A Financial Conditions Index for Norway: Can financial indicators predict GDP?

BY AKSEL SAKAROĞLU ØRBECK AND MARIANN TORVANGER

NORWEGIAN UNIVERSITY OF LIFE SCIENCES
Department of Economics and business
Master Thesis 30 credits 2011

PREFACE

This master thesis completes of our masters program in business administration at the UMB School of Economics and Business.

We would like to extend our sincerest gratitude to our supervisor Ole Gjølberg at the University of Life Sciences for all help during the process of writing this thesis.

Finally, we would also like to thank Senior Economist Bjørn-Roger Wilhemsen at First Securities, Lasse Holboell Nielsen at Goldman Sachs and OECD for providing us with data material and forecasts as benchmarks.

Oslo, 15.05.2011

....

Aksel S. Ørbeck Mariann Torvanger

Summary

In this thesis, we used financial indicators to construct a Financial Conditions Index (FCI) aimed at predicting Norwegian GDP.

Our analysis started out by surveying previous work on FCIs and leading indicators. The majority of existing FCIs include some measure of interest and exchange rates, asset prices and risk premiums. We followed this consensus, and constructed two sets of single equation log-log regression models with up to four lags; one for each of the five single indicators proven to have leading characteristics in previous literature, and one for each FCI using an equally weighted sum approach. We then calculated sub-indices for each financial indicator and added the sub-indices together, resulting in five FCIs.

Next, we conducted several out-of-sample predictions of the period 2006(1)-2010(4) based on estimated weights from the basis period 1980(2)-2005(4). To test our FCI's forecasting power we examined some alternative forecasts as benchmarks, five single indicators, a naïve model and FCI predictions by OECD and Goldman Sachs. To compare the different prediction models' preciseness, we chose the RMSE and MAPE measures.

The results are contradictory, and dependent on whether RMSE or MAPE is the criterion of selection. However, neither of the FCIs, nor any of the other single indicators or benchmarks was able to provide consistently superior estimates. None of the models that provide the lowest RMSE or MAPE values are statistically significant, and no model is superior in both criteria. The best FCI model based on RMSE is our static FCI, but chosen on MAPE the best model is the FCI with one lag. We emphasize the RMSE criterion due to the assumptions underlying OLS estimation, and recommend the static FCI. Further improvement suggestions and future research potential are treated at the end.

Table of Contents

Overview of Graphs, Tables and Figures

Graph overview

Table overview

Figure overview

1 Introduction

Since 2007, financial stress has contributed to a downturn in the world economy. The latest financial crisis demonstrated how important financial conditions are for real economic growth and showed us that the predictive power from past indicators has been limited. Economic researchers expressed a need to develop a broader measure to capture changes in economic growth, and consequently after the financial crisis several new Financial Condition Indices (FCIs) were constructed.

Even though the banking system in Norway is probably more robust than other countries', the Norwegian economy was also affected by the international financial crisis. The Gross Domestic Product is commonly used to measure the effects and consequences of such a crisis. Fluctuations in GDP are generally interpreted as a measure of a country's future wealth and development, and indicators that are able to signal or measure fluctuations in consumption, real capital investments, and the trade balance at an early stage, is therefore valuable. The theory of economic growth, examines how increases or decreases in the GDP depends on these indicators, their importance and underlying sources driving economic growth. Studying the implications and characteristics of economic growth is therefore important for the understanding of macroeconomic relationships.

In order to capture and predict movements in the different constituents that is GDP, several methods have been employed in past. Single financial indicators were used for many years, followed by the construction of Monetary Conditions Indices (MCIs), but GDP is a too complex and multi-faceted a size to be captured by these measures alone. None of the existing models currently in use seem to have been able to signal the real extent of the previous financial crisis in Norway. Therefore we believe that constructing an index for Norway, which consists of several single monetary and financial indicators, would perform better in terms of GDP forecasts.

To our knowledge no attempt at creating an FCI for Norway has been conducted. In this paper we therefore seek to construct an FCI with Norwegian data, based on research on leading financial indicators in international FCIs, to capture and predict quarterly Norwegian GDP fluctuations between 2006(1) and 2010(4). To examine the FCI's predictive power we analyze five single indicators in addition to two previous FCIs and a naïve model, and compare our FCI to these benchmarks to measure forecast performance.

We hope this thesis can serve as a guideline and contribute positively to those successors who seek to conduct further research on the drivers of economic growth in Norway. In addition, we hope this thesis can be of interest to actors in the market, such as policy makers, investment firms, private investors and banks. Our aim at the outset of this thesis is therefore that the FCI we develop is easily employed and updated, and can help those in need of GDP predictions to predict GDP themselves, without an MBA in finance. We have therefore chosen to keep the technical estimation of the FCI relatively simple, and only included variables that are readily and freely available to the general public. Anyone interested in and/or dependent on the variations in GDP might find this FCI a useful tool in their total predictions of the immediate future. An investment bank for example, might have investment opportunities whose profit is dependent on GDP fluctuations, in which case reliable predictions of GDP can be a valuable asset, and aid in investment decisions.

We start by specifying and presenting our thesis in chapter 1. In chapter 2 we review previous attempts to construct FCIs, mostly for the U.S. economy. We examine estimation approaches, indicators that have been included and how well these FCIs have predicted GDP measures. In chapter 3 we first study previous research within the field of leading indicators and GDP growth. We start off with the international research and proceed with resembling Norwegian literature. Secondly, we present possible connections between GDP and financial indicators in Norway to define a set of indicators that might be reasonable to include in our FCI.

In chapter 4 we present the model setup for the five single indicators and our FCI which includes eight variables spanning back to 1980. We employ a VAR model and start by constructing a static model which we expand to a dynamic model to include four lagged values of the various indicators. In the FCI we also include up to four lagged values of GDP changes. In total we estimate five models for each single indicator and five FCIs, and we employ the same out-of-sample analysis for all models. We estimate the weight attached to each indicator based on changes in the basis-period 1980(2)-2005(4), and hold them constant throughout the prediction period. For each model we multiply the weight with its respective quarterly value, into a (sub-) index, which in the case of the single indicator models represents the actual predictions. Correspondingly, for the FCIs we summarize the sub-indices, which then become the prediction values of GDP changes.

Furthermore, we conduct several tests to detect potential violations of the OLS assumptions. We also discuss measurements of forecast preciseness to compare prediction results. At the end of the chapter we present the models underlying the predictions from Goldman Sachs and OECD, and a Naïve model.

In chapter 5 we analyze the five single financial indicators. We start by examining the underlying models and estimate t values, F scores, R square values and tests for autocorrelation, mis-specification, multicollinearity and heteroscedasticity. Secondly, we conduct a comparison to determine which model provides the most precise forecasts, using MAPE and RMSE as criteria.

In chapter 6 we present the analysis of our FCIs' model setup including four lags. We continue and test for linearity and long-run solutions in the models by employing the Wald test. To determine the number of lags that provide the most information, we conduct a lag significance test.

In chapter 7 we compare MAPE and (R)MSE values from all the various benchmarks, our FCIs and the long-run static solution models. We also conduct a break point Chow test to look for structural instability in the weights. At the end we give a general recommendation, and determine which model we believe to be the most likely candidate for real-life predictions.

Finally, in chapter 8 we provide comments and ideas for future research, as well as learning outcomes from the process of writing this thesis.

1.1 Specification of the problem thesis

The definition of our problem thesis is as follows:

Can we construct an FCI of financial indicators to predict quarterly GDP changes in Norway?

Our idea for writing this thesis was the perceived need to develop a broader measure to capture changes in economic growth after the financial crisis. Several new Financial Condition Indices (FCIs) have been constructed for the U.S. and the Euro Area. Our intention was to combine Norwegian financial indicators, whose counterpart has been proven to have an effect on GDP in these FCIs, to construct a similar FCI to predict GDP for Norway.

To answer these questions, and to obtain a superior understanding of the empirical literature which we survey in chapter 2 and 3, the following sections will give a further description of Norwegian GDP, inflation and effects of changes in the key policy rate.

1.1.1 The Gross Domestic Product of Norway (GDP)

GDP is defined as the total monetary value of all finished goods and services within a country in a specific time period[1]. In GDP, private and public consumption (C), government spending (G), investments (I) and net exports' (NX) are all included; GDP = C + G + I + NX. Three definitions are used to define GDP; the production method (I), the expense method (II) and the income method (III). These approaches exhibit different underlying variables, and due to lack of perfect information, the calculations within each approach may give discrepancies. Statistics Norway (SSB) has calculated Norwegian GDP since 1953, and the numbers are based on real national accounts, thus reflecting real GDP. Three different macro sizes are defined for the GDP:

- GDP, measured by its market value
- GDP Mainland Norway, measured by its market value
- GDP Mainland Norway, measured as basis value

GDP is measured in market value, defined as the sum of value added of all industries measured in basic prices, i.e. the value of all goods or services after taxes and subsidies are taken into account, plus all taxes on manufactured goods less all product subsidies. The GDP numbers are published quarterly, approximately 50 days after the end of the quarter. The final annual national accounts provide the basis year for which constant-price estimates are calculated.[2]

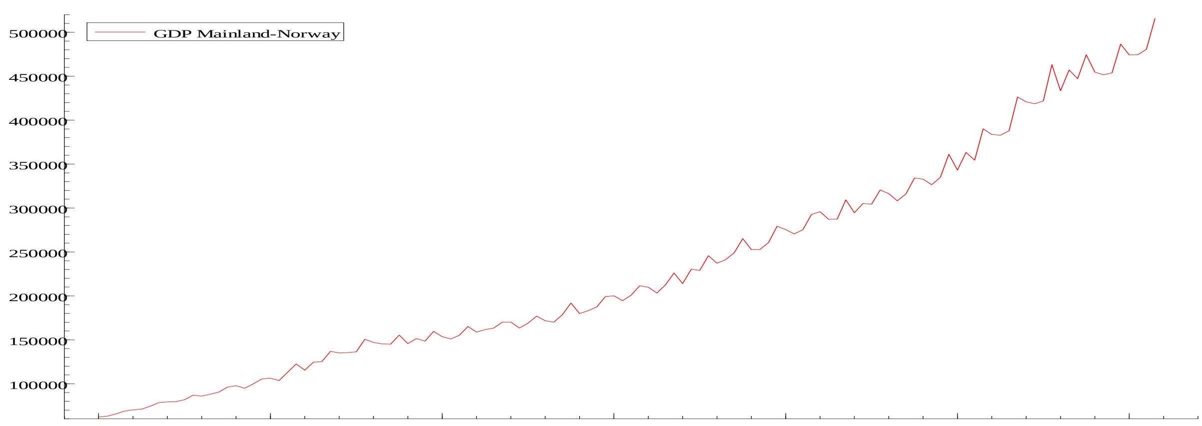
The GDP is also calculated in terms of current prices, annual changes in volume (%) and prices (%), and as seasonally adjusted estimates[3]. The GDP of Mainland Norway consists of all domestic production activity except the extraction of crude oil and natural gas, services incidental to oil and gas extraction, transport and shipping[4]. Economic activity can also be analyzed looking at the output gap or the price deflator. The former is defined as the difference between a potential value and actual GDP, while the latter measures the ratio of current GDP (nominal) to GDP adjusted for inflation (real).

Since we were mainly interested in detecting Norwegian indicators with leading properties on GDP, we regarded the oil and gas producing sector of the economy as exogenous in this setting. We believed that including the oil and gas producing sector would incorporate international influences, which we wanted to exclude. We therefore chose GDP for Mainland Norway. In this thesis we use quarterly GDP observations from 1980 and onwards, based on current prices. (All our references to GDP Mainland Norway in the following are denoted as GDP). In the next section we elaborate on the characteristics on the total period.

1.1.2 Total Period Characteristics

Graph 1.1 shows historical movements in quarterly GDP Mainland Norway from 1980 until the last quarter of 2010.

Graph 1.1 GDP changes Mainland Norway 1980-2010

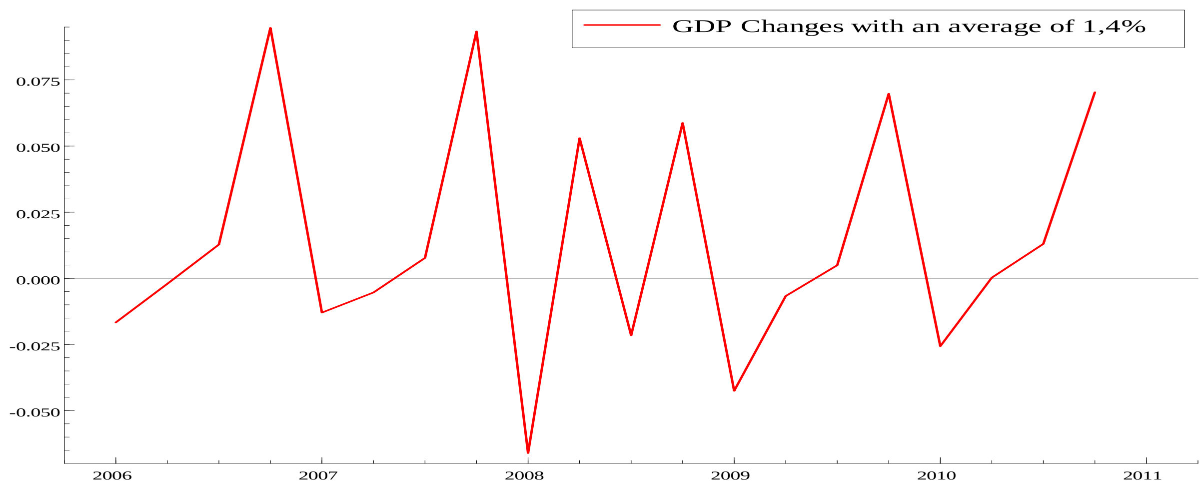


Source: SSB

The graph shows an upward trend in GDP Mainland Norway. The mid 1970s and the early 1980s were associated with lower and more variable growth in productivity, and according to Norges Bank, unstable macroeconomic factors were explanatory reasons. Economic growth recovered in the early 1990s and experienced further growth from 2000 onwards, probably due to the service sector's increased use of IT systems. In 2005 growth decreased and even got negative in 2008[5].

We aim to predict percentage changes rather than level values, because economic growth is typically reported as changes in real or inflation adjusted GDP. Looking at changes, the pattern in GDP Mainland Norway is different. This is exhibited in graph 1.2. Since we were interested in predicting the sub-period 2006(1)-2010(4) only changes from this period are shown.

Graph 1.2 Quarterly GDP changes Mainland Norway from 2006 (1) until 2010 (4)



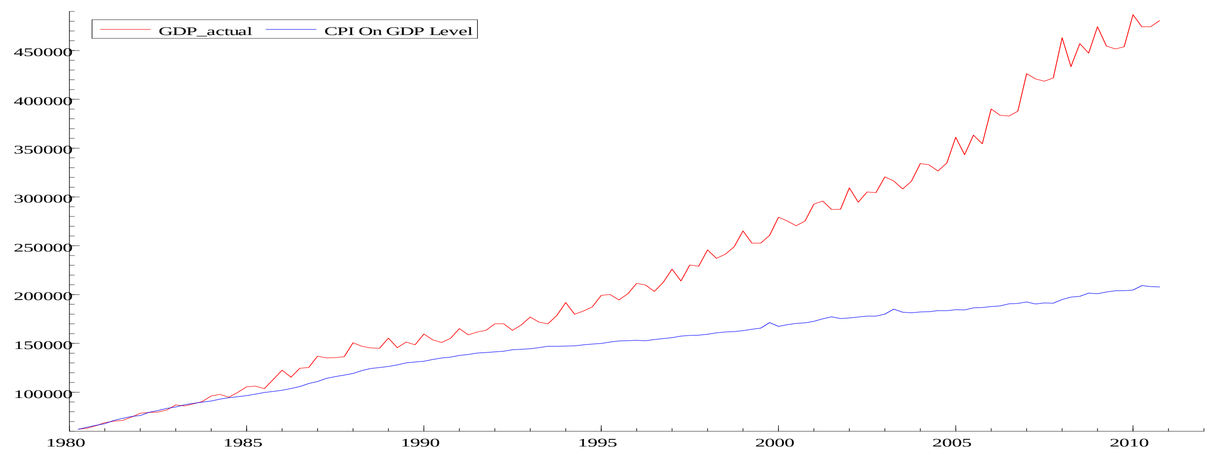
The pattern of changes in GDP Mainland Norway appears more volatile. The average was 0,015 % throughout the period. The first quarter of 2008 experienced the lowest economic growth with negative changes in GDP of 0,067 %. Conversely, the last quarters of 2006 and 2007 had the highest positive growth, 0,094 % and 0,093 % respectively. Only the second quarter of 2008 experienced positive growth, and only the third quarter of 2008 experienced negative growth, over the total period. Overall, the highest growth appeared in the last quarter of every year, while the first quarter each year experienced negative growth, suggesting that GDP is affected by seasonal fluctuations.

1.1.3 Inflation and the Consumer Price Index (CPI)

Inflation is a measure of the persistent growth in the general price level. Usually, inflation is measured by the Consumer Price Index (CPI) which measures the actual percentage price changes for basic household goods and services including charges and fees. Goods and services whose prices are recorded, are all offered to Norwegian consumers, and is therefore a good measure of basic private consumption costs. When the CPI reaches high levels consumers have less available funds to spend on non-basic needs, and the part of the economy who supplies non-basic (elastic) goods and services are most likely to experience the negative effects. If more of the households´ income is spent on basics, less is left for other consumption. This might not necessarily lead to a reduction in overall spending, but it will probably have implications for the consumption allocation in the economy as a whole. This may again slow down GDP growth, as unemployment may increase and consumption is shifted to inelastic goods.

As inflation targets are set by Norges Bank at 2,5 % per annum, so is the CPI expected to increase by roughly the same amount (Boskin et al. 1998). Graph 1.3 shows inflation, measured by quarterly changes in the CPI, and real GDP from 1980(1) to 2010(4). This graph shows that the CPI seems to a follow slower increase than GDP in the period.

Graph 1.3 the Consumer Price Index and GDP



Source: CPI numbers are provided by SSB, 1998=100[6]

1.1.4 Expected affects of changes in the key policy rate

Norges Bank influence economic development by setting the key policy rate. Norges Bank’s monetary policy influences inflation through three possible channels – the Demand channel, the Currency channel and the Expectation channel – which are expected to have these corresponding relationships:

Table 1.1 Expected effects of changes in the key rate

Increased Key Policy Rate	Effect
Short term interest rates	+
Exchange rate NOK	+
NIBOR	+
Consumption and Investment	-
Inflation	-

Short-term interest rates in the money market and the banks' deposit and lending rates, will normally be highly influenced by changes in the key policy rate. Generally, an increased key policy rate will reduce the price growth negatively and conversely, a decreased key policy rate will increase inflation and supposedly increase the price growth. The relationship between the key rate and inflation is however not stable over time because the effect on inflation occurs with a lag and may vary in intensity. Therefore, other factors can have an impact and cause changes in inflation and GDP. Increased demand after goods and services or increased input prices can also contribute to a higher CPI.

2 Financial Condition Indices

An examination of past indices was essential in order to construct an FCI for Norway. Not only to avoid previous mistakes, but also to obtain a broader perceptive of estimation methods that have been applied, financial indicators that have been included, and finally, how well these FCIs have performed in predicting GDP. The majority of the studies focus on larger economies such as the U.S. and the Euro area, which has a greater availability and diversity in the range of financial indicators. These articles have served as a starting point in the construction of our FCI, in particular a newly published study by Hatzius et al. (2010).

In this chapter, we first review how the previous FCIs have emerged in section 2.1. We then present several studies on established international FCIs in section 2.2. In section 2.3 we discuss the estimation of the weights in the models, and the two main approaches to determine them. Table 2.1 at the end of section 2.3 gives a short summary of the main characteristics of the various FCIs.

2.1 FCIs in the international literature

Previously, single indicators such as the yield curve, were used to forecast economic conditions. In the mid 1990s Bank of Canada introduced the first Monetary Conditions Index (MCI). The MCI was calculated as a weighted sum of changes in the short-term interest rate and the exchange rate, and was used to adjust for macroeconomic instability and served as an appealing operational target and guideline for monetary policies.

Several central banks in different countries, including Norges Bank, applied MCIs in their monetary policy. Despite its advantages, the index was not without limitations. Weights were unknown and estimated from different econometric models and consequently strong assumptions were underlying the different parameters. This led to several operational problems such as model specification errors and model specification bias. Lack of ability to judge whether the existing monetary policy fitted the policy objective was another disadvantage associated with the MCIs. In Norges Bank's MCI, the weights were highly sensitive to the estimation sample period and different weight estimates resulted in various inferences about the monetary condition. Even though the problem areas above were dealt with, external shocks not related to domestic monetary policy, such as the oil price, could affect the interest rate and the exchange rate. The MCI could be misleading as a monetary guide.

Eventually more and more indicators were added to the MCIs, and the broader measure became known as FCIs, in order to distinguish them from the MCIs. FCIs have emerged as a broader measure of economic activity and most FCIs include some measure of short-term interest rates, long-term interest rates, risk premium, equity market performance, and exchange rates (Hatzius et al. 2010). This means that the new FCIs incorporate the characteristics of the older MCI and single indicators, in addition to more recent assumptions about economic co-variations.

2.2 Empirical application of previous FCIs

In the following sub-sections we will describe some well known established FCIs; Bloomberg FCI, Deutsche Bank FCI, Goldman Sachs FCI, Federal Reserve Bank of Kansas City FCI, OECD FCI for the U.S., Bank of Canada's new FCI for the U.S and The new American FCI by the National Bureau of Economic research. We discuss indicators that have been included, econometric approaches that have been applied and how the financial indicators have been weighted.

2.2.1 Bloomberg Financial Condition Index for the U.S. (BFCIUS)

BFCIUS contain yield spreads and sub-indices from the Money Markets, Equity Markets, and Bond Markets. Each of these sub-indices are equally weighted (1/3) and consists of ten equally weighted underlying indicators with monthly data from 1991 onwards. The FCI's final values are interpreted as standard deviations and signal if the present financial condition is above or below a mean value. If the index falls into negative terrain it indicates stress within one or all of the three sub-markets or potential stock market weakness (www.bloomberg.com).[7]

BFCIUS neither captured the recent sub-prime crisis in the U.S., nor the low levels of nominal and real long-term interest. This led to the introduction of the new index in 2009, BFCIUS+, which in addition includes real estate prices and several yields – measured as the ratio between Nasdaq / S&P 500, S&P

Homebuilders / S&P 500, 5 Yr Treasury Yield less nominal GDP Growth and the Real BAA Corporate Yield. Each sub-index is given a new set of weights, 1/5 respectively (Rosenberg 2009).

The two indices seemed to correlate most of the time, but the new index proved better in capturing the abnormal high values in 2004–2007. BFCIUS+ had 2 standard deviations above its average in this period. The new index also gave clearer signals of the financial crisis that started in 2007 and was generally a better measure of the overall economic performance (Rosenberg 2009).

2.2.2 Deutsche Bank Financial Condition Index (DB FCI)

Deutsche Bank's index has been available since 1983, and DB combines both the principal component and the weighted-sum approach. Exchange rates, house prices, asset prices and bond indicators are all included in their index, a total of seven U.S. financial variables. The principal components (PC) are derived from these variables and the index is constructed as a weighted sum (WS) of the short-term interest rate (federal funds rate) and of its PCs (www.db.com).[8]

In their Global Economic Perspectives (2010), Deutche Bank presented an updated version of DB FCI – the Monetary Policy Forum Financial Conditions Index (MPF FCI) that includes several additional indicators. In total, 45 financial indicators are incorporated in the new MPF FCI, including interest rate spreads, Treasury yields, asset prices, volatility measurements and economic surveys. Financial stock and flow variables are also given substantial weight. The index is constructed using unbalanced panel estimation techniques, because some of the new variables has limited history and are released with a time lag, such as the lending survey from the Federal Reserve Board.

One particular set of variables contributed to the substantial increase in explanatory power in the new index compared to the original index. According to the rapport the top "performers" explaining the index' upside were banks' willingness to lend, a market cap ratio – used to determine whether the overall market was under- or overvalued, a tightening in commercial and industrial (C & I) Loans, the 10-Year Treasury Yield and a spread between Jumbo/30yr Conventional. On the other hand, the VIX index – a summary for market volatility, the Real Broad Trade-Weighted Dollar – the value of the U.S. Dollar relative to other world currencies, the Wilshire 5000 – an index for the U.S. equity market and finally, a non-mortgage ABS Issuance were the five best "performers" in predicting the downside.

2.2.3 Goldman Sachs Financial Condition Index (GS FCI)

According to Hatzius et al. (2010) Goldman Sachs incorporates a short-term bond yield, a long-term corporate yield, the exchange rate and a stock market value in their FCI, which utilize a weighted sum approach. The weights are determined by Goldman Sachs' own model and the Federal Reserve Board's large scale macroeconomic model. An increased index indicates tighter financial conditions, and correspondingly, if the index decreases it signals improved financial conditions. The GS FCI differs from other indices because it utilizes levels of financial variables rather than spreads or changes in the variables. The index has therefore experienced a noticeable downward trend.

2.2.4 Federal Reserve Bank of Kansas City Financial Stress Index (KC FSI)

In a study by The Federal Reserve Bank of Kansas City, Hakkio and Keeton (2009) presented a new index to capture five key aspects of financial stress. The Federal Reserve utilizes a PC approach to measure 11 financial indicators. Each variable have to represent at least one of the five key aspects of financial stress, reflect prices or yields on financial markets, and be available on a monthly basis since the 1990s.

Compared to the first index by Bank of Canada (BOC), the KCFCI does not include indicators such as exchange rate volatility (more important for small open economies like Canada) and the yield curve (which reveals more about the stance of monetary policy than financial stress). The KCFSI also excludes investment uncertainty about bank stock prices (www.kansascityfed.org/)[9].

The KCFSI suggests that high values coincide with periods of financial stress. Positive values indicate financial stress above the average and vice versa, negative values signal stress below the average. Since 1990, the KCFSI generally peaked during known episodes of financial stress. Only two periods had not been captured – the Mexican peso crisis in late 1994 – and the Asian Financial crisis in 1997. These crises were mainly international and expected to have less effect in the U.S. The results also indicate that financial stress can lead to a decline in the economy through three possible channels; (1) uncertainty about prices or other investors actions, (2) business and households financial spending and (3) tighter credit standards in banks either by raising (a) the interest rate or (b) minimum standards (Hakkio and Keeton 2009). The authors' suggestions as to when to tighten, is for policymakers to know if financial stress no longer poses a threat to the economy. The article did not address a critical level at which financial stress is a serious concern.

2.2.5 OECD Financial Condition Index for the U.S. (US FCI)

On behalf of the OECD, Guichard and Turner (2008) developed an FCI to capture financial conditions in the U.S. In addition to standard measures such as the exchange rate and short and long-term interests, they also include credit standards, various measures of bond spreads and stock market capitalization and real house wealth – expressed as a ratio to GDP. In total are six indicators from 1990(4) and up to 2007(4) included. The six indicators that comprise the FCI are equally weighted and each sub-indicator is again weighted according to its relative effect on GDP changes. The estimation is conducted using two models with overlapping data material: first, a reduced form equation model is used, and then a Vector Autoregression model (VAR) to account for any type of correlation between the variables.

The VAR model suggested that credit standards and high-yield spreads were correctly signed, and in most cases significant. The long and short-term interest rates were also correctly signed but often not significant, nor the exchange rate. Stock market capitalization was weakly significant.

One novelty with the index was the inclusion of a survey measure to capture the tightening in banks' credit lending standards. The survey was a Senior Loan Officer Opinion Survey of bank lending practices conducted by the Federal Reserve Board (FED). The results showed that a tightening in the survey response of 10 percentage point lead to a decrease in U.S. GDP by approximately ¼ percentage points after four to six quarters (www.oecd.no)[10].

The study was extended by Guichard, Haugh and Turner (2009) [11] to also include the Euro area, U.K. and Japan. Unlike the U.S., survey measures on credit conditions in the former countries had only been available for a limited time period, thus making data mining and comparison complicated. To overcome the problem, the authors utilized a bank lending survey for the Euro area conducted by the European Central Bank (ECB) based on data from 2003 onwards. The survey was extended further by using U.S. variables, the yield curve and a business survey regarding investment in France. For the U.K., a business survey developed by the confederation of British Industry (CBI) asking business executives if external finance was a limiting factor for investment were used. For Japan, the authors found a significant relationship between tighter credit standards in the loan survey and lending attitudes, and so included a study on lending attitudes of banks.

In all the countries, the survey measures showed that banks had tighter credit standards than previously experienced. The tighter lending standards and their effects on GDP could be doubted because bank lending to businesses experienced increased growth throughout 2007 and into the first half of 2008. The increased growth was justified by the delays that tend to exist between tighter lending standards and credit growth. The authors suggested that operators in the short-run respond to tighter lending standards with increased demand after credit, though in the long run there tend to be a reduction in the growth of bank lending. Guichard et al. (2009) did not exclude the financial crisis as a possible explanation for the delay.

2.2.6 Bank of Canada (BOC) – FCIs for the U.S.

In 2009, Beaton et al. developed two new American FCIs for the Bank of Canada (www.bankofcanada.ca/en)[12]. The first index, the SFCI, was constructed by a structural vector error correction model of similar art as VAR models, while the second index, the MFCI, was based on Bank of Canada's large-scale macroeconomic model MUSE. The indices are unique because they both measure contribution to growth from financial shocks in a given quarter and the tightness of financial conditions (Beaton et al. 2009). Positive values signal expectations of GDP growth, while negative values indicate reduced growth. Tighter financial conditions are identified as a decline in a positive value over time, and conversely, looser financial conditions are identified by an index that becomes less negative over time. The indicators in the two FCIs differ slightly, but both include interest and bond rates, the real exchange rate, housing and financial wealth measures and various lending measures.

The overall pattern depicted by the two FCIs was quite similar. Financial wealth contributed most in the SFCI, followed by loan standards for consumer spending, business borrowing spreads and the commercial interest rate. In the MFCI financial wealth contributed less than total loan standards, but more than the business borrowing rate. The SFCI was thus more volatile than the MFCI due to BOC's MUSE Model's forward-looking nature.

2.2.7 The National Bureau of Economic Research (NBER) - The New American FCI: A Fresh Look after the Crisis

Hatzius et al. (2010) developed a new American FCI containing 45 indicators. Each variable are analyzed quarterly and contain data spanning back to 1970. The authors utilize a standard PC approach, but their index feature three key innovations. Compared to other FCIs, additional indicators are included and cover a wider range of both quantitative and survey based indicators. In addition, the use of panel estimation techniques allow for unbalanced time series and by controlling for past GDP and inflation which further confer the predictive power of the FCI (www.nber.org)[13].

Compared to the OECD FCI for the U.S., this new index includes a broader range of survey-based measures. In total seven surveys covering bank lending, consumer and business credit conditions in the U.S. – conducted by the FED, the National Federation of Independent Business and the University of Michigan – are included. To analyze the survey's performance prediction tests two and four months ahead were carried out. The results show that the relative mean square error (RMSE) of the group of surveys were lowest when tracking growth four months ahead in the final period (2005 – End of Sample). In other words, there was less noise in the surveys' forecasting performance looking four months ahead.

The new index' biggest weakness was its size, which made estimation and updates more complex. The new index outperformed any of its major subcomponents, such as spreads and asset prices. It also showed that the explanatory power from a number of financial variables not included in earlier FCIs became stronger when the latest financial crisis was covered. Even though its predictive power was unstable in earlier periods, the new index outperformed various recent measures. Primary reasons for the index' improved explanatory power were the expanding numbers and variety of the variables and by excluding macroeconomic shocks. The latter contributed somewhat more to the FCIs forecasting power than the inclusion of the new variables, but the FCI's forecast performance seemed better in periods of financial instability coming from within the asset market. The new FCI also showed an abnormal condition in the credit market at the end of 2009, implying that the economy in 2010 still was affected by the financial crisis.

2.3 Determining the weights in the FCIs

In determining the weights attached to each variable, two methods have been applied in all the studies we surveyed above, either individually or combined. In the following we address the two approaches – the Weighted Sum (WS) approach and the Principal Component (PC) approach.

2.3.1 The Weighted Sum Approach

In the WS approach three different estimation techniques are commonly applied to validate the effect of financial shocks on GDP growth; Aggregate Demand (AD) models, Large Scale Macroeconomic models and a Vector Autoregressive (VAR) models.

In the AD model, movements in aggregate demand can be used to predict exogenous factors' effect on real GDP. The aggregate demand curve is a function of investment, consumption, government spending and net exports at a static level. Large Scale Macroeconomic forecasting models can handle a greater number of variables and observations, and includes multiple equations. In these models the variables are typically so numerous that the number of variables often exceeds the number of observations. Many banks utilize Large Scale Models, among others Norges Bank. Two macroeconomic models are used by Norges Bank to produce projections for key macroeconomic variables and these models are constructed based on several different types of models, exists in several variants, lag lengths and for different estimation periods.

VAR models are by far the most common estimation approach in the construction of the previous FCIs. According to Bjørnland (2000), VAR models are better than Large Scale Models in analyzing economic fluctuations because they are more flexible and capable of describing the dynamic structure between economic variables.

VAR models deal with both lagged values of the target variable and a vector of the independent variables (Gujarati and Porter 2009). There are several advantages with VAR modeling. First of all, all the endogenous variables are considered simultaneously, and each variable is explained by its lagged or past values, and the lagged values of the other endogenous variables in the model. All variables are treated equally, so there are no distinctions between endogenous and exogenous variables. VAR models are also easily applied and often provide better forecast estimates than other models (Gujarati and Porter 2009).

Despite its many advantages, VAR models do exhibit some disadvantages. VAR models are for example less suitable for policy making/analysis because of their emphasis on forecasting. Estimating too many parameters will provide fewer degrees of freedom, and can lead to several problems. For small samples, it can be challenging to define an appropriate number of lags. In addition, most time series variables are non-stationary and must be integrated by first-differentiating before they can be included in a VAR model. Consequently, these estimates may suffer from bias and the estimation results may be useless (Gujarati and Porter 2009).

2.3.2 Dynamic Factor Models (DF) – The Principal Component Approach

The second method, the PC approach, is the most common version of a factor analysis. An increasing number of researchers apply DF models to forecast key macroeconomic variables, such as GDP and inflation. The availability of data at a more disaggregated level and the opportunity to cope with many variables without running into degrees-of-freedom problems are, according to Jolliffe (2005), partly why DF models experience such increased popularity. Elimination of idiosyncratic movements that include measurement errors and local shocks has also supported the increased use of DF models.

The inspiration (according to Ziegler and Eickmeier 2006) dates back to Burns and Mitchell (1946). The central idea was to reduce dimensionality in large data sets, while retaining as much of the original variance as possible. That way, the bulk of variation of many variables could be explained by a small number of common factors or exogenous shocks. By utilizing an orthogonal transformation, the PC approach converts observations from variables that are possibly correlated into uncorrelated variables. The new variables, the PCs, are linear combinations of the original variables, but replace them by a smaller or equal number. The PCs are then sorted according to their variance, where the first components exhibit the highest variance. By doing so, the parameter uncertainty which is most likely induced by poor forecasting performance, is to some extent avoided. In addition, it allows for out-of-sample forecasting so that its performance could be assessed and compared to other approaches. The PC approach is also a common technique to recognize patterns in the data series.

2.3.3 A discussion of the two approaches

Camba-Mendez et al. (1999) investigated several leading indicators and GDP growth for the European countries France, Germany, Italy and the United Kingdom. Traditional VAR models only deal with a limited number of variables and to overcome the problems associated with limited degrees of freedom, the authors utilized a PC Dynamic factor analysis. By doing so, they could pool information from a large set of variables. Their results indicated, with a few exceptions, that the DF model outperformed the VAR model in both the in-sample and the out-of-sample period.

Jolliffe (2005) argued that, despite the DF-models' success and preciseness in terms of forecasting and analysis, some remarks must be illuminated. Firstly, its forecasting performance depends among other things, on the target variable, country, data sets, benchmarks and time horizon, but an assessment of the determinants of the forecast performance is still not available. Also, factors that explain less of the entire panel, like the fifth or sixth principal component, may be important for the target variable. Therefore, including just the first factors may not be sufficient. Finally, the selected variables in the data set are mostly ad hoc and often used to predict other variables. This may not be adequate as only variables that exhibit high explanatory power should be included.

Ziegler and Eickmeier (2006) analyzed the DF models' ability to forecast GDP output and inflation. Interestingly, the factor models seemed better suited to forecast GDP in the U.S. area than the Euro-area, but in the case of inflation there were no significant differences. Alternative methods, such as large scale models, were found to provide slightly better forecasts than factor models, but the latter generally outperformed small-scale models. Ziegler and Eickmeier (2006) also found some evidence that factor models were better suited for quarterly than monthly data and that the forecasting power for GDP worsened when the horizon increased. The estimation technique was also found to matter, in that the dynamic approach tended to outperform the static approach, even though the latter approach is more common, and easier to implement.

2.4 Summary and Final Comparison of the FCIs

Table 2.1 summarizes the main findings from all the studies of the FCIs.

Table 2.1 Summary of past FCIs

Study	Financial Indicators	Approach
Bloomberg BFCIUS+	5 sub-indices (#'s in parentheses); the Money Market (3), the Bond Market (5), the Equity Market (3), Asset Bubbles (2) and the Yield Gap (2).	WS approach; Each sub-index is weighted 20%. Indicators underlying the Money market are weighted 6.7 %, bond market indicators 4 %, tree last sub- indices indicators are each weighted 10%
Goldman Sachs	Stock market Capitalization /GDP	Weights based on GDP effects derived from the FED macro model and GS own modeling.
Deutsche Bank MPF FCI	45 indicators; interest rates (1), yield spreads (14), the exchange rate (1), stock market (2), house prices (1), Market volatility/risk (3), Survey of lending standards (7) and Quantitative stock and flow indicators (15).	Combination of both the PA and the WS approach - The PCs is extracted from the various variables, and then the index is equally weighted by its PCs and the Fed rate.
OECD FCI for the U.S.	6 indicators; Real short-term and long-term interest rates, bond spreads (2) credit standards tightening (survey-based measures), the real exchange rate and a ratio to GDP/ de-trended. (Stock market capitalization and real house wealth)	WS of the six indicators, then weighted according to their %-age effect on ΔGDP
Bank of Canada The new FCI for the U.S	SFCI included; Commercial paper rate, Business borrowing spread, Loan standards for consumer Spending, Financial wealth. In addition the MFCI included; Fed rate (instead of the Commercial rate), business borrowing rate (instead of the spread), Loan standards for residential and business investment, Mortgage rate, Real effective exchange rate (REER).	WS approach -Two FCIs constructed by a Vector error correction model (SFCI) and a large scale macro model (MFCI)
The New American FCI	45 variables including; interest rates (15), prices (5),flow and stock quantities (15), surveys (7) and 2nd moments (3)	PC approach with three new features; (1) Broader range of flow/stock and survey measures, (2) unbalanced panels, (3) Control for past GDP growth/inflation.
Federal Reserve Bank of Kansas City KC FSI	11 indicators; TED (3mnd LIBOR/T-bill), various spreads (swap spreads, treasury spreads, High yield bonds spreads, 6 in total), stock-bond correlation, stock market volatility (2), Cross-section dispersion of bank stock returns	PC approach

Sources: The various FCIs discussed in section 2.2.1 to 2.2.7

Overall, in most FCIs short-term interest rates, exchange rates, house prices, asset prices and different types of spreads are included. The WS VAR approach appears superior to the PC approach in most cases, although newer FCIs seem to prefer the PC approach.

3 Empirical applications of various indicators forecasting performance

A considerable amount of empirical work has focused on single financial indicators' predictive power on economic growth and inflation. Even though international and national data cannot be perfectly homogenous, we use international research as a starting point for the indicator selection. In section 3.1 and 3.2, we present several international and Norwegian empirical studies. Our main findings are listed in table 3.1 and 3.2. In sub-section 3.3 we present financial indicators thought to have leading characteristics, to get an idea of indicators that might be appropriate to include in our FCI.

3.1 International research

A considerable amount of economic studies has focused on the yield curve. The U.S. yield curve, also called the term spread – the difference between long-term and short-term interest rates – has frequently been measured as the spread between the fed funds rate and the 10 year Treasury bond yield. Laurent (1989) analyzed the spread between interest rates on long treasury bonds and the fed funds rate in the U.S. First, the author reviewed general characteristics of leading indicators and specific properties of the spread. Secondly, he analyzed how the term spread had recently performed in forecasting economic growth, and finally, if and how the spread could forecast economic growth in the future. The results showed that there existed a relationship between the term spread and GDP growth, a widening in the term spread was followed by accelerations in GDP growth, and opposite, a narrowing in the term spread was followed by decelerations in the spread. Laurent (1989) found no significant or promising proof of the term spread as a precise forecast to economic growth, but he argued that it could be useful as a guide for movements in GDP growth.

Estrella (1998) argued that financial variables, such as prices of financial instruments, are commonly associated with expectations of future economic events. Estrella (2005) also tested the term spread's performance in the U.S. Contrary to Laurent (1989), he found that the yield curve had predicted every U.S. recession since 1950, except for the credit crunch and slowdown in production in 1967. Estrella and Mishkin (1997) established evidence (according to Estrella 2005) that this predictive relationship also existed in other countries, particularly Germany, Canada and the U. K.

Goodhart and Hofmann (2001) analyzed the predictive power of asset prices for monetary policies, the output gaps and inflation, and constructed FCIs for the G7 countries. The analysis was based on the sample period 1973(1) to 1998(4). They utilized both an AD model and a VAR model and included an equally weighted sum of the short-term interest rates, the real exchange rate, equity prices (an All-Share Index) and a real estate price index in their FCIs. In the VAR model the oil price was also included and the authors allowed for a maximum of five lags for each indicator. They conducted an out-of-sample analysis and calculated Root-Mean Square Errors (RMSEs) to assess the forecasting performance of the FCIs. The results showed that all asset prices had a significant effect on the output gap in the AD model. In the VAR model short-term interest and the majority of the house prices had a significant effect both on the output gap and inflation. The exchange rate was only significant in about fifty percent of the cases, but was always correctly signed.

Borio and Lowe (2002) studied gaps in asset prices and their effect on both economic growth and monetary policies. The study was conducted by analyzing asset prices' single effect on GDP growth and the combined effect of asset prices and the investment gap (the deviation between actual- and a trend value), the credit gap and credit growth. The idea was to identify indicators that could predict banking crises through upper threshold values. Data from 34 countries during the sample period 1960-1999 were analyzed and the authors proved that combinations of indicators gave the best predictions. The credit gap was the most reliable indicator, while gaps in real equity prices, investment and credit growth had less predictive power. Expanding the time horizon gave real equity prices and credit growth improved predictive power.

Gropp et al. (2002) analyzed equity prices and the bond market as early leading indicators for bank vulnerability in the Euro area. In their dataset they used monthly observations for the period 1991(1) - 2002(2). The authors found that both indicators performed quite well as leading indicators.
They also obtained significant results indicating that both the bond spreads and the equity prices – measured as distances-to-default – has leading properties on a 2 to 4 quarter horizon. The results also pointed towards a significant difference between the two indicators. The equity prices had less predictive power in the last days before maturity, while bond spreads had the more predictive power closer to default.

Banerjee and Marcellino (2003) compared various single equation models' ability to capture GDP growth and inflation in the U.S. and discussed variables and characteristics to look for in indicator selection. Their first sub-problem regarded the selection of leading indicators and lag lengths. The authors employed a method (developed by Hendry and Krolzig in 1999) that included an information criteria, significance tests of the parameters and model specification tests of the residuals. Their second- and third sub-problem regarded pooled indicators, groups of indicators and pooled forecasts provided by single indicators. The study was extended by Banerjee et al. (2003) who analyzed leading indicators, inflation and GDP in the Euro area between 1975 and 2000 using both series from the Euro area and the U.S. 46 Euro-area variables and 16 U.S. variables were analyzed. The results suggested that the short-term interest rate, public expenditure, total industrial production, world GDP and demand growth could be used as leading indicators. Employment and unemployment were also included in the set of good indicators. The results matched the findings by Banerjee and Marcellino (2003) that the best single indicator systematically beat the best group, and that no indicator served to be best more than twice.

Montagnoli and Napolitano (2005) considered how asset prices influence monetary policies and constructed an FCI for the U.S., Canada, Euro Area and the U.K. The authors argued that asset prices have a forward-looking nature and contain information about future demand and subsequent inflation, and could therefore serve as a good indicator of economic growth. Based on the efficient market hypothesis (EMH), they suggested that current asset prices should contain all past information and therefore no lagged values of asset prices were included. Additional indicators captured by their index were short-term interest rates, exchange rates, inflation, the output gap and a house price index, with six lagged values respectively. The FCIs in all the countries, except for the Euro area, were significant and positive as a short term guide for monetary policy, meaning that asset prices played a positive role for monetary purposes. Lack of significance for the European FCI was explained by the more complex European banking system and that financial markets in Europe are not as well integrated as in the single countries. The article gave positive support to asset prices as a leading indicator. Moreover, it coincided with Goodhart and Hofmann´s results from 2001 which showed that short-term interest rates, the exchange rate, inflation, the output gap and house prices could serve as leading indicators.

Table 3.1 Summary from various international studies on leading indicators and GDP

SUMMARY OF INTERNATIONAL RESEACH		
Author(s)	Study	Main result(s)
Laurent (1989)	Tested the term spread's effect on GDP growth in the U.S.	-A wider term spread led to accelerations in GDP growth, and a narrowing led to decelerations in GDP growth. -Not significant as a precise forecast to economic growth, but could be used as a guide
Estrella (2005)	Tested the term spread in the U.S	-Significant in predicting almost every U.S. recession since 1950 -Established evidence of similar relationship in Germany, Canada and the U. K
Goodhart and Hofmann (2001)	Analyzed asset prices and its effect on the output gap and inflation and constructed FCIs for the G7 countries	-All indicators had a significant effect on the output gap -Short-term interest and the majority of the house prices had a significant effect on the output gap and inflation. -The exchange rate was only significant half of the time
Montagnoli and Napolitano (2005)	Analyzed asset prices and Monetary Policies, and constructed FCIs for the U.S., Canada, Euro Area and the UK including; short-term interest rates, exchange rates, inflation, the output gap and a house price index, with six lagged values respectively	-Asset prices played a positive role for Monetary purposes, except for the Euro Area -All the indicators could serve as leading.
Borio and Lowe (2002)	Analyzed gaps in asset prices, investments and credit, and credit growth, single and combined effect on GDP	-Credit gap most reliable indicator -Gaps in asset prices and investments, and credit growth had less predictive power.
Gropp, Vesala and Vulpes (2002)	Analyzed Equity and the Bond Market as early leading indicators for bank vulnerability in the Euro area	-Equity prices and bond spreads had leading properties on a 2 to 4 quarter horizon.
Banerjee et al. (2003)	Analyzed leading indicators, inflation and GDP in the U.S. and Euro area	-Short-term interest rate, public expenditure, total industrial production, world GDP and demand growth could function as leading indicators - Employment and unemployment were also good indicators - The best single indicator beat the best group -No indicator served best more than twice.

3.2 Research on Leading Indicators for the Norwegian Economy

In this section we expand by surveying Norwegian literature on leading indicators and GDP.

Husebø and Wilhelmsen (2005) analyzed 30 Norwegian variables' effect on the output gap, and examined their behavior against the U.S. and the Euro area. The analysis contained data from the period 1982 - 2003. The authors found evidence of a somewhat similar relationship between several economic indicators and GDP both in the U.S. and the Euro area, both in terms of strength and whether they seem to lead, lag or coincide with GDP. Even though the Norwegian economy cannot be compared to the U.S. and the Euro area directly, Husebø and Wilhelmsen (2005) suggested that there exist several similarities. Indicators that measure consumption, investment and labor market levels all had strong correlations

with GDP and coincided with U.S. estimates, thus implying that there exist some similarities between smaller economies like Norway and larger economies such as the U.S.
Riiser (2005) analyzed gaps in real house prices, real equity prices, investment and credit, and their effect on GDP in Norway between 1819 and 2005. The study by Borio and Lowe (2002) were used as a starting point for this analysis. Real house price gaps peaked from one to six years before the outbreak of a crisis and generally narrowed down at the beginning and remained negative throughout the crisis. The investment gap showed a similar pattern. In contrast, Borio and Lowe (2002) found that the investment gap had less predictive power. The credit gap typically followed the other gaps. The credit gap results were affected by the relatively short dataset and could therefore be misleading. Lack of data regarding equity prices also made the evaluation complicated. Overall, all the gap indicators were found useful in signaling imbalances prior to a crisis. At least two of the gap indicators simultaneously had high values prior to the banking crises, suggesting that combinations of indicators could increase the predictive power of the analysis.

Riiser (2008) presented updated figures from Riiser (2005). The major difference was the decreased critical value for the credit gap. The analysis also revealed that all the gap indicators had high critical values in 2007. Gaps in house prices and credit both exceeded their critical limits, while the investment gap approached its critical level. All the gap indicators, except for the credit gap, fell in 2008, a pattern also observed in Riiser (2005). Riiser (2008) also analyzed the gap indicators ability to signal a crisis in the future. The gap indicators only revealed imbalances associated with dept and did not signal imbalances in other markets than those included. This was defended by the potential relationship that exist between liquidity risk and financial debt imbalances, and therefore the credit gap indicators should be able to capture and signal such imbalances as well.

Riiser (2010) presented revised figures for the gap indicators of Norway up to 2009. The path of both investment and the credit gap were broadly unchanged so earlier evidence still applied. The house gap, the investment gap and the credit gap all had signs of increased financial instability. In 2009 only the credit gap had a higher value than its critical level, compared to all in 2007, suggesting that corrections in financial instability take a long time. We did not find this unreasonable as credit growth affect the economy with a lag.

Gerdrup et al. (2006) argued that the correlations between the economy and financial sizes change over time. Therefore, information from previous periods will contain less relevant information about the future. The nineteen eighties were also characterized by a more liberal banking system and the banking crisis between 1988 and the early 1990s. The analysis was therefore conducted on data from 1993-2005 and included several indicators; real money supply (M1 and M2), real credit growth (K2 and K3), real house prices, real asset prices, real short-term interest rate, real exchange rate and the spread between 5 year treasury bills and 3mnd NIBOR .

The authors found that the coefficients between GDP growth and the lagged values of real credit and real money supply were either low or negative throughout the period. Real growth in C2 companies and C3 mainland companies affected GDP growth with a delay. This was not striking since C2 and C3 are published with a lag, one and two months respectively. In other words, growth in GDP could serve as a leading indicator for changes in real credit. Only M1 seemed to provide leading information about GDP-growth. The correlation was strongest in the previous and same quarter, thus reflecting the fact that M1 could be related to the economy in the short term. The correlation between real share prices and GDP-growth was strong, especially in the same and the previous quarter. Real short-term interest rate (NIBOR) was negatively correlated with GDP throughout the period, while the spread between 5 Yr Treasury bills and 3 month NIBOR was positively correlated, and a better indicator than the short- term interest rate. The real exchange rate was positively correlated with BNP-growth.

To overcome some of the limitations associated with simple correlation analysis and to account for the fact that many variables could affect GDP with different time lags, the model was extended to a Simultaneous Equation Model (SEM) that included several explanatory variables and contained data from 1990 to 2005.

The chosen model included lagged values of asset prices, several various measures of credit growth (C2) and GDP growth. Growth in real credit to companies (C2 companies) gave information about GDP in the same and previous quarter, although the latter had less explanatory power. Growth in real domestic credit (C2 domestic) was affected by its lagged value (C2 domestic_1). Gerdrup et al. (2006) also detected a constant long-term relationship between real credit (C2) and real share prices. Growth in real stock price was affected positively by increased GDP growth in the same and previous quarter. A shock in GDP growth could therefore affect real stock growth, which in turn could affect real credit growth (C2) to companies and GDP growth. The model predicted GDP-growth and real credit growth (C2) relatively well the first eight to six months into the future. However, it had a problem in the predictions of growth in real credit (C2) in the second quarter of 2005, probably caused by increased demand for foreign credit. The model's lack of capability to explain the powerful growth in real stock prices in 2004 and 2005 was justified by exogenous factors such as high oil prices, which were thought to have an influence. The forecasting errors, when compared to uncertainty in the estimates, were small for both credit and stock prices. Overall, Gerdrup et al. (2006) found evidence that stock prices, credit growth (C2), money supply growth (M1), real exchange rate, and the spread between long- and short term interest rate were good leading indicators.

Table 3.2 Summary of Norwegian literature

SUMMARY OF NORWEGIAN RESEARCH		
Author(s)	Study	Main result(s)
Husebø and Wilhelmsen (2005)	Analyzed 30 Norwegian variables' effect on the output gap, and examined their behavior against the U.S. and the Euro area.	-Consumption, investment and labor market indicators had strong correlations with the output gap and coincide with U.S. estimates
Riiser (2005)	Analyzed gaps in real house prices, real equity prices, investment and credit, and credit growth, single and combined effect on GDP.	-All the gap indicators were found useful in signaling imbalances prior to a crisis. -At least two of the gap indicators had high values simultaneously
Riiser (2008)	Updated figures from Riiser (2005)	-All the gap indicators had high critical values in 2007 -Gaps in house prices and credit exceeded their critical limits -The investment gap approached its critical level -All the gap indicators, except for the credit gap, fell in 2008
Riiser (2010)	Revised figures for the gap indicators of Norway up to 2009.	-The house gap, the investment gap and the credit gap all had signs of increased financial instability -The credit gap had higher value than its critical level in 2009
Gerdrup et al. (2006)	(1)Correlation analysis between single indicators and GDP growth (2)Simultaneous Equation Model (SEM) that included lagged values of asset prices, credit growth (C2) and GDP-growth	**(1)**C2 and C3 affected GDP growth with a delay. -Current and lagged M1 provided information about GDP growth -Current and lagged asset prices had a strong correlation with GDP growth. -The spread and the exchange rate were positively correlated with GDP, while NIBOR was negatively correlated. **(2)** C2 gave information about GDP in the same and previous quarter -Growth in real stock prices were affected positively by GDP growth in the current and previous quarter -Overall stock prices, credit growth (C2), money supply growth (M1), real exchange rate, and the spread were good leading indicators.

3.3 Financial indicators and GDP

Our objective was to build a financial conditions index of leading indicators, and in this section we discuss possible connections between financial variables and GDP for Mainland Norway. Proven correlation or leading characteristics was of importance in order to select indicators to include in our Norwegian FCI. The studies discussed in chapter 2 and section 3.1 and 3.2 have served as our reference point. Earlier financial FCIs have shown to contain limitations because they contain a short history and include a small number of underlying series that may ignore important financial conditions.

We place special emphasis on the newly published study by Hatzius et al. (2010). They included a broader range of quality and survey-based indicators, such as different surveys capturing the tightening conditions in the credit market. These indicators proved to be more important under the recent financial crisis than they had been in the past. Most of the indicators in the new FCI were selected based on empirical results in previous FCIs. Some of their variables were however chosen after the crisis because they had abnormal values and were proven to have an effect on GDP growth. We intended to incorporate many of the methods employed in this study, but to avoid including indicators only because they had abnormal values within the period of prediction.

In empirical surveys meant for prediction purposes it is desirable to have as many observations as possible. For practical purposes it is also desirable to keep the level of complexity to a minimum, especially since this index is meant as a tool for actual real-life predictions, and not a technical exercise. Our FCI should give as precise predictions as possible, while still being fairly easy to use, and the input variables should be readily available. The issue of complexity therefore, functions as a guideline for the number of variables that was included in our FCI. Fortunately, finding an equilibrium between these two contradicting elements did not turn out to be a challenge.

Gerdrup et al. (2006) provided several criteria for selecting indicators contributing to GDP growth in Norway and suggested that financial variables were suitable as leading indicators if they were either; (1) Priced on the basis of expectations about the future, (2) Affected the economy with a lag, or were (3) published more often and faster than GDP numbers. We applied the same criteria for the indicator selection, and variables included in our final selection are published more frequently than GDP numbers. Some of the indicators are priced on the expectations about the future, while some are thought to affect the economy with a lag, in other words all our indicators meet at least two of criteria.

3.3.1 Indicator selection and historical coverage

Ideally, we hoped to find data series reaching as far back as the 19th century in order to draw inferences between crises in Norway. Unfortunately, due to the reconstruction of the banking system and the liberalization in the nineteen eighties this was fairly difficult. In addition, we have less access to financial data in Norway than in the U.S. Many markets have emerged over time and new series have become available while some even overlap existing ones. Thus, we do not have as long a history and as many indicators as desired, and therefore our effort to overcome limitations of previous FCIs are affected. As will become evident later in this chapter, not all of the variables that we wanted to include could be included, and these are presented in section 3.4.

Among the independent variables we considered initially, we ended up including NIBOR 3m, a 5 year Spread between short and long-term interest rate, the Oil price, OSEAX, A Real Estate Index, the Real Exchange Rate, Money supply and Credit growth. GDP in Norway is calculated based on real national accounts, thus reflecting real numbers. All the variables we have included are adjusted for inflation, using the changes in CPI, hence we account for each variable's real effect on economic growth. In the following we present detailed information of the various indicators, summarized in table 3.3 at the end of this section.

3.3.1.1 Interest rates

NIBOR 3M

The Nibor 3m is the three month interest rate at which inter-bank lending without collateral occurs. The rate is set in the three month currency SWAP market, and is decided by aggregate supply and demand in the money market. There are Nibor rates with other horizons in addition to the three month rate, and all are swap-based rates decided in the same manner[14]. This inter-bank rate is useful because it serves as an indicator of the cost of capital in the inter-bank market, as well as having an impact on other interest rates such as those offered to borrowers or lenders[15].

There are a variety of reasons why this variable should be included. Its positive and negative changes are likely to have an impact on GDP, some of which was proven by Gerdrup et al. (2006). An increase in the Nibor 3m rate would make capital more expensive in the short term capital market, an effect most likely to be evident in the increase of the short-term lending rate offered to customers by any given bank. As inter-bank credit conditions tighten, less credit is provided in the marketplace, which in turn is likely to reduce both investment and consumption levels. As these levels are reduced we expect to see a negative short-term effect on GDP, and the NIBOR 3m is therefore thought to have a negative correlation with GDP.

The opposite is true for a decrease in the Nibor 3m rate. If the rate is reduced, short-term credit conditions loosen, making credit available at a lower cost of capital. As a result, short-term investment and consumption effects are expected, both components in the Keynesian model for GDP estimation. Another plausible effect of a decrease in the Nibor 3m can be an increase in inflation due to the increase in available money and purchase power. The effect of a decrease in the Nibor 3m rate on GDP is likely to be positive, as long as the inflation is lower than the Nibor 3m, which it always is. If it was not, no lending would occur, as profits would be diminished by the inflation and the banks would incur a loss.

5 year Spread between short and long-term interest rate

The long-short spread is the difference between short-term interest rates and the long-term interest rates, as indicated by Schiller and Campbell (1990).

Long-term interest rates are interest rates paid on securities with maturities exceeding one year. The long term interest is decided by several factors, but mainly the short-term rate and an error term. The long-term rate is a possible indicator of long-term assumptions about economic stability and interest rate levels in general. If the long-term interest is low, then there is little fear of future shocks, but shocks can alter the rate in both directions. We believe long-term interest rates give a good indication of how analysts and banks perceive the future. The influence that the long-term interest rate alone might have on GDP is assumed to be relatively small. However, the indications it gives with regards to future interest rates could be an early warning of possible despair to come.

The long-short spread tells us something about the difference in expectations for the short and the long term, and expectations about the long-term stable interest rate. If the spread is small, it indicates that the short-term interest rate is close to the perceived long-term interest rates. Gerdrup et al.'s 2006 article suggests that the 5 year spread has a significant impact on GDP, and is a good leading indicator. We have therefore chosen to include it.

Alternatively we could have included a spread with a shorter horizon, but we chose not to, as this had no previous support in the literature we surveyed.

3.3.1.3 Prices: The Oil Price, House Prices and the Real Exchange Rate

Traditionally, the price of Brent Crude is highly correlated with the Oslo Stock Exchange All-Share Index (OSEAX), which can cause problems in the development of the FCI, had we been unaware of it. Real Estate prices also contain effects from many different variables, such as interest rates, inflation and the general variations in household wealth (Goodhart and Hofmann 2001). Consequently, it can be sensitive to changes in several of the other input variables we have considered. The Oil price may also suffer from such correlation issues.

The Oil Price

The price of oil in this context is the price of Brent Crude from the North Sea traded at London International Financial Futures and Options Exchange (LIFFE). The price of oil is determined by a wide variety of conditions all over the world, and therefore contains a lot of information. It is out of the scope of this thesis to go further into what comprises the oil price, but Mabro (1992) will enlighten those with a particular interest. Instead, we concentrate on its effects on GDP.

Due to all the information contained in the price of oil, it can serve as leading indicator with high informational value. Because Norway is a net exporter of petroleum products, the usual effects of changes are opposite to most other European countries, which are net importers. If the price increases it is likely to have a net positive effect on GDP, both in terms of increased profits and profit expectations from the oil sector of the OSEAX, and in terms of government wealth. However, due to the risk of increased inflation, heightened profits from the government petroleum production are not excessively pumped into the economy through increased government spending, such as it is in many other countries with a small population and vast oil production. We therefore expect the petroleum price fluctuations to have moderate to small implications on government spending related effects on GDP. We chose not to include the oil and gas producing sector in basis GDP observations, but because the oil price has an impact beyond that of the producing sector and is an input in a majority of production activities, we believe that including the oil price is not a controversy.

OSEAX

The OSEAX is adjusted for dividends, and contains all stocks traded at the exchange[16]. A stock price can be seen as an incorporation of all expectations regarding both the future profitability of any given firm and firm specific future interest or discounting rates. A stock price is therefore an aggregate expectation function for the future of that firm, and the index is the sum of the expectations for all firms traded, at any given time. Stock prices can be useful indicators of what financial markets and institutions expect in the future. Based on the EMH, stock prices should contain all private and publicly available information of all stocks traded at the OSE. We can therefore make use of all the analysis made by professionals and others, which should be incorporated in the index as well as in the price of any given stock, at any given time.

According to Hatzius et al. (2010) equity prices are one of the most common financial variables to include in an FCI, and its U.S. equivalent – the S&P 500 – is included in the Conference Board's index of leading indicators. However, as an indicator it loses its accuracy when the horizon of the predictions exceeds nine months, according to Estrella and Mishkin (1998).

The correlation and effect of OSEAX fluctuations on GDP is likely to be positive; if the OSEAX increases over a quarter, we expect to see an increase in the GDP for that same quarter as well. The interpretation of an increase in the OSEAX is that it bears news of a brighter future, and we would -therefore expect an increase in GDP, caused by investment and consumption.

The Real Estate Index

The Norwegian Real Estate Index[1] is estimated as a combination of the sale price of individually owned real estate, and real estate owned by co-operatives. It contains price information regarding sale values of real estate sold in the open Norwegian real estate market. The index is a measure of NOK per m^2, and signals not only the level of demand for Norwegian real estate, but also the population´s expectations with regards to interest rates in the long term.

The current source of all input data in the index is the Norwegian web site Finn.no. However, in the past, inputs were collected from several different sources. Due to the many real estate indices that are regularly published by several different organizations, we considered the index estimated by SSB, as it is both readily available and contains the largest public historical database. Two points were important though: 1) The index will overvalue housing as an investment over time, as it does not exclude the increased value of renovations and refurnishing 2) The input data from Finn.no only contains price information of about sixty percent of all traded real estate[17]. An increase in the index could therefore signal a change in one or several conditions:

[1] SSB's House Price Index

[2] The notation is taken from Hatzius et al. (2006)

- A loosening of credit conditions or a change in the bank´s lending practices
- An expectation or a plain reduction in inflation
- A lower interest rate on long-term loans
- Increased exchange rates and a strong economy
- Bright expectations for the future, such as a steady rise in wages and stability in the economy as well as long-term increase in the m^2-price of housing.

We considered including house prices in our FCI not only because amongst others Hatzius et al. (2010), Estrella and Mishkin (1998) and *Deutsche Bank´s Financial Condition Index* all includes it, but we also believed that real estate prices probably are a suitable signal of general expectations, especially regarding private consumer long-term conditions in the overall economy.

The Real and Nominal Exchange Rate of the Norwegian Krone (NOK)

The nominal exchange rate is the NOK's value measured against any other currency. The real rate of the NOK is the nominal rate after inflation adjustment. The rates are decided as an aggregate of supply and demand for the NOK, set in the international currency exchange market[18].

These rates are important for exports and imports, and the exchange rate towards the main trade partners of Norway determines not only the balance of imports and exports, but also the amount of goods being imported or exported. The exchange rate has served as an indicator of GDP expectations for a long time, not surprising as the trade balance is included as a variable in the Keynesian GDP estimation.

If the value of NOK increases, exports will become more expensive for the buyers of Norwegian goods and services, and so aggregate demand is likely to fall (Svensson 2000). If the increase is sustained over a period of, say 2 quarters or more, it is likely to have a negative effect on the employment levels in the exporting sector. The importing sector is likely to see positive effects as foreign goods and services become relatively cheaper. The balance of imports and exports will be the final determinant as to whether the net effects on GDP are positive or negative. Because the Norwegian petroleum products are priced in dollars, the exchange rate between NOK and USD is especially important.

The opposite is true for a decrease in the exchange rate. If the value of NOK is reduced, goods and services to be paid in NOK will become cheaper in the market in question, and it is therefore likely to incur an increase in aggregate demand. This may result in a positive effect on employment levels in this sector, especially if the decrease is sustained for some time. The importing sector is faced by increased prices, and so again, the trade balance is the determinant of the final effect on GDP.

Evidently, the conditions for the sector of the economy whose survival is dependent on either exports or imports, such as the Norwegian salmon exports, can be vulnerable to currency fluctuations. The Real Exchange Rate was among one of the two indicators included in the MCIs and is included in nearly all FCIs. Gerdrup et al. (2006) also found the Real Exchange Rate to be leading for GDP changes in Norway. Naturally, the Real Exchange Rate was considered as valuable input in our Norwegian FCI.

3.3.1.4 Monetary indicators

We believe all money supply and credit growth indicators could serve as indicators of future economic activity. Since the Money Supply indicators, M1 and M2, and the credit growth indicators, C2 and C3, probably are highly correlated, including all of them could violate the OLS assumptions. In the study by Gerdrup et al. (2006) M1 and C2 were found as good indicators for GDP. On this background we chose to include only M1 and C2 in our FCI. All indicators are however highlighted in the following sub-sections.

Money Supply (M1 and M2)

M1 measures NOK held in cash and bank accounts by the population and financial institutions, other than banks and government lending institutions. M2 also incorporates unrestricted bank deposits and bank certificates, and is usually referred to as the population´s liquidity. All funds that can be accessed immediately and converted to cash without any extraordinary costs are included in these measures.

High values of the monetary indicators would mean less spending (the funds are held or deposited) and signals a slower circulation of funds in the overall economy. This has a negative impact on both private and public spending. High values could also be a signal of worsened expectations about future economic activity. In addition, high values can indicate that inflation and interest rates are high, and that owners of capital lack higher yielding investment opportunities. Consequently, it might impact GDP growth negatively.

If the monetary indicators decrease, it can signal positive expectations about the future, increased spending and a higher circulation speed of funds in the economy. That in turn, it expected to have a positive effect on GDP. On the other hand, it could also mean that higher-yielding investment opportunities have now come along, and that the population is re-allocating their funds to these higher yielding opportunities. This argument also works in the opposite direction: the population might be losing their faith in the banks, and are withdrawing their funds to be stored out of the bank´s reach. The last possibility is that it can signal a shift in the population´s priorities or preferences, from consumption in the future to consumption in the present. This possible cause of change is expected to have a positive impact on GDP as well.

Norwegian Domestic Credit Growth (C2) and Norwegian Credit Growth (C3)

C2 is an indicator of the growth, decline and size of domestic debt, held by households, non-financial businesses and public sector county administration. C3 measures gross debt, in other words total domestic and foreign dept hold by households, non-financial businesses and public sector county administration. The debt is held in NOK and foreign currency, in domestic as well as foreign banks [19]. C2 and C3 numbers are published monthly, within one month and eight to nine weeks after the statistic month respectively[20].

Growth would be perceived as a positive sign of an expanding economy, and positive expectations for the future. The possibility of a correlation with the real estate index means we draw some of the same conclusions with regards to the effects of fluctuations. If C2 increases, expectations of wage increases, stability in the economy and future profitability, is positive. Expectations for CPI are also likely to be positive or unaltered, although most likely positive, where positive means prices will not increase. For these conclusions to hold however, we do need to assume that the borrowers intend to pay their loans at the time of borrowing.

As the population becomes more global-minded, we assume their financial decisions do as well. Technological improvements such as the internet and improved cell phone coverage has made it considerably easier to trade and make cross-border (or even cross-continent) transactions. An increase in the C3 without a replicating increase in C2 would imply that Norwegians are borrowing more abroad, which could have a positive effect of GDP, if the borrowed funds are either spent in Norway or spent buying Norwegian goods and services abroad. There are of course several reasons why Norwegians would prefer to borrow abroad, which we will not treat in this thesis.

The final indicators to be included in our FCI are listed in the table below. Sources are listed on the right hand side. Each indicator has data history from 1980(1) until 2010(4).

Table 3.3 Final Indicators to be included in our FCI

Number	Description	Source(s)
	Interest rates:	
1	3 month NIBOR	Ecowin/Norges bank
2	5 Yr Treasury Bills/3 month NIBOR	Norges bank
	Prices:	
3	Oil price USD	Ecowin
4	OSEAX	Ecowin/Oslo Børs
5	Real Estate Index	SSB
6	Real Exchange Rate	OECD
	Monetary indicators:	
7	Real M1	SSB
8	Real C2	Ecowin/Norges bank

3.3.2 Possible Connections Between Financial Indicators and GDP

In this section we discuss two indicators which we considered to include in our FCI, but were not included due to a variety of reasons. We believe it is important to highlight these indicators as they can provide valuable information for researchers in order to select financial indicators to include in an FCI.

3.3.2.1 Norges Bank's Survey of Bank Lending Practices

Several surveys on how to capture the tightness of the credit market have been conducted, especially in the U.S. Credit as a means of payment is increasing, and if lending practices become stricter less credit will be offered in the market. When less credit is provided in the market, both consumption and investment levels are likely to decrease, which will result in an increase in the price of credit. This might result in an increase in the rate of default on loans. Assuming that borrowers intend to pay their loans, a reduction in the demand for credit bears information about negative expectations for future economic growth. The effect of defaults and tighter credit conditions on GDP is likely to be negative.

In the FCI by Guichard and Turner (2008) the Senior Loan Officer Opinion Survey on Bank Lending Practices conducted by the U.S. Federal Reserve was included. It provided evidence of the tighter credit conditions in the economy. The survey, which was also included in a broader measure of lending standards in the FCI by Hatzius et al. (2010), is conducted by interviewing senior loan executives from approximately sixty large U.S.-based banks and twenty four U.S. branches and agencies of foreign banks.

The Federal Reserve has conducted the survey quarterly since 1997, in which questions cover changes in the standards and terms of the banks' lending and the state of business and household demand for loans.
The sub-prime loans in the U. S. that triggered the financial crisis had underlying securities in other banks in other countries and we believe there is a correlation between the U.S. credit market and other foreign credit markets. It was also proved by Guichard, Haugh and Turner (2009) that similar studies to the one conducted by the Federal Reserve about Bank Lending signaled the tighter credit conditions in the U.K., Japan and the Euro Area.

Norges Bank's Survey of Bank Lending provides qualitative information about demand, supply and terms of new loans in Norway. We believe this survey could provide useful information about the tightness of the credit market in Norway, and hence contribute with explanatory power for GDP changes in Norway. Unfortunately, this survey has only been carried out since 2007 so the data history is too short to be included as a variable in our index. For researchers developing FCIs containing unbalanced panel data, it could be useful to include this indicator.

3.3.2.2 High-Yield Bond Spread

The spread is the difference between yields from high-risk high-yield bonds and government bonds (or other benchmark) with the same maturity. Credit spreads in high-yield bonds can be a good indicator of the least risk averse debt holder's probability of default. This is because credit spreads usually increase when investors are concerned about the quality of corporate debt and the future of the corporation in general. Risk is determined by the rating of the security, for example between BAA-rated bonds, and AAA-rated bonds[21]. If the risk of default by the issuer is high, so too will the risk premium or spread be. The first to default in a potential crisis is the risky firms, and so a high spread is probably a good early warning indicator of financial troubles ahead. An increase in the spread is therefore likely to affect GDP negatively. A reduction in the spread might not yield short term effects, but as financial conditions improve we expect the reduction to affect GDP positively.

According to Hatzius et al. (2010), this variable is also fairly common and usually contains high informational value. Unfortunately, and to our knowledge, there are no similar spreads in the Norwegian bond market. We could therefore not include this indicator in our FCI.

4 Analysis Assumptions

To make as precise forecast estimates as possible and to draw any conclusions from the final results we saw it necessary to make some assumptions. We start by presenting the standard model setup, weights, and lag lengths in section 4.1. We then proceed with a discussion of the assumptions regarding OLS estimation and OLS definition tests in section 4.2. In section 4.3 we discuss the most common measures of predictive preciseness. Finally, in section 4.4 we discuss several benchmarks, to which we will later compare our FCI.

4.1 Model Setup

The most systematic estimation approach in our view was to start with a static model without lags. According to EMH, all indicators should be continuously updated and contain all available public and private information.

The empirical research we surveyed contained evidence that some indicators' lagged values provide essential information. Thus, including several lags of each variable could improve the forecasting performance. Based on the same EMH argument and the literature we surveyed we thought that including more than four lags was unnecessary, and would contribute more in terms of complexity than explanatory power.

As previously discussed, two main approaches have been used in the construction of previous FCIs to determine the weights. Despite advantages associated with the PC approach, Ziegler and Eickmeier (2006) suggested that this approach were better suited for the U.S. rather than the Euro area. Based on this argument, and VAR models ability to deal with both lagged values of the dependent variable and the explanatory indicators, we chose to use a weighted sum VAR model. In addition VAR models are simple to put into practice. VAR models are however limited in that they can only handle a small number of explanatory variables simultaneously. In the single indicator models, one indicator's single effect on GDP growth is examined at a time. Thus, the limitation will not be a problem. Since we only include eight indicators in our FCI, problems associated with degrees of freedom are not substantial.

To analyze the various single indicators' and our FCI's performance we first constructed a static model which was then expanded to a dynamic VAR model including several lags. In total we established five model versions for each single indicator and our FCI, including from zero to four lags.

In all models, the following remains unchanged:

- Lagged values are denoted P_x, so that $P_x = \{0,1,2,3,4\}$[2].
- The models contain a constant (α)[3] to capture possible exogenous factors
- The models include an error term (e_t) to deal with possible noise
- Changes in GDP at time t are denoted ΔY_t

4.1.1 Model setup for the single indicator's analysis

For each single indicator the following static (1) and dynamic (2) equation models were developed:

$$\Delta Y_t = \alpha_i + \sum_{i=1} \gamma_t \Delta X_t + e_t \tag{1}$$

Where; ΔX_t denotes changes in the single financial indicator at time t and γ_t denotes the weight attached to each indicator. The static model includes no lagged values.

$$\Delta Y_t = \alpha_i + \sum_{i=1}^{Px} \gamma_i \Delta X_{t-i} + e_{t-i} \tag{2}$$

The difference between the equations is the addition of lagged changes of the single indicators in (2) so that ΔX_t becomes ΔX_{t-i}. The weight attached to each indicator is also lagged γ_i.

4.1.2 Model setup for our FCI

We constructed the following static (3) and dynamic VAR model (4) for our FCI:

$$\Delta Y_t = \sum_{i=1} \vartheta_t FCI_t + e_t \tag{3}$$

The ϑ coefficient represents the weight attached to the FCI_t. The FCI_t parameter represents an equally weighted sum of the various indicators in table 3.3 in chapter 3, and a constant term. The static FCI include no lagged values.

[2] The notation is taken from Hatzius et al. (2006)

[3] The excel spreadsheets on the CD provided with this thesis gives a precise elaboration how all calculations were conducted.

$$\Delta Y_t = \sum_{i=1}^{Px} \vartheta_i FCI_{t-i} + e_{t-i} \quad (4)$$

The only difference between equation (3) and (4) is the addition of lags in (4), both lagged values of the weights ϑ_i and the FCI_{t-i}. Lagged GDP changes and changes in the various indicators are also included.

4.1.3 Determining the weights – the weighted sum approach

As discussed initially in this chapter, we employed the WS approach to estimate each indicator's weight for both the single indicators and our FCIs. Our *mode d'emploi* for this analysis was somewhat inspired by the out-of-sample analysis conducted by Hatzius et al. (2010).

In determining the weight attached to each **single indicator model's** variable, we regressed the basis period 1980(2)-2005(4) on ΔGDP. In both the static and the dynamic model these weights were kept constant over the entire period. Since we used changes rather than levels, the basis period started in the second quarter of 1980. The weight for each indicator was then multiplied with each quarter's observation in the sub-period 2006(1)-2010(4).

In the dynamic models the basis periods 1980(2)-P_x-2005(4) were lagged, so for example, by including two more lags the basis period would not start until 1980(4). The observations were lagged as well, so that the weight for each lag was multiplied with the lagged values of the indicator: *Coefficient*$_i$ *x Observation*$_{t-i}$. We estimated different weights for each indicator, depending on its lagged value, and this was conducted for each lag and quarter respectively.

We used the same approach to determine the weights attached to the indicators contained in our FCIs, thus estimating coefficients based on the period 1980(2)-P_x-2005(4). The weight for each indicator was multiplied with its respective quarterly observation into sub-indices. These sub-indices were then equally weighted, and summarized into an FCI.

4.2 Violating OLS assumptions

In this section we survey some of the assumptions underlying OLS estimation. This is of special importance in order to properly interpret our test results, especially t-tests and F-tests.

The R-square value shows how well the model has explained the overall variance in the basis period. To account for the weights' overall significance we estimated F-statistics. The F-statistics shows the total significance of all indicators combined. Corresponding t-values were estimated to account for each indicator's individual effect on changes in GDP. The critical values were chosen based on a 95 % confidence level[4]

4.2.1 Testing for Model specification and wrong functional form

One underlying assumption of the OLS specifies that the model needs to be "correctly specified" or else there is a chance of conducting spurious regressions. Including or omitting (ir)relevant variables, using the wrong functional form, measurement errors, incorrect model specification or non-normality in the error term are all types of errors that can arise (Gujarati and Porter 2009). The consequences and how these errors should be detected differ depending on the nature of the problem. Our explanatory variables were mainly chosen on theoretical grounds, and our intention was to include only variables that were thought to have an effect on changes in GDP, and thus not accounted for by other indicators. One should however not discard potential inhibit errors in the model.

To test for model specification we used the RESET test, basicly because it is easily applied. A disadvantage with the test however, is its lack of capability to specify alternative models if it turns out that the model is mis-specified. In other words, the test suggests that something is wrong but does not give clear signals on why and where the model is wrong. If the computed RESET test scores were statistically significant at a five percent level, we accepted the hypothesis that our models were mis-specified.[5]

When dealing with multiple regression analysis, such as our FCIs, we go beyond the world of linearity, thus violating the OLS assumption that the parameters or the regression models must be linear. We

---

[4] We defined the following null hypothesis: $H_0: \emptyset_i = 0$ against the alternative H_A: $\emptyset_i \# 0$

Where: $i =$ Numbers of lags P_x and in the F-statistics $\emptyset_i = \alpha_i = \gamma_i = \vartheta_i = 0$

If the t and F-values were significant at a 5 % level we rejected the alternative hypothesis

[5] *RESET Test*

We specified the null hypothesis H_0: no mis-spesification against the alternative H_A: mis-spesification
From our specification models:
(1) $\Delta Y_t = \alpha_i + \sum_{i=1} \gamma_i \Delta X_t + e_t$
(2) $\Delta Y_t = \alpha_i + \sum_{i=1}^{px} \gamma_{t-i} \Delta X_{t-i} + e_{t-i}$
(3) $\Delta Y_t = \sum_{i=1} \vartheta_t FCI_t + e_t$
(4) $\Delta Y_t = \sum_{i=1}^{px} \vartheta_{t-1} FCI_{t-i} + e_{t-i}$
We obtained estimated changes in GDP, $\Delta \ddot{Y}_t$, from the models (1), (2),(3) and (4) respectively. The equations where then reformed to include $\Delta \ddot{Y}_t$ as a right hand side variable, so that model (1) became: (1a) $\Delta Y_t = \alpha_i + \sum_{i=1} \gamma_i \Delta X_t + \gamma_2 \Delta \ddot{Y}_t + e_t$
Similar reformations were utilized for model (2),(3) and (4). The R^2 values from the new equations became the new R^2 values.
We conducted the F-test = $\frac{(R^2 new\ R^2 - old)\ number\ of\ new\ regressors}{(1-R^2 new)/(n-numbers\ of\ parametes\ in\ the\ new\ model)}$
If the computed F values were significant at a 5% level we rejected the alternative hypothesis that our models were correctly specified.

therefore wanted to test the hypothesis that our regression models might not be linear. Since we also deal with FCIs with a dynamic structure, concepts such as equilibrium solutions, steady-state growth paths and mean lags of response are of interest. To make predictions with our FCIs these issues should be accounted for. We therefore utilized a Wald test to see if our models had a long-run solution[6]. This test was conducted only for the dynamic FCIs.

4.2.2 Normality

The normality assumption states that factors not included in the model are captured by the error term and do not affect the dependent variable (Gujarati and Porter 2009). This implies that the model does not suffer from specification bias or specification error and that the regression model is correctly specified. If the computed p-values are sufficiently low (< 0, 05) the normality assumption is rejected. Consequently, corresponding results of the t and F tests are not reliable[7].

4.2.3 Serial-and Autocorrelation

The assumption of no autocorrelation is often violated when dealing with time series data. We employed a Durbin Watson to test for autocorrelation. Our FCIs dynamic models also exhibit lagged values of the dependent variable meaning that they are not exogenous. Autocorrelation could therefore not only be a problem between the independent variables, but also between lagged values of the dependent variable. These effects were accounted for by conducting AR (1-5) tests.

Volatility is not an unusual phenomenon in time series data such as stock prices and exchange rates, and this non-constant variance may be autocorrelated. We tested for this type of correlation using an AutoRegressive Conditional Heteroscedasticity test (ARCH1-4). Autocorrelation may however not be pure autocorrelation; it can be caused by model mis-specification such as wrong functional form and omitted variables. The ARCH test served as an additional test to examine whether autocorrelation is due

[6] *Test for long-run static solution*
We tested the null hypothesis H_0: All long-run coefficients are equal to zero, except the constant term, against the alternative H_A: all long run coefficients are different from zero i.e. a long run static solution. Significant Wald test values indicate that we have a static long-run solution and so we discard the null hypothesis.

[7] *Test for Normality of ε_t*
If normally distributed, then for model (1) we have: $E(\varepsilon_t|X_i) = 0 = E(Y_i|X_i) = \alpha_i + \sum_{i=1}\gamma_i \Delta X_t + \varepsilon_t$. The same applies for the other models. Normality in ε_t also assumes no covariance between variables, and hence, the assumption implies that the second assumption $Cov(\varepsilon_t, X_i) = 0$ holds. Moreover, any linear function of normally distributed variables is itself normally distributed, so that if $\varepsilon_t \sim N(0,\sigma^2)$, then $\hat{Y}_1$ is normally distributed (Gujarati 2009). We defined the null hypothesis H_0: Normality, against the alternative H_A: no Normality. We applied a Multivariate Normality test for small sample sizes that tests if skewness and kurtosis of the error term correspond to those of a normal distribution (http://www.pcgive.com/pcgive/index.html?content=/pcgive/volume2.html 29.04.2011 13:46). If the computed p-value were less than 0,05 we rejected the null hypothesis.

to pure autocorrelation or mis-specification of the model, if autocorrelation is present in the Durbin-Watson test[8]. As with non-Normality, if autocorrelation is present the usual test statistics such as t-tests, F-tests and chi-square tests, cannot be legitimately applied, and often the variance is underestimated.

4.2.4 Multicollinearity

Since we have included several lagged values in our dynamic models it might turn out that these indicators are correlated, thus recalling the OLS assumption no multicollinearity. Especially in time series data trend values are not unusual. As suggested earlier, the GDP showed tendencies of seasonal fluctuations, suggesting that there might be a trend. If multicollinearity is present precise estimation is difficult because the explanatory variables X_t would explain the same, thus one will have problems identifying the indicators individual effect on Y_t . Consequently, the standard errors and the variance of the parameters go infinite and this is a major problem since the entire intent is to capture the single effect. The confidence intervals also tend to be wider, increasing the chance of type II error. In addition, the t values are often non-significant and the overall R-square values are high. As long as there is no perfect multicollinearity, estimation is possible, but the estimation results are very sensitive to changes in the data set[9].

4.2.5 Heteroskedasticity

Even though cross sectional data are more exposed to heteroscedasticity, the variance of e_t in time series data may also be variable because of data collection techniques, outliers or model specification errors. No constant variance in e_t or homoscedasticity means that the Y_i corresponding to the different X_i values has the same variance. If this is not true we have a case of heteroscedasticy, which will affect the estimation and the test values negatively, so that the result may not be reliable. To test for

[8] *Test for autocorrelation*
We stated the null hypothesis H_0: zero autocorrelation (ρ=0, so DW=2) against the alternative H_A: positive first-order autocorrelation (ρ>0) and tested for autocorrelation by using the Durbin-Watson (DW) d statistic. If the computed value was [2<DW<4] then H_A: negative first-order autocorrelation (ρ<0) and so the DW value should be computed as 4-DW.

Our dynamic FCIs (with P_x=1,2,3,4) included lagged values of the dependent variable, changes in GDP. Consequently, the DW test cannot detect autocorrelation. To test whether lagged variables of Y_t could be included as explanatory variables X_t and so account for potential correlation that could exists between the explanatory variables and the lagged residuals, we utilized Durbin's h-test for autocorrelation of higher order, AR (1-5). If the computed p-values of the F-test were less than 0,05 autocorrelation was recorded.
Similar, if the computed p-values in the ARCH1-4 test were less than 0,05 we kept the alternative hypothesis.

[9] *Detection of Multicollinearity*
Multicollinearity is primarily a sample phenomenon and there is no unique method of detecting it. We used the following rule of thumb: if the R^2 values were very high ($> 0{,}8$) but there were few significant t-values, then multicollinearity was a serious problem.

heteroscedasticity we applied the White test[10]. However, since the test fails to distinguish between pure heteroskedasticity and specification errors in the cases where the test statistics are significant, we conducted an additional test for heteroskedasticity in the error terms (the hetero X-test).[11]

4.2.6 Stationarity

Non- stationary and stochastic trend values are commonly observed in time series. Asset prices, such as stock prices and exchange rates, are often said to follow a random walk, i.e. they are non-stationary (Gujarati and Porter 2009). Regressions based on non-stationary time series are spurious, and conclusions drawn on the basis of their behavior in the particular period are only valid in the same period. Hence, one cannot generalize the results to other time periods. Since our FCI was invented for forecasting purposes generalization to other periods with regards to validity and causality is an absolute requirement. To avoid generalization issues associated with non-stationary time series, they need to be differentiated so that they become stationary. Since we primarily dealt with non-stationary time-series, all our series were differentiated once (so that $X_t \sim I(0)$ = non-stationary goes to $\Delta X_t \sim I(1)$ = stationary).

We could have tested our time series for unit-root using the augmented Dickey-Fuller test, but as Perron (1989) argues, macro-economic variables are not unit-root processes; they are stationary series with a trend and structural breaks. The crisis of 2008-2009 is one example of such a structural break. If the test for unit-root processes was conducted, it would be biased towards the acceptance of unit-root in the series.

4.3 Measures of Forecast Preciseness

To test our model's predictive preciseness we considered several different measures because they contain various properties. Using such measures separately could lead us to a faulty decision in selecting the best model. Also, for the application of the t and F statistics, the measurements should be considered together to better evaluate our models' performance.

---

[10] *Test for heteroskedasticity*
We stated the hypothesis H_0: constant variance against the alternative H_A: no constant variance (heteroscedaticity). From our original models we obtained $\hat{\varepsilon}_t$ and ran auxiliary regressions, so that model (1): $\hat{\varepsilon}_i^2 = \alpha_i + \sum_{i=1} \gamma_i \Delta X_t \dots + \alpha_3 X_{2i}^2 + \alpha_4 X_{3i}^2 + \alpha_n X_{ni} X_{ni} + v_i$
If the White test's computed p-values were less than 0,05 we rejected the null hypothesis.

[11] This test is a general test for heteroscedastic errors and is conducted based on auxiliary regressions of the squared error term on all squares and cross-products of the original variables. (www.pcgive.no) If heteroscedasticity is present, it is related to the error terms and not the variables.

To find measures as precise as possible for our purposes, we considered an article by Armstrong and Collopy (1992) where they compared the six measures – the Root Mean Square Error (RMSE), Percent Better, the Mean Absolute Percentage Error (MAPE), the Median Absolute Percentage Error MdAPE, the Geometric Mean Relative Absolute Error (GMRAE) and the Median Relative Absolute Error (MdRAE) according to reliability, construction validity, outlier protection, sensitivity and their relationship to decisions. The measures were rated as good, fair or poor. None of the criteria were found outstanding e.g. that they achieved the "best rating" within all the five criteria.

For the first criterion – Reliability – the authors examined the extent to which an error measure would produce the same accuracy rankings when it was applied to different samples. Here, Percent Better was the only good measure, and all the others were found reliable except for RMSE.

To evaluate the next criterion – validity – the authors compared forecasting methods for each of the error measures and examined the extent to which the error measure measured what it was supposed to measure. All the measures were reasonably accurate when the number of series was fairly large. What is "fairly large" is however a subjective decision. MdRAE was found most reliable when small sets of data series were available.

Regarding the outlier criteria, RMSE was poor, probably due to its lack of ability to differentiate between whether the model have under- or overestimated the true value. MAPE is impaired because it gives relatively more weight to forecasts that deviate largely from the mean, thus serving as a poor measure for the outlier criteria. MdRAE was less affected by outliers and was rated as good, primarily due to its ability to control for differences in scale and sensitivity, by means of the amount of changes that occurs over the forecast horizon. MdAPE also served as a good proxy.

Regarding sensitivity and decision making, RMSE was rated as good in both cases, and was actually the only good measure for decision making. RMSE was not ideal for aiding decision making, although it was preferable by most decision makers compared to other percentage measures. MAPE and GMRAE were also rated as good measures for sensitivity, but regarding decision making, they only achieved scores of fair and poor respectively. MdRAE was a poor measure for sensitivity. Except for the sensitivity criteria, all the criteria were important in the selection of forecasting preciseness measures.

The reliability criteria were important for us as it addressed whether a repeated procedure would produce similar results, and we chose RMSE for our measure of preciseness. RMSE has also been widely used for comparing forecasting methods, and so a reader with "limited" technical insight should still be able to relate to this measure. Even though RMSE is not unit-free, as we dealt with changes rather than levels, this was not a concern for us. We also hoped to have an ideal measure for aiding decision making, and due to the RMSE's "good" test score in the article mentioned above, we assumed it would serve as an adequate measure.

We also found MAPE useful as it was rated "good" regarding sensitivity and construction validity and thus accounted for some of RMSE's limitations in measuring validity. We saw it as desirable to have an error measure to reveal as much of the effects of changes over time. MAPE is only relevant for ratio-scaled data, and since we only dealt with such data this was not a concern in our case. By using two measures we believe all the criteria are adequately covered and that the limitations associated with each criterion are accounted for. For each regression we estimate MSE, RMSE and MAPE. We will however put more weight on RMSE in the final model selection, because we have used the OLS approach in our regressions. The estimation method's emphasis on minimizing the sum of the squared errors $\sum \varepsilon_i^2$ means that RMSE is more aligned to the regression approach, and hence a more suitable measure than MAPE.

4.4 Benchmarks

In order to assess the performance of our FCI it is important to compare it against some established benchmarks. Single indicator predictions have been used heavily by researchers to compare their own models' preciseness in predicting. When forecasting a variable of interest, one is usually faced with a broad set of candidates and it not always easy to recognize the right candidates. Not only can the single variables be based on different information, it can also differ in terms of how that information is used. Nowadays, forecasts based on a broader information basis have emerged as more feasible. Improvements in computing and more available data are primary reasons for the growth of these "pooled forecasts". According to Ziegler and Eickmeier (2006) *"Pooled single indicator forecasts"* and forecasts based on *"best single indicator models"* or "*groups of indicators"* are often used as benchmarks. Generally, FCI forecasts resulting from the principal component approach has been compared to basic weighted sum AR-models' forecasts.

We compare our FCI against five single indicators, two previously established FCIs and a naïve model. A detailed analysis of the single indicator models is presented in chapter 5, but first we present the models underlying OECD and Goldman Sachs predictions for GDP changes in Norway and a simple Naïve model.

4.4.1 Goldman Sachs FCI's Predictions for Norway

GS has provided us with GDP predictions from 2005 until 2012. These predictions were given both one-year-ahead at the end of the previous year, and (one) quarterly updated to the one-year-ahead. Graph 4.1 gives a picture of GS GDP predictions for Mainland Norway one year ahead starting in 2004 until 2011.

Graph 4.1 GDP Mainland Norway and GS predictions 2004-2011

From the graph it seems like GS in 2006 to some degree predicted the downturn in Norwegian GDP that appeared in 2007. When predicting Norwegian GDP changes, GS utilize a global forecast model, which does not account for local specifics. The fact that GS underestimated the effect of the crisis, in that they predicted an upswing in mid 2007 which did not appear before mid 2009, is therefore not surprising.

4.4.2 OECD Projections for GDP changes in Mainland Norway

OECD provided us with quarterly GDP predictions for the period 1970(1) - 2012(4). These projections are adjusted for inflation, i.e. they are in real terms. In their predictions, OECD claims that they distinguish themselves from other economic forecasts. First off all, they focus on framing the policy debate in the chosen country by answering questions like: if the government in Norway implements mandated fiscal measures, what is likely to happen? OECD claims that such questions can achieve better predictions because their likelihood of identifying potential problems in the economy is greater. Secondly, they ensure that the projections are consistent at a superior level compared to rest of the world. Finally, they benefit from government and policy makers' participation and expertise in arriving at its projections (www.oecd.org/).[22]

To overcome sensitivity problems in their predictions OCED use a chain-weighted method so that the base year is continuously moving forward. This means that real GDP is non-additive. In other words, GDP is not equal to the weighted sum of its indicators, except for the benchmark year, and this requires caution. Additivity however, holds in growth terms and OECD utilizes a dynamic weighting structure that takes into account relative changes of each indicator, so that real GDP growth is computed as a weighted average of the growth of its indicators. [23]

4.4.3 A Naïve model

The naïve model assumes that history repeats itself, so all dependencies between variables exist due to the causal relationships between the variables in the model. Therefore, the possibility that we have excluded some variables causing dependencies in our model is denied. In reality, this assumption is rather naïve, and accordingly the model has its name.

In time series, the naive model extends the latest observation. We developed a simple Naïve model assuming that changes in GDP in the current period are the same as changes in GDP in the previous period.

5 Testing the Predictive Power of the Single Indicators

In this section we analyze the predictive performance of five single indicators. These indicators were selected because they, according to Gerdrup et al. (2006), had a significant effect on GDP growth in Norway between 1990 and 2006. This selection was also inspired by Hatzius et al. (2010) who stated that a spread, a money supply indicator and a Share Price Index are classified as leading indicators by the Conference Board's in the U.S. The chosen indicators are:

1. The spread between 5 year Treasury bonds/NIBOR 3m
2. Real Exchange Rate
3. OSEAX
4. Real Credit Growth
5. Real Money Supply Growth

5.1 Out-of-Sample Results

This first part of the analysis has one main divide, namely between model construction and predictions. We first present the properties of the single indicator models that were estimated in the basis period. For all models the weight of each lag and the constant term is presented. We conduct an F-test, a Durbin-Watson test and present the degree of fit to the model estimation basis period. The models are then tested for model specification error, auto- and serial correlation, heteroscedasticity and normality. Test results significant at a 95% confidence interval are marked with one star [*], while test results at a 99% significance level are market with two stars [**]. The models' degree of prediction preciseness is chosen on the lowest MAPE and RMSE values respectively. Where the results differ, we present both possibilities.

5.1.1 Single Indicator Model Properties

In the analysis of the single indicators, the results were disappointing. We hoped for a clear relationship and correlation between the different variables and GDP, but as table 5.1 shows, that was not the case:

Table 5.1 Properties of the Dynamic Single Indicator Models with Four Lags

SINGLE INDICATOR PREDICTION MODELS WITH FOUR LAGS					
Variable	**5 Year Spread**	**Real Exchange Rate**	**OSEAX**	**Real C2**	**Real M1**
Constant	-0,005 (-4,57)	**-0,006** **(-4,92)**	-0,005 (-0,001)	-0,005 (-4,73)	-0,005 (-4,76)
γ	-0,0004 (-0,915)	**0,167** **(-2,76)**	0,001 (-0,011)	-0,002 (-1,21)	-0,002 (-1,18)
γ_{t-1}	0,0003 (-0,688)	-0,021 (-0,34)	-0,012 (-0,012)	0,001 (-0,268)	0,0005 (-0,228)
γ_{t-2}	-0,001 (-1,21)	0,023 (-0,38)	-0,015 (-0,012)	-0,004 (-1,82)	-0,004 (-1,79)
γ_{t-3}	0,0003 (-0,618)	0,05 (-0,821)	-0,014 (-0,012)	-0,0003 (-0,129)	-0,0004 (-0,176)
γ_{t-4}	0,0001 (-0,286)	0,087 (-1,43)	0,006 (-0,011)	0,001 (-0,287)	0,001 (-0,286)
R^2	3,14 %	9,40 %	6,26 %	5,41 %	5,15 %
F -Test	0,604 [0,697]	1,931 [0,097]	1,242 [0,296]	1,063 [0,386]	1,011 [0,416]
DW	2,55	2,62	2,7	2,61	2,61
RESET test:	F-test = 1,057[0,3067]	F-test = 4,863[0,03]*	F-test = 0,077 [0,782]	F-test = 0,005[0,946]	F-test = 0,005 [0,942]
AR 1-5 test:	F-test = 2,878[0,019]*	F-test = 3,549[0,006]**	F-test = 4,116 [0,002]**	F-test = 2,844 [0,02]*	F-test = 2,852 [0,02]*
ARCH 1-4 test:	F-test = 0,987 [0,419]	F-test = 1,361 [0,254]	F-test = 0,522 [0,72]	F-test = 1,698 [0,158]	F-test = 1,613 [0,179]
Normality test:	Chi = 6,406 [0,041]*	Chi = 3,582 [0,167]	Chi = 8,331 [0,012]*	Chi = 7,1186 [0,0285]*	Chi = 6,98 [0,031]*
Hetero test:	F-test= 0,351 [0,964]	F-test = 0,73296 [0,6914]	F-test = 0,176 [0,998]	F-test = 0,652 [0,765]	F-test = 0,653 [0,764]
Hetero-X test:	F-test = 0,924 [0,56]	F-test = 0,783 [0,725]	F-test = 0,603 [0,898]	F-test = 0,525 [0,947]	F-test = 0,544 [0,937]

Note: t-values in parenthesis, single indicator model predictions are estimated with equation (1) and (2), γ is the coefficient.

The inclusion of up to four lags does not provide us with any significant results, except for the Real Exchange Rate's value in the same quarter. There are significant serial- and autocorrelation, proven by the significant AR 1-5 tests, and the Durbin-Watson test scores. The AR 1-5 test sheads doubt as to the reliability of the F test values, and the Durbin-Watson test score means we cannot trust the t-values either. The only model that has a significant coefficient (which we cannot trust) also has a significant RESET test, which indicates that the model is incorrectly specified. The Real Exchange Rate's fit to the basis period is also very low. For the models with a non-significant RESET test value, we have issues concerning non-normality in the error terms. There seems to be no issues of heteroscedasticity in the models' error terms, and the error term is not linked to or dependent on the regressors or their squares. Evidently, the potential for improvement is overwhelming even if some of the OLS criteria are met. We do not recommend using neither the Real Exchange Rate single indicator model for prediction purposes nor any of the other models, as they all appear flawed.

5.1.2 Single Indicator Model Forecasts

Based on the model properties that have been presented in the table 5.1 it is obvious that the predictions with the single indicator models cannot provide anything of worth. The prediction results are presented below, measured by their corresponding (R)MSE and MAPE values. The best prediction model of each indicator is marked with **bold** types and the best single indicator model overall is market with **bold and *italic* types**.

Table 5.2 Single Indicator Model quarterly prediction results, 2006(1)-2010(4)

Single indicators	Lags	MSE	RMSE	MAPE
5 Yr spread	$P_x = 0$	0,002290	0,04785	**2,053**
	$P_x = 1$	0,002308	0,04804	2,109
	$P_x = 2$	**0,002289**	**0,04784**	2,134
	$P_x = 3$	0,002290	0,04785	2,089
	$P_x = 4$	0,002290	0,04785	2,219
Real Exchange Rate	$P_x = 0$	0,002341	0,04838	2,243
	$P_x = 1$	0,002342	0,04839	2,359
	$P_x = 2$	**0,002326**	**0,04823**	2,074
	$P_x = 3$	0,002330	0,04827	2,168
	$P_x = 4$	0,002370	0,04868	**1,999**
OSEAX	$P_x = 0$	0,002272	0,04767	**2,075**
	$P_x = 1$	**0,002272**	**0,04767**	2,177
	$P_x = 2$	0,002355	0,04852	2,684
	$P_x = 3$	0,002336	0,04833	2,786
	$P_x = 4$	0,002343	0,04841	2,608
Real C2	$P_x = 0$	0,002265	0,04759	2,118
	$P_x = 1$	0,002270	0,04765	2,168
	$P_x = 2$	***0,002267***	***0,04761***	1,963
	$P_x = 3$	0,002267	0,04761	***1,946***
	$P_x = 4$	0,002271	0,04766	2,003
Real M1	$P_x = 0$	**0,002270**	**0,04765**	**2,105**
	$P_x = 1$	0,002276	0,04770	2,130
	$P_x = 2$	0,002273	0,04768	2,155
	$P_x = 3$	0,002274	0,04769	2,137
	$P_x = 4$	0,002277	0,04772	2,188

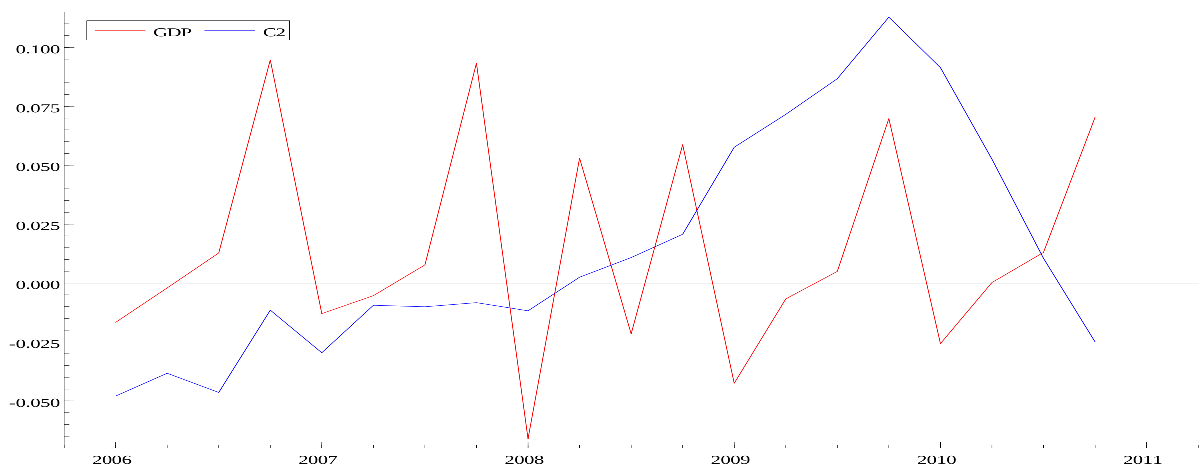
With MAPE as the criterion of selection, three out of five static models provide us with the most precise predictions. The exceptions according to MAPE are Real C2 and Real Exchange Rate, which are most precise with three and four lags respectively. If they are excluded from consideration, this result might be interpreted as confirming some of the findings of Montagnoli and Napolitano (2005) and to some extent Gerdrup et al. (2006). The superior performance of C2 and the Real Exchange rate could however be random.

To a certain extent therefore, the MAPE selection criterion suggests that there is no need for a dynamic model, although the dynamic C2 model with three lags is the most precise overall.

If (R)MSE is chosen as the selection criterion, the C2 model with 2 lags appears to be most precise overall, although the difference between the RMSE values is very small. Contradictory to the conclusions drawn on the basis of the MAPE values, four out of five RMSE values suggests that the dynamic models provide more precise predictions.

Regardless of selection criterion C2 with 2 lags provides the most precise single indicator predictions as we can see in table 5.2. We proceed with the C2 model, despite all the previously mentioned problems, which were quite numerous. Graph 5.1 below shows the extent of those problems.

Graph 5.1 GDP changes (in red) compared to the predictions estimated with the C2 model with 2 lags



Graph 5.1 shows that the correlation between the C2 model and GDP appears to be close to zero.

We believe that because we have based our indicator selection on theory rather than econometric testing, our models might still be used for predictions, despite test results in the basis period indicating the opposite. Whether this conclusion holds will be further investigated in the analysis of our FCIs and comparable models, where we will study how well the single indicator models perform compared to the other possible models we have for prediction.

6 Evaluation of our Financial Condition Index

In this analysis we present the different FCIs that we have constructed. We employ the same procedure as with the single indicators, where we first present the total result of the analysis, and then the prediction results. We also test for a possible long-run static solution, and present only significant results provided by the Wald test. At the end, we present an evaluation of all predictions, along with a recommendation of prediction model.

We could have presented each FCI model with their different lags separately. We do however believe that, since the differences in the coefficients are so small, the exercise of separate presentation would neither aid the main purpose of developing a suitable prediction model nor increase the analysis' explanatory power. Firstly though, the coefficients and model tests:

Table 6.1 the Dynamic Model with Four Lags

Variable name	Weights	Standard Error	t-values
Constant	-0,77	0,17	**-4,5**
GDP_1	-0,74	0,19	**-3,9**
GDP_2	-0,49	0,18	**-2,7**
GDP_3	0,14	0,14	1,0
GDP_4	0,03	0,01	3,0
Oil	0,02	0,03	0,7
Oil_1	-0,06	0,03	**-2,0**
Oil_2	0,05	0,03	1,7
Oil_3	-0,05	0,03	-1,7
Oil_4	0,04	0,03	1,3
Nibor	0,07	0,04	1,8
Nibor_1	-0,03	0,05	-0,6
Nibor_2	0,05	0,04	1,3
Nibor_3	-0,02	0,04	-0,5
Nibor_4	-0,01	0,05	-0,2
Real Exchange Rate	-0,13	0,18	-0,7
Real Exchange Rate _1	-0,01	0,2	-0,1
Real Exchange Rate _2	-0,06	0,18	-0,3
Real Exchange Rate _3	-0,18	0,19	-0,9
Real Exchange Rate _4	0,29	0,19	1,5
Real Estate Index	0,2	0,16	1,3
Real Estate Index _1	0,1	0,16	0,6
Real Estate Index _2	0,06	0,15	0,4
Real Estate Index _3	0,23	0,14	1,6
Real Estate Index _4	-0,06	0,14	-0,4
Oseax	-0,07	0,04	-1,8
Oseax_1	0,05	0,05	1,0
Oseax_2	-0,02	0,05	-0,4
Oseax_3	0,09	0,05	1,8
Oseax_4	-0,04	0,05	-0,8
5 Year Spread	-0,001	0,002	-0,5
5 Year Spread _1	0,001	0,002	0,5
5 Year Spread _2	-0,001	0,002	-0,5
5 Year Spread _3	-0,001	0,002	-0,5
5 Year Spread _4	0	0,001	0,0
C2	0,36	0,21	1,7
C2_1	0,22	0,25	0,9
C2_2	0,48	0,23	**2,1**
C2_3	0,42	0,21	**2,0**
C2_4	0,07	0,19	0,4
M1	-0,36	0,21	-1,7
M1_1	-0,23	0,24	-1,0
M1_2	-0,47	0,23	**-2,0**
M1_3	-0,42	0,21	**-2,0**
M1_4	-0,07	0,18	-0,4
R2			79,20 %
F-test			4,673 [0,000]**
RESET test:	F(1,53) =		0,62650 [0,4322]
AR 1-5 test:	F(5,49) =		1,9278 [0,1066]
ARCH 1-4 test:	F(4,46) =		0,10801 [0,9791]
Normality test:	Chi^2(2) =		5,0031 [0,0820]
Hetero test:	Chi^2(88)=		83,899 [0,6039]

Note: Models are estimated using equation (3) and (4), significant t values are written in bold types

By a first glance, the FCI model appears to have everything in order. The R-squared is relatively high, the F-test is significant and unlike the previous single indicator models there seems to be no signs of serial correlation. In general we might now be able to trust the t-values, but because many of them are very low and R-squared is high, we might have multicollinearity issues. There seems to be no issues concerning heteroscedasticity, lack of normality in the error terms or model specification. We also wished to include the Hetero-X test in this analysis, but because there were not a sufficient number of observations for the test to be calculated, it had to be omitted.

We are aware of the many issues in the models, but choose to predict with them nonetheless. We are after all constructing FCIs based on literature rather than econometric testing, and we are not trying to explain the basis period, but the period 2006-2010.

6.1 Static Long-run Solutions

For all the models we conducted a Wald test to find any long-run static solutions. The first significant Wald test appeared in the model with three lags, and the second in the model with four lags. We present both in order to highlight the point of small differences made above. In both tables we present the characteristics of the significant static long-run solutions, for the FCIs with three and four lags:

Table 6.2 Static Long-Run Solution Models

Solved static long-run equation for GDP				
	FCI with three lags		FCI with four lags	
Variable name	Weight	Standard Error	Weight	Standard Error
Constant	0,011 (8,58)	0,001	0,01 (5,72)	0,002
Oil	-0,018 (-1,25)	0,014	0,002 (0,1)	0,021
Nibor	0,024 (1,67)	0,014	0,02 (1,14)	0,02
Real Exchange Rate	-0,091 (-0,9)	0,102	-0,03 (-0,22)	0,147
Real Estate Index	0,2 (5,35)	0,037	0,19 (4,11)	0,047
Oseax	0,017 (0,98)	0,017	0,01 (0,26)	0,024
5 Year Spread	-0,001 (-0,65)	0,001	-0,0002 (-0,14)	0,001
C2	0,421 (4,63)	0,091	0,54 (4,02)	0,135
M1	-0,422 (-4,72)	0,089	-0,54 (-4,09)	0,133
WALD test: Chi^2(8):	66,1388 [0,0000] **		45.1941 [0.0000] **	

Note: deduced from equation (4), t values in parenthesis

As we can see in the table, there is not much to gain in terms of t-values, as only about fifty percent of the variables are statistically significant. The Wald test statistic indicates that the models are statistically significant, and we make predictions with these models in the same manner as with the FCIs.

Before we proceed to the predictions, one last table is needed to clarify potential problems and inaccuracies in the FCI models. This next table is an analysis of the significance of the lag structure:

Table 6.3 Significance of each lag

Test values			
Lag 4	F(9,54) =	1,004	[0.4483]
Lag 3	F(9,54) =	2,9984	[0.0057]**
Lag 2	F(9,54) =	2,5814	[0.0150]*
Lag 1	F(9,54) =	4,4522	[0.0002]**

This table shows that the fourth lag seems to be non-significant, a conclusion also induced in table 6.1, where none of the variables' fourth lags are significant. This conclusion is also confirmed by the fact that the inclusion of a fourth lag only provides approximately 3 % increased fit to the basis period.

7 Forecast Tests

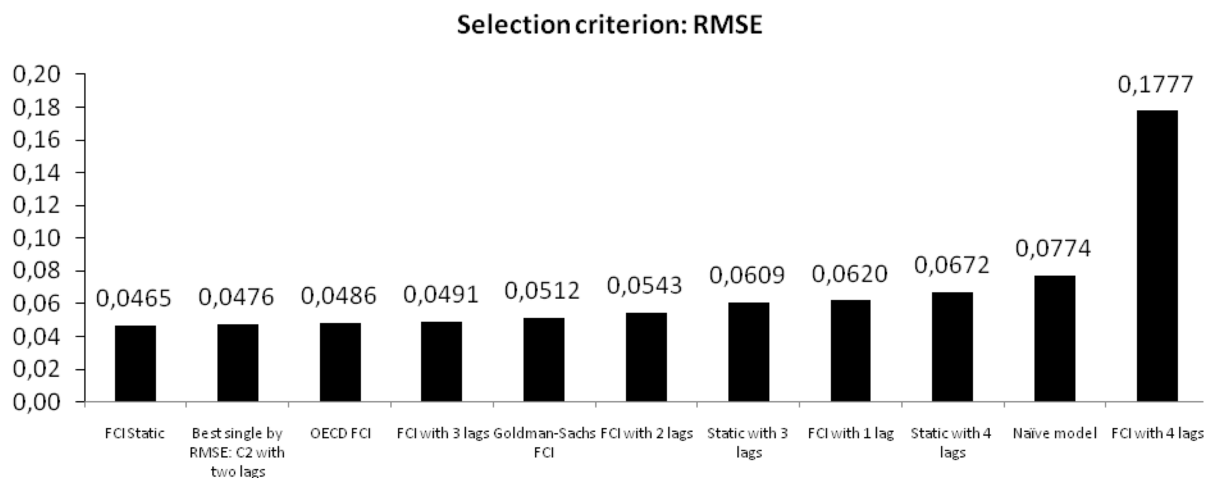
In this chapter we present the results of all the out-of-sample predictions. The Naïve model and the predictions made with OECDs and GS' FCIs are all good ways of comparing model performance. The Naïve model is a simple non-technical way of predicting, the FCIs are the opposite. A comparison against these models should therefore help determine whether our FCIs have any value or not. Table 7.1 presents the results of all the predictions:

Table 7.1 Prediction results

Model	MSE	RMSE	MAPE
FCI Static	0,0022	0,0465	8,5
FCI with 1 lag	0,0038	0,0620	3,5
FCI with 2 lags	0,0030	0,0543	23,4
FCI with 3 lags	0,0024	0,0491	23,2
FCI with 4 lags	0,0316	0,1777	60,2
Naïve model	0,0060	0,0774	8,1
OECD FCI	0,0024	0,0486	4,0
Goldman-Sachs FCI	0,0026	0,0512	6,2
Static with 3 lags	0,0037	0,0609	8,6
Static with 4 lags	0,0045	0,0672	9,3
Best single by RMSE (MAPE): C2 with two (three) lags	0,0023	0,0476	1,9

In table 7.1 one major divide is of great importance. The selection criterion will determine which of the models provide the most precise predictions. As RMSE might lead to a different decision than MAPE, we present both approaches.

Figure 7.1 Selection based on RMSE

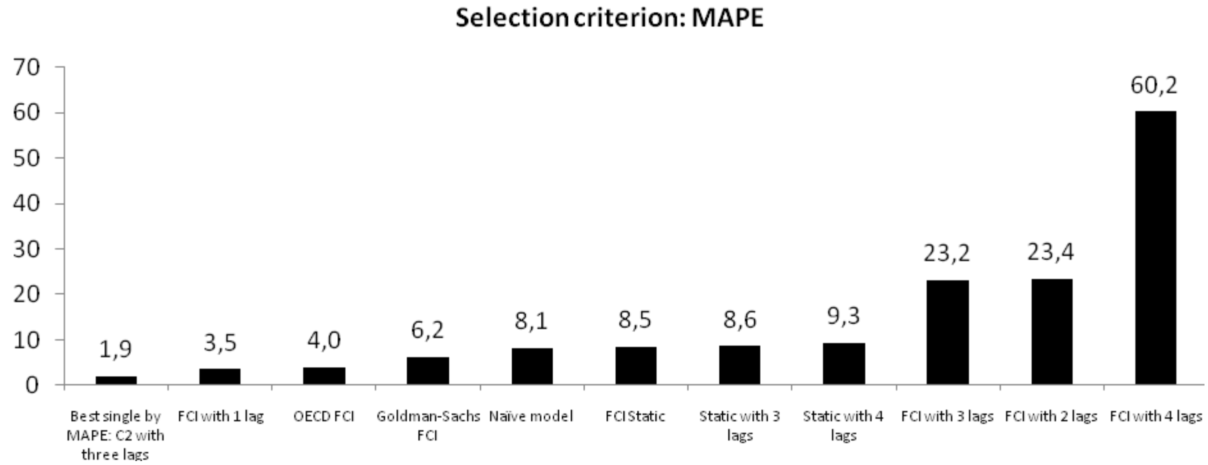


If the criterion selected is RMSE, the static FCI provides the most precise predictions for the period 2006-2010. But as we mentioned earlier, the model is completely unreliable. The static FCI is closely followed by the C2 with two lags, which was non-significant. As we can see in table 6.1 it seems like it is the lags of GDP, C2 and M1 that drives the main prediction results in the dynamic FCI models. It is therefore interesting that the FCI model with three lags does worse than the C2 model, even though it should contain more information and hence provide better predictions. We are therefore inclined to assume that the extension of the model provides more noise than information.

In the other end of the scale, we find the Naïve model, and our own FCI with four lags. This is not surprising as the weights of the FCI model with four lags were mostly non-significant, and the fourth lag was proven to be non-significant overall. It is even more imprecise than the static solution which originates from it. As we can see, the OECD FCI is the most precise benchmark, and the GS FCI which is second in line, is beaten by our own FCI with three lags. Our best model, the static FCI, predicts more precise than any of the (externally developed) benchmarks, which we find encouraging.

The next figure presents the alternative results of the analysis:

Figure 7.2 Selection based on MAPE

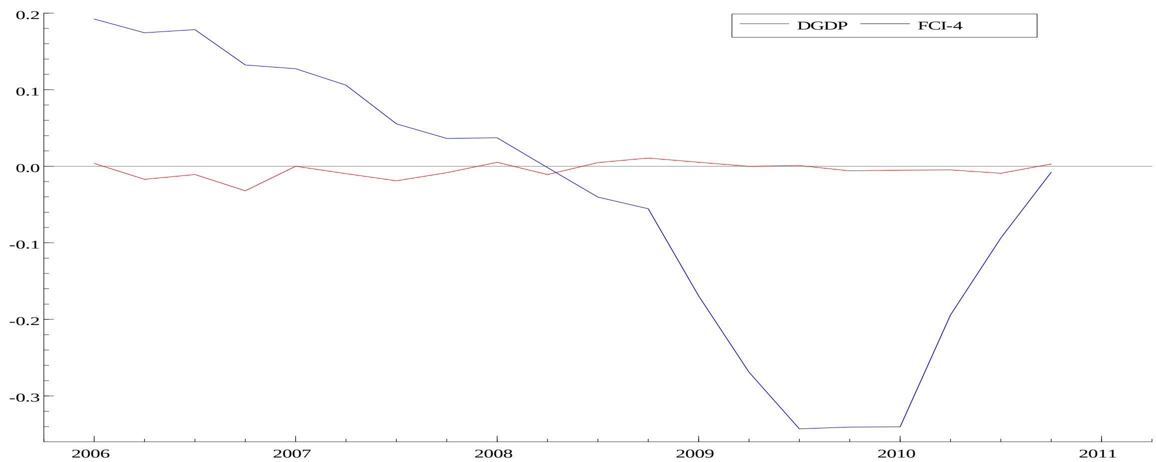


When the criterion is changed the story is as well. The single indicator C2 model has now become the most precise. Our own FCI model with one lag is the most precise of the FCIs judging by MAPE, and we take pleasure in beating OECDs and Goldman Sachs' FCI predictions again, although with a different model. As with the previous measure, the C2 model which should contain less information than the FCIs, is better, suggesting that the total result of including more variables and lags does not contribute positively to prediction preciseness.

Ideally, the two different measures of preciseness should indicate that the same model is the most precise. Even if the static FCI is the best of the FCIs according to RMSE, it is still not the best overall and still contains no statistically significant weight. Taking the diverging indications by RMSE and MAPE into consideration, we believe that all the FCIs we have constructed are still pretty imprecise.

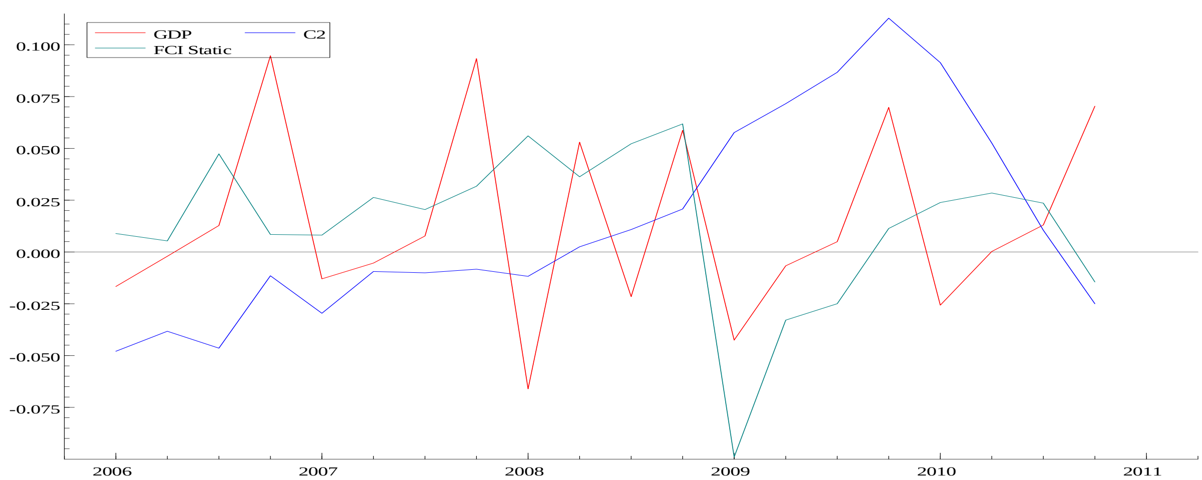
Figure 7.2 also shows that when MAPE is the selection criterion, both of the static long-run FCI models are more precise than the models from which they originate. The results of the lag significance test proved that the fourth lag of the FCI models was non-significant, and as we can see, this model performs significantly worse than the rest, independent of preciseness measure. With an unsurprisingly high MAPE value around 60, this is by far the worst performing model. This suggests that none of our included variables effects GDP with a 9 to 12 month delay. The reason why the two measures provides such conflicting results may be the "outliers" that are present in the datasets. To show this clearly, we present the graph below, which compares our least precise FCI's predictions to actual GDP changes (GDP is as usual in red).

Graph 7.1 GDP Changes and the FCI with Four Lags



As we can see in the graph, there are large differences between the actual and predicted values of GDP, and these gaps or outliers are probably what cause the reported impreciseness of the model with four lags. Before we proceed to the conclusions and recommendations, a graphical representation of the best models is needed:

Graph 7.2 Actual GDP, the static FCI and the C2 model with two lags



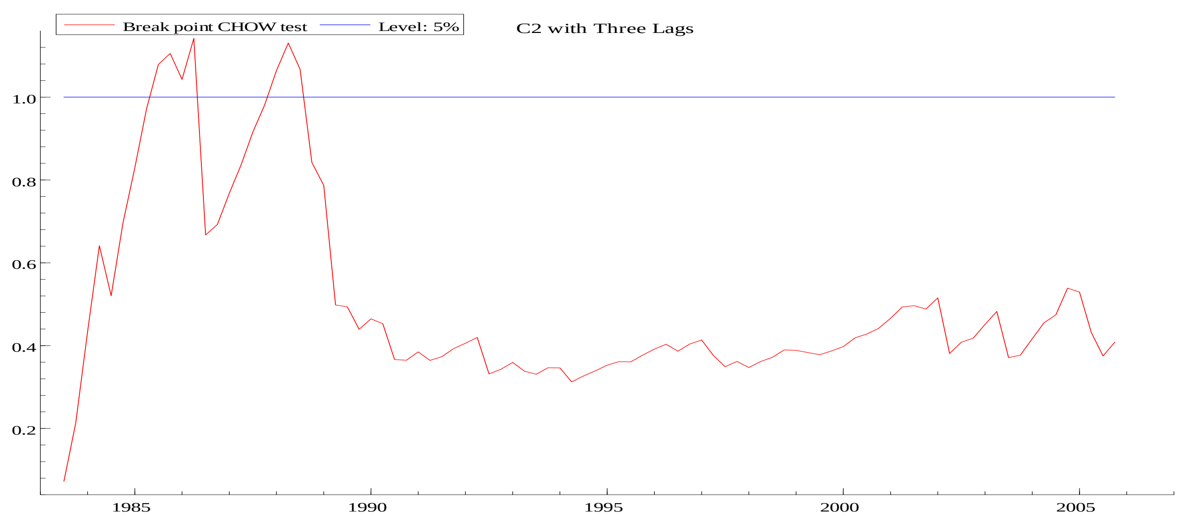
The graph might be interpreted as showing that the static FCI is the most precise, followed by the C2 model with three lags.

7.2 Basis Period Characteristics

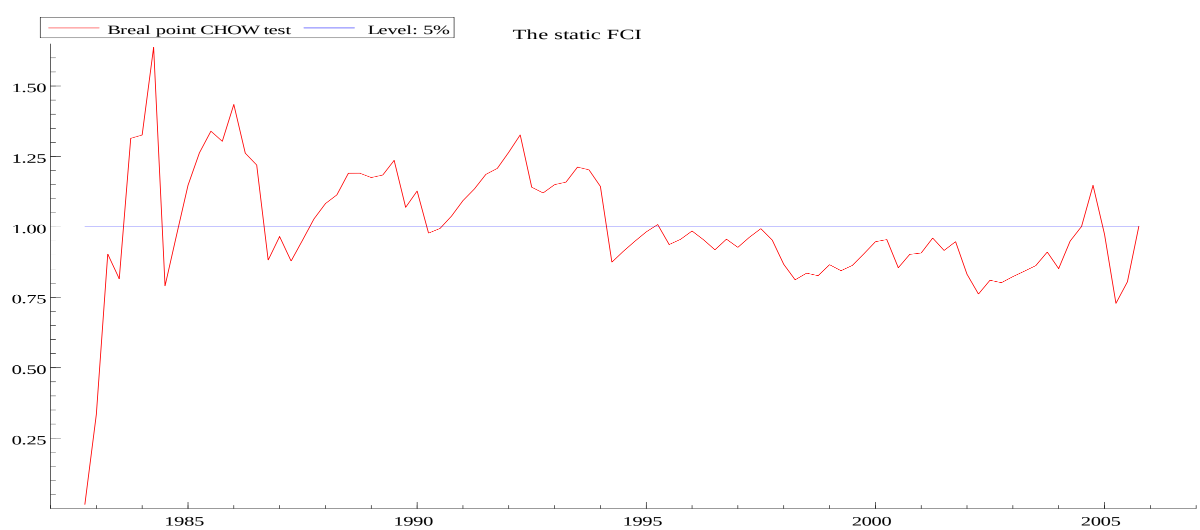
Through the analysis, in which we examined basis period characteristics and the models that result from it, it has become clear that all the models we have constructed still have improvement potential.

According to Perron (1989) time series data consists of trends and structural breaks. We therefore conducted a Chow test for structural breaks, on C2 with three lags and the variables included in the static FCI, to investigate whether this was also the case in our dataset. The results shown in graph 7.3 and 7.4 shows the extent of the instability issues:

Graph 7.3: Break point Chow test for C2 with Two Lags



Graph 7.4: Break point Chow test for the static FCI



The first graph show that the basis period of the best single indicator is unstable and the coefficients are therefore likely to be unstable, and hence not trustworthy. This assumption corresponds to conclusions drawn from table 6.1. The lower graph show why the fit to the basis period for the FCIs are low in the models; there are more structural breaks than there are stable quarters on the 5 % level. This is the exact opposite of what we would like the test result to look like, as breakpoints are unhelpful when we try to maximize predictive preciseness. They supply interferences in the assumed linear relationship between the variables, and because of the linearity assumptions in OLS estimation, this can induce us to accept spurious relationships when we estimate the weights of the prediction models based on this basis period.

7.3 Conclusions Based on the Single Indicator and FCI analysis

The analysis has proven that most of the variables included in our data set do not appear to have leading characteristics that are able to explain the fluctuations in Norwegian GDP. Those few weights which seemed to have leading characteristics and were significant were in bold types in table 6.1.

The C2 indicator models provided the best fit compared on the MAPE values, but did slightly worse when compared by RMSE. The conclusion we draw from this, combined with table 5.2, is that judging by MAPE, there seems to be an increasing degree of relevant information in the lags up to three quarters. However, the non-significance issues of the C2 single indicator models, and the results in table 6.1 are clear indications of a general lack of leading characteristics for a majority of the variables, both the singles and the ones included in our FCIs. In other words, *"much that once was, is lost, for none now live who remember it"* (Tolkien 1954, p.1).

For the FCIs, many of the same conclusions apply. If the criterion is RMSE, the static FCI that we constructed is the most precise. However, if the criterion is MAPE, the C2 single indicator model is the most precise. If one was to choose one of our FCIs based on MAPE, the most precise is the FCI with one lag. As mentioned above the two measures should ideally indicate that the same model is superior in terms of prediction preciseness. Because the indications of the two measures diverge and because none of the FCI weights can provide significant values, we believe that none of the models that we have constructed should actually be used for predictions. If the models are used anyway, the intended use should determine which model is selected, and hence which selection criterion is the most heavily weighted.

With regards to the static models, there is a slight problem though. If one intends to predict Norwegian GDP, using numbers from the same quarter can be regarded as a challenge. When one wishes to predict, having the numbers from the quarter one is trying to predict is difficult at best. However, we get around this logical inconsistency by the fact that GDP numbers are published with a considerable delay. This means that it is possible to use a static model to predict, because most of the indicators are published with a delay which is shorter than that of the GDP. The shorter publishing intervals are also one of the criteria for inclusion to begin with (Gerdrup et al. 2006). We initially set out to construct these FCIs based on the assumption that they might be able to aid in the prediction models already in use in asset management firms, and this is why we do not see the potential timeline issue as a real problem.

If our resistance to recommend was stomped and we were forced to recommend one of our FCIs, say at gunpoint or something of similar severity, we would recommend our static FCI which we have named the Torvanger-Ørbeck FCI. The reason for this choice is that the logic we base the choice on, is very similar to that of several other previous research papers. In addition, the EMH would indicate that this model should contain all relevant information. Even if we know that the static FCI is imprecise, in a majority of all of the tests conducted above the static model's superior performance could be interpreted as lending some support at least, to this conclusion[12]. We have also chosen to emphasize the RMSE measure the most, because we use OLS regressions. We have therefore chosen the RMSE measure as the determinant in our final model recommendation.

One probable explanation for the static model's superior performance could be the ever-increasing flow of information that modern computer technology is able to provide, with which information spreads at lightning speeds. One quarter is after all a considerable amount of time, and should be more than enough for all available public information to be included in the financial variables, and thereby aligning with the assumptions of EMH.

7.4 A Comparison of Previous Findings and Our Results

In section 3 we went through several assumptions made in previous research papers, their findings, and our own assumptions about the relationships between the different variables included in the analysis. In this section we compare previous results and assumptions with the findings of the analysis above.

[12] We would also like to stress that the inclusion of several other domestic and international variables, such as those included in Hatzius et al.'s article, or inclusion of several more lags, does not improve the models predictive power.

Short-term interest rates such as the Nibor 3m were assumed to have leading properties by Montagnoli and Napolitano (2005), and by Goodhart and Hofmann (2001). In our case, this turned out to be false. Goodhart and Hofmann (2001) also suggest that the Real Exchange Rate was significant in fifty percent of the cases they surveyed, but as a single indicator we found it to be significant in the same quarter only, throughout the period. In the FCIs the Real Exchange Rate did not produce significant results, contradicting the previous paper altogether. The authors also suggested that a Real Estate Index (house prices) could serve as a leading indicator, but the Real Estate Index published by Statistics Norway did not achieve any statistically significant leading properties in our analysis. This is shown in table 6.1.

The paper by Husebø and Wilhelmsen (2005), which analyzed the period 1982 – 2003 found evidence of a relationship between economic indicators and GDP in the U.S. and the Euro area, both in terms of strength and whether they seem to lead, lag or coincide with GDP. As we can see from the Chow test in Graph 7.4, they might have been able to avoid some of the structural instability that we included in our data set by starting their basis period in 1982. That might help explaining why they got significant results, and we did not.

Gerdrup et al. (2006) argued that the correlations between the economy and financial sizes change over time. We found this to be true using the Chow-test. The authors also found that the coefficients between GDP and lagged values of real credit and real money supply were either low or negative throughout the period. The findings in table 6.1 contradict these results. Both the C2 and the M1 have relatively high coefficients, and C2 is positive while M1 is negative. The negative M1 weights probably mean that the consumption effect mentioned in section 3.3.1.4 is the driving factor of the correlation. Both indicators seem to have significant leading properties with two and three lags, although Gerdrup et al. (2006) found leading characteristics only for the M1, and only in the same and previous quarter. They also found a correlation between real share prices and GDP, but we were not able to confirm these findings, neither in the single indicator analysis nor in the FCIs. With regards to Nibor, where Gerdrup et al. (2006) found negative correlation, we encountered both positive and negative correlation in our analysis, although non-significant. The authors also found the 5 Year Spread to be statistically significant and leading, but our analysis shows no correlation or significance for this spread. They also found the Real Exchange Rate to be positively correlated to GDP, while we proved there to be a negative relationship between the two if the Real Exchange Rate was a single indicator, but no relationship when it was included as part of an FCI.

The last included financial indicator is the Oil Price, which we thought would have a positive correlation with the fluctuations in GDP. In table 6.1 the non-significant correlation is both positive and negative, dependent on the number of lags, suggesting a fairly random relationship between the two.

8 Learning Outcomes

When we started writing this thesis, we did not have any previous experience on FCI construction or identification of indicators with leading properties. At the outset several problems were encountered. Because we based our indicator selection on American literature mainly, many of the indicators found to have leading characteristics in the U.S. did not have a Norwegian counterpart. This meant that several significant U.S. relationships, that could have been significant in Norway as well, had to be omitted. Obtaining data series spanning back to 1980 turned out to be problematic for some of the indicators we initially wanted to include, and they therefore had to be left out. Even for those variables we did include, though historic observations are published regularly, the intervals of the observations differed. We also discovered that different institutions published conflicting observations for the same indicators. Where we had a choice of data source we strived to obtain data from the official government web pages, Norges Bank or Statistics Norway. These are all publicly available data free of charge. We emphasized the data's availability without payment in order to encourage possible employment and later improvement and updates of the FCIs, as well as the general principle of reconstructability in academic research.

After these issues had been treated, new ones emerged. Obtaining benchmark FCIs turned out to be a challenge, because these prediction services were subject to a fee. We were lucky enough to receive predictions for the Norwegian GDP from Goldman Sachs and OECD through our personal contacts, without payment.

8.1 Further Research Opportunities

In hindsight, there are several ideas that have potential for further research on this topic.

Firstly, one could use GDP estimates that include the oil and gas producing sector. Since this sector is largely driven by international influences, prediction results could have been improved by including variables that capture co-movements between the Norwegian and foreign economies.

In addition, shorter intervals between observations might improve predictive power. Some of the impreciseness could have been avoided if one was to use the real level value of some of the indicators, instead of their changes. The actual level of the Oil price, Nibor and the Exchange Rate is likely to have more of a significant and even leading influence on GDP, compared to their changes in each quarter. Unfortunately we were not aware of this at the outset. If we had been, issues related to treating indicators measured in different units would also have to be investigated and dealt with, in addition to potential problems with non-stationary data series in regressions.

Another improvement possibility is to give more weight to observations closer in time, when estimating the models' weights. In the period that we have chosen to base our models on, there have been judiciary changes, as well as several crises or structural breaks, evident in the Chow tests. This weakens the significance of the estimated weights, and contributes to the instability in the parameters. Some of these issues could have been remedied by a shorter or different basis period. Also, our measures of predictive preciseness both suffer from the same weakness with regards to outliers, so additional preciseness measures might help determine which and what kind of model contains the most predictive power. Combined with the measures already in place and a Winsorizing[13] of the variables, one might be able to overcome some of these issues.

There are probably also some gains to be had by utilizing principal components. Surveys on adequate measures for the components of the Keynesian equations for example, might be one way of determining which indicators to combine into such principal components. Conducting an analysis of assumed correlations would probably aid in indicator selection.

There are also the issues of unbalanced panel estimation. When the different time series start at different points in time, issues arise, which are solved using unbalanced panel estimation methods. Hatzius et al. utilizes such methods to account for the different starting points of their indicator observation series. We did not employ neither principal components nor unbalanced panel estimation techniques because that would contradict our principle of simplicity and hence, potential real-life predictions with the Torvanger-Ørbeck FCI.

[13] a.k.a Boot-strapping or setting extreme values equal to a fixed level

References

Armstrong, J.S., and Collopy, F. (1992), *Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons,* International Journal of Forecasting No 8, pp. 69-80.

Banerjee, A. & Marcellino, M. (2003), *Are There Any Reliable Leading Indicators for U.S. Inflation and GDP Growth?*, Working Papers 236, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University, Italy.

Banerjee, A., Marcellino, M., & Masten, I. (2003), *Leading Indicators for Euro-area Inflation and GDP Growth*, Working Papers 235, IGIER, Bocconi University, Italy.

Beaton,K., Lalonde,R. and Luu, C. (2009), *A Financial Conditions Index for the United States,* Bank of Canada Discussion Paper 11

Bjørnland, H.C. (2000) "*VAR Models in Macroeconomic Research*", Statistics Norway, Research Department.

Borio, C. and Lowe, P. (2002), *Asset prices, financial and monetary stability: exploring the nexus,* Monetary and Economic Department Working Papers No 114, Bank of international Settlement.

Boskin, M.J., Dulberger,E.R., Gordon, R. J., Griliches, Z. and Jorgenson, D.W. (1998), *Consumer Prices, the Consumer Price Index, and the Cost of Living,* The Journal of Economic Perspectives, Vol. 12, No. 1 pp. 3-26

Campbell, J.Y. and Shiller, R. J. (1990), *Yield Spreads and Interest Rate Movements: A Bird's Eye View,* The Review of Economic Studies, Vol. 58, No. 3, Special Issue: The Econometrics of Financial Markets, pp. 495-514

Eika, K.H., Nymoen.R. Ericsson and R.Nymoen (1996), "Hazards in implementing a monetary conditions index" Board of Governors of the Federal Reserve System, 568, pp 1-34.

Estrella, A. and Mishkin, F.S. (1998), *Predicting U.S recessions: Financial variables as leading indicators,* The Review of Economics and Statistics, Vol. 80, No. 1, pp.45-61.

Estrella, A. (2005), *Why Does the Yield Curve Predict Output and Inflation?*, The Economic Journal Volume 115, Issue 505, pp. 722–744

Gerdrup, K.R., Hammersland, R., and Naug, B.E. (2006), *Finansielle størrelser og utviklingen i realøkonomien,* Penger og kreditt 2/2006 Norges Bank.

Goodhart, C. and Hofmann, B. (2001), *Asset Prices, Financial Conditions, and the Transmission of Monetary Policy*, Paper prepared for the conference on 'Asset Prices, Exchange Rates, and monetary Policy', Stanford University.

Gropp, R., Vesala, J. and Vulpes, G. (2002), *Equity and bond market signals as leading indicators of bank* fragility, Working paper No 150, European Central Bank.

Guichard, S. and Turner, D. (2008), *Quantifying the Effect of Financial Conditions on US activity*, OECD Economics Department Working Papers, No.635, OECD Publishing.

Guichard, S. Haugh, D. and Turner, D. (2009), *Quantifying the Effect of Financial Conditions in the Euro Area, Japan, United Kingdom and United States*, OECD Economics Department Working Papers, No.677, OECD Publishing

Gujarati, D.N. & Porter, D.C. (2009). *Basic Econometrics,* 5'th edition, New York, Mc Graw Hill.

Hakkio, C.S. and W.R. Keeton (2009), *Financial stress: What is it, How Can It Be Measured, and Why Does It Matter*?, Federal Reserve Bank of Kansas City Economic Review.

Hatzius, J., Hooper, P., Mishkin, F., Schoenholtz, K.L. and Watson, M.W. (2010), Financial Conditions Indexes: A Fresh Look after the Financial Crisis, NBER Working Paper No. 16150, National Bureau of Economic Research.

Hooper, P., Mayer,T., Slok, T. and Biggs, M. (2010). *Global Economic Perspectives - Financial Conditions Weakest Since Late '08*, Global Markets Research 2010, Deutsche Bank.

Husebø, T.A and Wilhelmsen, B.R. (2005), *Norwegian Business Cycles 1982-2003*, Staff Memo No. 2/2005 Economics Department, the Norwegian Bank.

Jolliffe, I. T. (2005) “Principal Component Analysis” 2nd Edition, Encyclopedia of Statistics in Behavioral Science, United Kingdom, Springer Publishing.

Laurent, R. (1989), *Testing the Spread*, Federal Reserve Bank of Chicago, *Economics Perspectives*, 12:1, pp.3-14.

Mabro, R. (1992), *OPEC and the Price of Oil*, Oxford Institute for Energy Studies, Oxford University Press, pp. 1-17

Montagnoli, A. and Napolitano, O. (2005), *Financial Condition Index And Interest Rate Settings: A Comparative Analysis*, Working Papers No 8, D.E.S. (Department of Economic Studies), University of Naples "Parthenope", Italy

Norges Bank (2004), Finansinstitusjonene, chapter. 5 i Norske finansmarkeder – *pengepolitikk og finansiell stabilitet*, Norges Banks Skriftserie No 34, pp. 66-74

Perron,P. (1989), *Testing For A Unit Root In A Time Series With A Changing Mean*, Papers 347, Princeton, Department of Economics - Econometric Research Program.

Riiser, M.D. (2005), *House prices, equity prices, investment and credit– what do they tell us about banking crises? A historical analysis based on Norwegian data*, Economic Bulletin 3/2005, Norges Bank, pp.145-154.

Riiser, Magdalena D. (2008), *Asset prices, investment and credit – what do they tell us about financial vulnerability?*, *Economic Commentary 6|2008*, Norges Bank.

Riiser, M.D. (2010), *Prices, investment, credit and financial vulnerability*, Economic commentaries no. 4, Norges Bank.

Rosenberg, M.R. (2009), *Global financial market trends & policy*, Bloomberg Financial conditions watch Volume 2 No.6, Bloomberg.

Rousseas, Stephen. 1985, *A markup theory of bank loan rates*, Journal of Post Keynesian Economics, Vol. 8, Issue 1, Pages 135-144

Svensson, Lars E.O. 2000, *Open-Economy Inflation Targeting*, Institute for International Economic Studies, Stockholm University, Journal of International Economics, Volume 50, Issue 1, Pages 155-183

Tolkien, J.R.R. (1954), *The Lord of the Rings: The Fellowship of the Ring, 1st edition,* George Allen & Unwin

Ziegler, C., and Eickmeier, S. (2006), *How good are dynamic factor models at forecasting output and inflation? A meta-analytic approach, Discussion Paper Series 1: Economic Studies 2006,42, Deutsche Bundesbank, Research Centre*

[1] http://www.ssb.no/emner/09/01/begreper/begreper.html 25.04.11 09:10
[2] http://www.ssb.no/vis/knr/om.html 26.04.11 11:26
[3] http://www.ssb.no/vis/knr/om.html 26.04.11 11:26
[4] http://www.ssb.no/emner/09/01/begreper/begreper.html 25.04.11 09:55
[5] http://www.norges-bank.no/Upload/77504/Produ (10.04.11) 18:03
[6] http://www.ssb.no/english/subjects/08/02/10/kpi_en/tab-01-en.html 30.04.11 13.22
[7] http://www.bloomberg.com/apps/quote?ticker=BFCIUS:IND 21.04.2011 09.49
[8] Deutsche Bank Global Markets Research 2010
http://online.wsj.com/public/resources/documents/deutschereport0623.pdf 18.03.2011 14.05
[9] http://www.kansascityfed.org/publications/research/er/er-2009.cfm 06.01.11 14:10
[10] http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=eco/wkp(2008)43&doclanguage=en 11.02.11 10.22
[11] http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=eco/wkp(2009)18&doclanguage=en 15.02.11 14.10
[12] http://www.bankofcanada.ca/en/res/dp/dp(y)_2009.html 21.04.2011 11:39
[13]http://www.nber.org/s/search?q=Financial+Conditions&btnG=Search&entqr=0&output=xml_no_dtd&sort=date%3AD%3AL%3Ad1&client=test3_fe&ud=1&oe=UTF8&ie=UTF8&proxystylesheet=test3_fe&site=default_collection&whichsearch=ftall&restrict_papers=yes&restrict_books=yes&fullresults=1 21.04.2011 12:02
14 http://www.norges-bank.no/no/ord-og-uttrykk/#N (20.02.2011) 15:06
15 http://www.regjeringen.no/nb/dep/fin/dok/nouer/2011/nou-2011-1/18.html?id=631344 (07.02.2011) 12:32
16 http://oslobors.no/markedsaktivitet/stockIndexOverview?newt__ticker=OSEAX (21.01.2011) 14:54
17 http://www.ssb.no/bpi/om.html (21.01.2011) 15:33
18 http://www.norges-bank.no/templates/article____67671.aspx (20.02.2011) 15:35
19 http://www.ssb.no/k2/ (01.02.2011) 11:44
[20] http://www.ssb.no/vis/emner/11/01/k3/om.html 21.04.2011 16:58
21 http://www.burneyco.com/PDF/W5_MM_CS_VIX_Note_v5.pdf (21.01.2011) 12:34
[22] http://www.oecd.org/faq/0,3433,en_2649_33733_1798284_1_1_1_1,00.html 03.05.2011 12:59
[23] http://www.oecd.org/document/46/0,3746,en_2649_34109_33701806_1_1_1_1,00.html#t_1 03.05.11 13.00