40	0,47	0,41	0,81	0,57	0,61
2008Q4	-2,30	0,02	0,18	0,79	-0,71	0,87
2009Q1	-0,50	-0,19	-0,05	0,70	-0,88	1,06
2009Q2	0,10	0,36	0,20	0,73	-0,52	1,13
2009Q3	-0,30	0,82	0,53	0,66	-0,23	0,99
2009Q4	0,70	1,02	0,71	0,63	0,21	0,89
2010Q1	1,60	1,05	0,69	0,75	0,32	0,90
2010Q2	-0,30	1,03	0,73	0,78	0,44	0,94
2010Q3	0,40	1,23	0,88	0,70	0,50	0,84
2010Q4	-0,20	1,30	1,05	0,66	0,54	0,86

Table 5: Predicted values for each indicator and GDP-growth for the period 2018Q1-2020Q1. All indicators are aggregated to quarterly with start data 1992Q2.

Date	GDP	BTS	CCI	LFS	FNI	Ratespread
2018Q1	0,60	0,76	0,77	0,68	1,06	0,73
2018Q2	0,50	0,95	0,71	0,67	1,10	0,71
2018Q3	0,20	1,00	0,62	0,68	0,78	0,68
2018Q4	1,20	0,92	0,59	0,63	0,85	0,68
2019Q1	0,50	1,30	0,57	0,67	0,53	0,67
2019Q2	0,60	1,35	0,59	0,63	0,88	0,63
2019Q3	0,60	0,70	0,62	0,66	0,82	0,60
2019Q4	0,10	0,57	0,61	0,66	0,61	0,61
2020Q1	-2,10	0,38	0,60	0,64	0,47	0,61

Table 6: Recession dating by Aastveit et al. (2016).

Dates	Peak/Trough	MS-FMQ
1986-1989	Peak	1987:Q2
1990-1994	Trough	1991:Q4
1995-2001	Peak	2001:Q1
	Trough	2001:Q3
2002-2003	Peak	2002:Q3
	Trough	2003:Q1
2004-2010	Peak	2008Q2
	Trough	2009Q3

8.4 IN-SAMPLE EVIDENCE

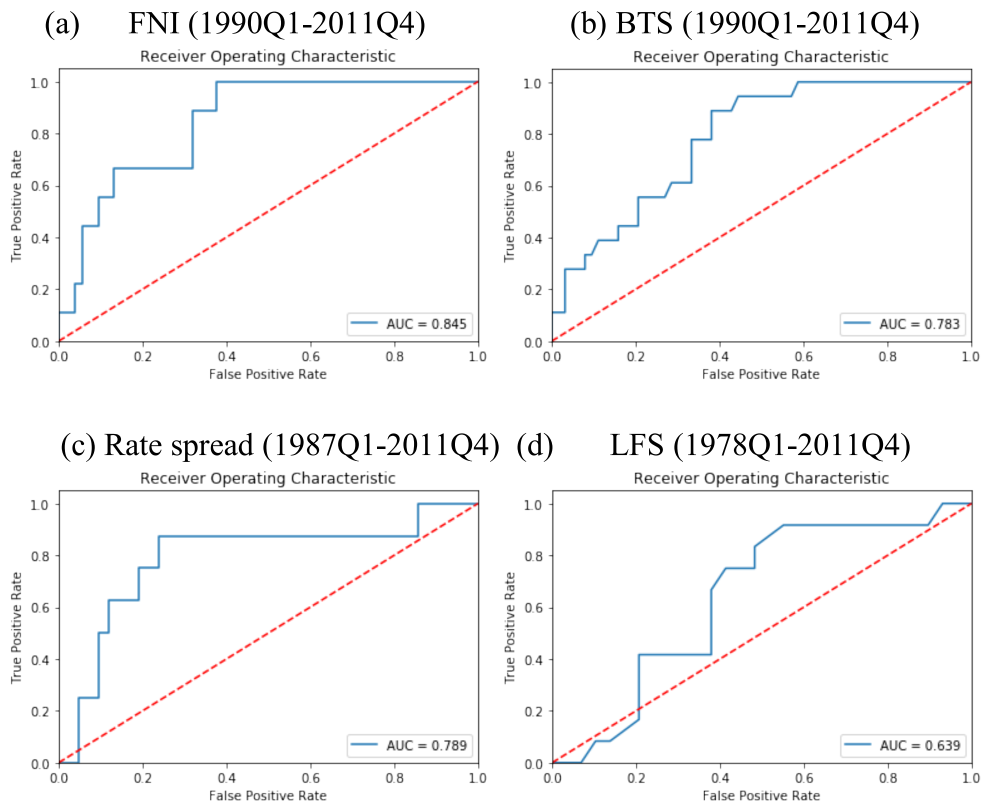


Figure 13: Receiver Operating Characteristics results for Financial News Index, the Business Tendency Survey, Rate Spread and Unemployment Rate. From each indicator own start date to 2011Q4. All indicators are aggregated to quarterly frequency

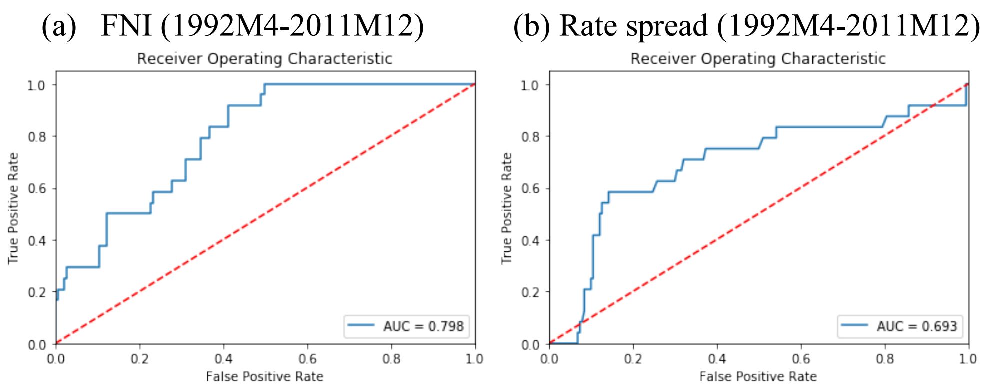


Figure 14: Receiver Operating Characteristics results for Financial News Index and Rate Spread. From 1992Q2-2011Q4, both aggregated at monthly frequency.

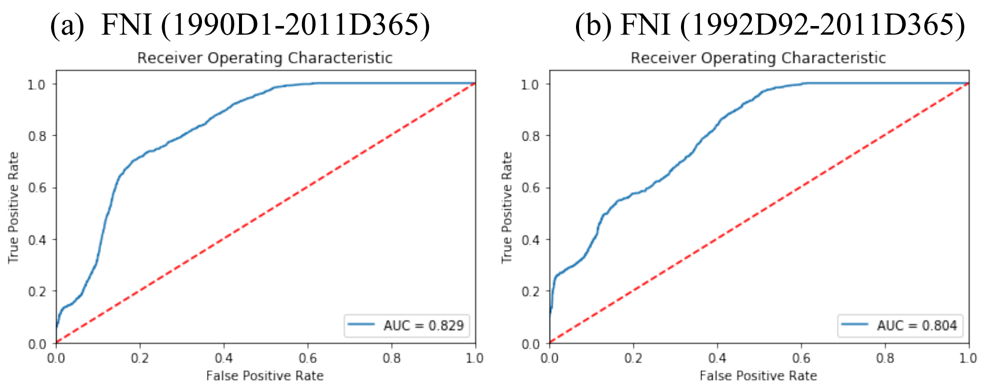


Figure 15: Receiver Operating Characteristics results for Financial News Index. From its own start date and 1992Q2 to 2011Q4. Indicator estimated with daily frequency.

8.5 PYTHON CODES

For the plot of each indicator with recession dating from Aastveit et al. (2016). Using LFS as an example for the code:

```
import stats_to_pandas as stp
import pandas as pd
ALH = stp.read_all(table_id = '08518')
ALH = ALH[ALH['sex']=='Both sexes']
ALH = ALH[ALH['age']=='15-74 years']
ALH = ALH[ALH['contents']=='Unemployed (1 000 persons)']
ALH['quarter'] = ALH['quarter'].str.replace('K','Q')
ALH = ALH.set_index('quarter')
ALH.index = pd.to_datetime(ALH.index)
ALH = ALH.rename(columns={'value':'qUPR'})
ALH = pd.DataFrame(ALH.qUPR)
ALH.head()
En = \
            ((ALH.index >= '1987-04-01') &
             (ALH.index <= '1991-10-01')).astype(int)
To = \
            ((ALH.index >= '2001-01-01') &
             (ALH.index <= '2001-07-01')).astype(int)
Tre = \
            ((ALH.index >= '2002-07-01') &
             (ALH.index <= '2003-01-01')).astype(int)
Fire = \
            ((ALH.index >= '2008-04-01') &
             (ALH.index <= '2009-07-01')).astype(int)


ALH['Recession'] = En + To + Tre + Fire
```

```
ALH.tail()
ALH = ALH[24:]
ALH = ALH[:-34]
ALH.index.names = ['Quarterly']
recs = ALH.query('Recession==1')
recs_2k = recs.ix['1987-04-01':'1991-10-01']
recs_2k8 = recs.ix['2001-01-01':'2001-07-01']
recs_2k9 = recs.ix['2002-07-01':'2003-01-01']
recs_2k10 = recs.ix['2008-04-01':'2009-07-01']
recs2k_bgn = recs_2k.index[0]
recs2k_end = recs_2k.index[-1]

recs2k8_bgn = recs_2k8.index[0]
recs2k8_end = recs_2k8.index[-1]

recs2k9_bgn = recs_2k9.index[0]
recs2k9_end = recs_2k9.index[-1]

recs2k10_bgn = recs_2k10.index[0]
recs2k10_end = recs_2k10.index[-1]
import matplotlib.pyplot as plt
def plot_var(y1):
    fig0, ax0 = plt.subplots()
    ax1 = ax0.twinx()

    y1.plot(kind='line', stacked=False, ax=ax0, color='red')
    ALH['Recession'].plot(kind='area', secondary_y=True, ax=ax1, alpha=.2,
color='grey')
    ax0.legend(loc='upper left')
    ax1.legend(loc='upper left')
    plt.ylim(ymax=0.8)
    plt.axis('off')
    plt.xlabel('Date')
    plt.show()
    plt.close()
plot_var(ALH['qUPR']
```

For the in-sample evaluation. Using CCI as an example:

```
from pandas import read_csv
from pandas import datetime
from pandas import DataFrame
from statsmodels.tsa.arima_model import ARIMA
from matplotlib import pyplot
import pandas as pd
# load dataset
df = pd.read_excel("./data/CCI.xlsx",index_col=0,parse_dates=True)
df.to_csv("./data/CCI.csv")
df.dropna(inplace=True)
En = \
            ((df.index >= '1987-04-01') &
             (df.index <= '1991-10-01')).astype(int)
To = \
            ((df.index >= '2001-01-01') &
             (df.index <= '2001-07-01')).astype(int)
Tre = \
            ((df.index >= '2002-07-01') &
             (df.index <= '2003-01-01')).astype(int)
Fire = \
            ((df.index >= '2008-04-01') &
             (df.index <= '2009-07-01')).astype(int)
df['Recession'] = En + To + Tre + Fire
df.head()
df = df[:-33]
X = pd.DataFrame(df.iloc[:,[0]])
```

```python
y = pd.DataFrame(df.iloc[:,[2]])

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.9,random_state=1)

from sklearn.linear_model import LogisticRegression
logmodel = LogisticRegression()
print(logmodel.fit(X_train,y_train))
y_pred = logmodel.predict(X_test)
from sklearn.metrics import confusion_matrix
confusion_matrix = confusion_matrix(y_test,y_pred)
print(confusion_matrix)
from sklearn.metrics import classification_report
print(classification_report(y_test,y_pred))
from sklearn.metrics import roc_auc_score
from sklearn.metrics import roc_curve

fpr, tpr, thresholds = roc_curve(y_test,logmodel.predict_proba(X_test)[:,1])
AUC = roc_auc_score(y_test,logmodel.predict_proba(X_test)[:,1])
import matplotlib.pyplot as plt
plt.figure()
plt.plot(fpr, tpr, label='AUC = %0.3f' % AUC)
plt.plot([0,1], [0,1], 'r--')
plt.xlim([0.0,1.0])
plt.ylim([0.0,1.05])
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic')
plt.legend(loc="lower right")
plt.show()
```

For out-of-sample evaluation. Using BTS as an example (2018Q1-2020Q1):

```python
import stats_to_pandas as stp
import pandas as pd
KJB = stp.read_all(table_id = '08267')
KJB = KJB[KJB['type of adjustment']=='Smoothed seasonally adjusted']
KJB = KJB[KJB['contents']=='Confidence Indicator']
KJB = KJB[KJB['industry/main industrial grouping']=='Consumer goods (total)']
KJB['quarter'] = KJB['quarter'].str.replace('K','Q')
KJB = KJB.set_index('quarter')
KJB.index = pd.to_datetime(KJB.index)
KJB = KJB.rename(columns={'value':'qKJB'})
KJB = pd.DataFrame(KJB.qKJB)
KJB.head()
import stats_to_pandas as stp
import pandas as pd
GDP = stp.read_all(table_id = '09190')
GDP = GDP[GDP['contents']=='Change in volume from the previous quarter, seasonally adjusted (per cent)']
GDP = GDP[GDP['macroeconomic indicator']=='Gross domestic product Mainland Norway, market values']
GDP['quarter'] = GDP['quarter'].str.replace('K','Q')
GDP = GDP.set_index('quarter')
GDP.index = pd.to_datetime(GDP.index)
GDP = GDP.rename(columns={'value':'qGDP'})
Q_GDP = pd.DataFrame(GDP.qGDP)
Q_GDP.head()
Q_GDP = Q_GDP[58:]
KJB = KJB[10:]
KJB = KJB[:-1]
Q_GDP['KJB'] = pd.DataFrame(KJB.iloc[:,[0]])
```

```
Q_GDP.head()
Q42017 = Q_GDP[:-9]
Q42017.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model = LinearRegression()

X = pd.DataFrame(Q42017.iloc[:,[1]])
y = pd.DataFrame(Q42017.iloc[:,[0]])
model.fit(X, y)
model = LinearRegression().fit(X, y)

print('intercept:', model.intercept_)
intercept: 5.633333333333329
print('slope:', model.coef_)
Q_GDP[:-8].tail()

model.intercept_+(model.coef_*5.4)
Q12018 = Q_GDP[:-8]
Q12018.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model1 = LinearRegression()

X1 = pd.DataFrame(Q12018.iloc[:,[1]])
y1 = pd.DataFrame(Q12018.iloc[:,[0]])
model1.fit(X1, y1)
model1 = LinearRegression().fit(X1, y1)

print('intercept:', model1.intercept_)
intercept: 5.633333333333329
print('slope:', model1.coef_)

model1.intercept_+(model1.coef_*8.0)
Q_GDP[:-7].tail()
Q22018 = Q_GDP[:-7]
Q22018.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model2 = LinearRegression()

X2 = pd.DataFrame(Q22018.iloc[:,[1]])
y2 = pd.DataFrame(Q22018.iloc[:,[0]])
model2.fit(X2, y2)
model2 = LinearRegression().fit(X2, y2)

print('intercept:', model2.intercept_)
intercept: 5.633333333333329
print('slope:', model2.coef_)
Q_GDP[:-6].tail()

model2.intercept_+model2.coef_*8.9
Q32018 = Q_GDP[:-6]
Q32018.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model3 = LinearRegression()

X3 = pd.DataFrame(Q32018.iloc[:,[1]])
y3 = pd.DataFrame(Q32018.iloc[:,[0]])
model3.fit(X3, y3)
model3 = LinearRegression().fit(X3, y3)

print('intercept:', model3.intercept_)
intercept: 5.633333333333329
```

```
print('slope:', model3.coef_)
Q_GDP[:-5].tail()
model3.intercept_+model3.coef_*8.0
Q42018 = Q_GDP[:-5]
Q42018.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model4 = LinearRegression()

X4 = pd.DataFrame(Q42018.iloc[:,[1]])
y4 = pd.DataFrame(Q42018.iloc[:,[0]])
model4.fit(X4, y4)
model4 = LinearRegression().fit(X4, y4)

print('intercept:', model4.intercept_)
intercept: 5.633333333333329
print('slope:', model4.coef_)
Q_GDP[:-4].tail()

model4.intercept_+model4.coef_*13.32517
Q12019 = Q_GDP[:-4]
Q12019.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model5 = LinearRegression()

X5 = pd.DataFrame(Q12019.iloc[:,[1]])
y5 = pd.DataFrame(Q12019.iloc[:,[0]])
model5.fit(X5, y5)
model5 = LinearRegression().fit(X5, y5)

print('intercept:', model5.intercept_)
intercept: 5.633333333333329
print('slope:', model5.coef_)
Q_GDP[:-3].tail()

model5.intercept_+model5.coef_*14.08179
Q22019 = Q_GDP[:-3]
Q22019.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model6 = LinearRegression()

X6 = pd.DataFrame(Q22019.iloc[:,[1]])
y6 = pd.DataFrame(Q22019.iloc[:,[0]])
model6.fit(X6, y6)
model6 = LinearRegression().fit(X6, y6)

print('intercept:', model6.intercept_)
intercept: 5.633333333333329
print('slope:', model6.coef_)
Q_GDP[:-2].tail()

model6.intercept_+model6.coef_*4.8
Q32019 = Q_GDP[:-2]
Q32019.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model7 = LinearRegression()

X7 = pd.DataFrame(Q32019.iloc[:,[1]])
y7 = pd.DataFrame(Q32019.iloc[:,[0]])
model7.fit(X7, y7)
model7 = LinearRegression().fit(X7, y7)

```

```
print('intercept:', model7.intercept_)
intercept: 5.633333333333329
print('slope:', model7.coef_)
Q_GDP[:-1].tail()

model7.intercept_+model7.coef_*3.0
Q42019 = Q_GDP[:-1]
Q42019.tail()

import numpy as np
from sklearn.linear_model import LinearRegression
model8 = LinearRegression()

X8 = pd.DataFrame(Q42019.iloc[:,[1]])
y8 = pd.DataFrame(Q42019.iloc[:,[0]])
model8.fit(X8, y8)
model8 = LinearRegression().fit(X8, y8)

print('intercept:', model8.intercept_)
intercept: 5.633333333333329
print('slope:', model8.coef_)
Q_GDP.tail()

model8.intercept_+model8.coef_*0.4
```

LIST OF FIGURES

LIST OF TABLES