A COMPOSITE LEADING INDICATOR OF THE FINNISH ECONOMY

Henry Jacob Beniard

University of Helsinki

Faculty of Social Sciences

Department of Economics

Master's Thesis

August 2010

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto – Fakultet/Sektion – Faculty	Laitos – Institution – Department
Faculty of Social Sciences	Department of Political and Economic Studies

Tekijä – Författare – Author
Beniard, Henry Jacob

Työn nimi – Arbetets titel – Title
A Composite Leading Indicator of the Finnish Economy

Oppiaine – Läroämne – Subject
Economics, econometrics

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	September 2010	60

Tiivistelmä – Referat – Abstract

This thesis researches empirically, whether variables that are able to reliably predict Finnish economic activity can be found. The aim of this thesis is to find and combine several variables with predictive ability into a composite leading indicator of the Finnish economy. The target variable it attempts to predict, and thus the measure of the business cycle used, is Finnish industrial production growth.

Different economic theories suggest several potential predictor variables in categories, such as consumption data, data on orders in industry, survey data, interest rates and stock price indices. Reviewing a large amount of empirical literature on economic forecasting, it is found that particularly interest rate spreads, such as the term spread on government bonds, have been useful predictors of future economic growth. However, the literature surveyed suggests that the variables found to be good predictors seem to differ depending on the economy being forecast, the model used and the forecast horizon.

Based on the literature reviewed, a pool of over a hundred candidate variables is gathered. A procedure, involving both in-sample and pseudo out-of-sample forecast methods, is then developed to find the variables with the best predictive ability from this set. This procedure yields a composite leading indicator of the Finnish economy comprising of seven component series. These series are very much in line with the types of variables found useful in previous empirical research.

When using the developed composite leading indicator to forecast in a sample from 2007 to 2009, a time span including the latest recession, its forecasting ability is far poorer. The same occurs when forecasting a real-time data set. It would seem, however, that individual very large forecast errors are the main reason for the poor performance of the composite leading indicator in these forecast exercises. The findings in this thesis suggest several developments to the methods adopted in order to produce more accurate forecasts. Other intriguing topics for further research are also explored.

Avainsanat – Nyckelord – Keywords
econometrics, business cycle, leading indicator, forecasting, economic growth

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto – Fakultet/Sektion – Faculty	Laitos – Institution – Department	
Valtiotieteellinen tiedekunta	Politiikan ja talouden tutkimuksen laitos	
Tekijä – Författare – Author		
Beniard, Henry Jacob		
Työn nimi – Arbetets titel – Title		
A Composite Leading Indicator of the Finnish Economy		
Oppiaine – Läroämne – Subject		
Kansantaloustiede, ekonometrian linja		
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Pro gradu	syyskuu 2010	60

Tiivistelmä – Referat – Abstract

Tässä työssä tutkin empiirisesti, onko mahdollista löytää Suomen taloutta luotettavasti ennustavia muuttujia. Tutkimukseni päämääränä on pyrkiä löytämään tällaisia muuttujia ja muodostaa näistä Suomen talouskasvua ennakoiva suhdanneindikaattori. Muuttuja, jota suhdanneindikaattori pyrkii ennustamaan ja jolla suhdannetta kuvataan, on Suomen teollisuustuotannon kasvu.

Talousteoriasta on löydettävissä lukuisia mahdollisesti talouskehitystä ennakoivia muuttujia esimerkiksi kulutusdatan, teollisuuden tilausdatan, kyselytutkimusdatan, korkojen ja osakeindeksien joukosta. Suuri määrä aiempaa talousennustamista koskevaa empiiristä kirjallisuutta, jota tässä tutkimuksessa käsittelen, osoittaa, että korkojen tuottoerot, esimerkiksi valtion obligaatioiden kohdalla, ovat olleet erityisen hyviä talouskasvun ennakoijia. Kirjallisuuskatsauksen perusteella vaikuttaa kuitenkin siltä, että tutkimuksissa hyviksi havaitut muuttujat vaihtelevat riippuen taloudesta jota ennustetaan, käytetystä mallista ja ennustehorisontista.

Kirjallisuuskatsauksen perusteella valitsen yli sata muuttujaehdokasta. Työssäni kehitetään menetelmä, jonka tarkoituksena on löytää parhaat teollisuustuotannon kasvun ennustajat tästä muuttujajoukosta. Menetelmän seurauksena saadaan Suomen talouden ennakoiva suhdanneindikaattori, joka koostuu seitsemästä eri muuttujasta. Nämä komponenttimuuttujat ovat hyvin samanlaisia aiemmassa empiirisessä kirjallisuudessa hyväksi havaittujen muuttujien kanssa.

Käytettäessä muodostettua ennakoivaa suhdanneindikaattoria ennustamiseen otoksessa 2007–2009, joka pitää sisällään myös uusimman taantuman, ennustetarkkuus heikkenee huomattavasti. Sama havaitaan ennustettaessa teollisuustuotannon ensijulkistuksista muodostettua reaaliaikasarjaa. Vaikuttaa kuitenkin siltä, että yksittäiset hyvin suuret ennustevirheet ovat pääsyy ennakoivan suhdanneindikaattorin heikkoon ennustetehoon näissä tapauksissa. Tämän tutkielman havainnoista saadaan useita kehitysehdotuksia käytettyihin menetelmiin ennustetehon parantamiseksi. Myös muita kiinnostavia lisätutkimuskysymyksiä käsitellään.

Avainsanat – Nyckelord – Keywords
ekonometria, suhdannevaihtelut, talouskasvu, taloudelliset ennusteet

CONTENTS

1 Introduction 2
2 Measuring economic activity and the forecasting model 3
2.1 The choice of the target variable 3
2.2 The forecast model 5
3 The leading indicators 7
3.1 Examples of composite leading indicators 7
3.2 Non-financial variables 10
3.3 Financial variables 12
3.4 Discussion 19
4 The data 21
4.1 The candidate variables 21
4.2 Data transformations 24
5 The variable selection process 26
5.1 Step 1: Forming the base set of variables 27
5.2 Step 2: The three-stage variable selection procedure 30
6 Results 33
6.1 The composite leading indicator 33
6.2 Evaluating the performance of the composite leading indicator 38
6.2.1 Forecasting in an unused sample and the latest recession 39
6.2.2 Forecasts using a real-time data set 44
6.2.3 Forecasting a composite coincident indicator 48
6.2.4 Stability of the CLI model 49
7 Conclusions 49
REFERENCES 53
Data Appendix i
Table A1: The time series used and the data source. i
Table A2: The questions of the business and consumer surveys. v

1 Introduction

Forecasting future economic activity has again become topical due the recent global recession. Having a forehand estimate on whether the economy will be growing or contracting in the future, however, is important to anyone making decisions in the economy; policy makers, investors and consumers alike. Instead of looking at single variables to forecast economic activity with, the main objective of this study is to create a composite leading indicator of the Finnish economy.

A composite leading indicator (CLI), as the name suggests, is composed of several variables that contain information on future economic activity. In this thesis such variables are called leading indicators. This CLI, combined of variables with forecasting power with respect to economic activity, could then ideally provide early signals of the rate at which the economy will be growing or contracting.

This thesis aims at finding the variables that, when combined to a single model, are most apt to forecast the growth of the Finnish economy six months ahead. Initially, a portion of the vast empirical literature on forecasting economic activity is reviewed. On the basis of this discussion a pool of candidate variables, potentially containing predictive ability with respect to the Finnish economy, are gathered. A variable elimination procedure is then developed to choose the best forecasters to use jointly in the forecast model. Applying this procedure to the pool of candidate variables results in the final composite leading indicator.

This thesis is organized as follows: Section 2 discusses the choice of the dependent variable measuring economic activity and presents the forecasting model used. The following section then addresses the leading indicators, discussing variables used in previous studies, their empirical results and the economic theory underlying the use of these variables. The data used in this study and the transformations applied to the time series are the topic of Section 4. In Section 5, the selection process of the leading indicators is described. Section 6 presents the final CLI resulting from this process and further evaluates its performance. The final section concludes and presents relevant topics arising for further research.

2 Measuring economic activity and the forecasting model

2.1 The choice of the target variable

Before presenting the forecasting model used in this thesis, a brief discussion on measuring the state of the economy is in order. The classical definition of a business cycle is the co-movement of several macroeconomic variables, including production, employment etc. (e.g. Burns and Mitchell 1946). To measure economic activity and consequently the business cycle, typically, the gross domestic product of the country is used. The GDP, as a measure of everything produced in the country, broadly covers the different sectors of the economy.

For the purposes of this study, employing GDP as the dependent variable is problematic: When constructing a composite leading indicator to forecast a relatively short time ahead, say less than a year, if GDP were used, only few observations would fall into that time span, as the GDP is typically published quarterly. Furthermore, most leading indicators suggested in literature are available at monthly or higher frequencies. Thus it would be preferable to use a dependent variable that provides information of economic activity on a monthly basis. Recently, there has been discussion in favor of starting to publish the GDP monthly instead of quarterly, and an indicator of monthly GDP does exist in the U.S. (see Macroeconomic Advisers 2008). In Finland, a statistic known as the trend indicator of output is published. It is based on data similar to the ones used to calculate the GDP and attempts to portray the monthly development of the national economy. Unfortunately, the series of the trend indicator of output starts in January 1996, whereas this study has a longer time span, beginning in 1988.

Having thus ruled out GDP and its monthly counterpart, a few alternative measures of economic activity are available. One is the use of a proxy variable for GDP, one that closely follows GDP but is available on a monthly basis. This is the method adopted for this thesis: industrial production is the dependent variable chosen. It is a component of GDP and available on a monthly basis. In addition, it correlates greatly with GDP and is arguably one of its most volatile components. It is also the variable being forecast by, among others, the OECD composite leading indicators, which exist for all of the OECD member countries and a few other ones.

The problem with forecasting a single variable is that it can produce a very one-sided view of the economy or exhibit local shocks that are not economy-wide business cycle phenomena. A further issue with the use of industrial production is that its share of total production is diminishing. This naturally means that the variable being predicted is becoming a less and less important factor of economic activity.

An alternative way to avoid the problems inherent with GDP as a dependent variable would be to forecast a composite coincident indicator (CCI) instead. The purpose of composite coincident indicators is to provide a broader measure of the status quo of the economy by combining information in several variables. Existing such indicators in the U.S. include the Conference Board (2009) CCI and the Stock and Watson (1989) CCI, both of which incorporate measures of personal disposable income, industrial production, manufacturing and trade sales, and the amount of work.

For the euro area, the Centre for Economic Policy Research, CEPR, produces a model-based CCI called Eurocoin (www.cepr.org) which incorporates similar variables as the above CCIs, but also several financial variables as well as price and survey data.

If simplicity, or comparison between countries, is of importance, or no better variables are available, industrial production is a reasonable and commonly used alternative. Yet it is obvious that a broader coverage of the economy on a monthly level may be useful for forecasting business cycle movements. For this reason, and to evaluate the robustness of the obtained results on forecasts of industrial production, the forecasting ability of the developed CLI with respect to a composite coincident indicator is evaluated in Section 6.2.3. Such an indicator has been recently constructed for the Finnish economy (see Lanne and Nyberg 2009). It is a model-based composite index composed of monthly industrial production, employment, imports, exports, as well as GDP on a quarterly level.

For the time being, however, this thesis limits itself to forecasting industrial production. A transformation is imposed on this dependent variable and as a result the variable to be forecast takes the form:

$$x_{t+6|t} = 200 \ln\left(\frac{Industrial\ production_{t+6}}{Industrial\ production_{t}}\right). \tag{1}$$

The target variable is thus the approximate growth rate of industrial production over the six-month forecast horizon, standardized to annual percentage growth rates. The following sub-section presents the forecasting model used in this thesis to develop the CLI.

2.2 The forecast model

While it is possible to create a composite leading indicator in a non-model based manner, by means of normalizing, smoothing, weighting, aggregating etc., the composite leading indicator constructed in this thesis is based on a simple linear forecasting model. The former method is the one adopted by, for example, the OECD in the construction of their composite leading indicators for all of their member countries. This non-model based approach causes problems from a statistical viewpoint, however. The main one being that one does not have a sound statistical framework and thus cannot provide e.g. standard errors. The fact that the weights of the individual time series are arbitrary is also problematic.

In this thesis, a linear regression model is used to estimate the relationship between leading variables and the coincident one being forecast, industrial production. Due to the persistence typical in macroeconomic variables, lags of the variables often contain relevant information. Thus the model used here, following among others Marcellino, Stock and Watson (2003) and Marcellino (2006), is a direct forecast of the dependent variable 6 periods into the future of the type

$$x_{t+6|t} = c + e(L)x_t + \boldsymbol{F}(\boldsymbol{L})'\boldsymbol{y_t} + u_t, \tag{2}$$

where the dependent variable x_t is the six-month growth rate of industrial production as in (1), $\boldsymbol{y_t}$ is a vector of leading indicators, c is a constant, $e(L)$ is a scalar lag polynomial and $\boldsymbol{F}(\boldsymbol{L})$ is a vector lag polynomial. Using heteroscedasticity and autocorrelation consistent standard errors is crucial because the error terms u_t are serially correlated by construction, and these standard errors are later used in statistical tests in the variable selection process. Newey-West standard errors are used throughout this study.

An alternative forecasting method, often used in literature, would be to estimate a one-period-ahead model and then iterate that model forward to the desired horizon. The advantage of the specification used here, is that no model is required for the leading indicators. When iterating one period at a time, the leading indicators $\boldsymbol{y}_t$ need to be forecast simultaneously, for example using a VAR specification. Forecasting $\boldsymbol{y}_t$ using its past values and those of the coincident variable can be somewhat questionable from an economic theory viewpoint. Furthermore, in a case where there is specification error in the one-step ahead estimating, this direct forecasting model will be more robust. Although, when no such error is present, the iterated method does estimate the coefficients more efficiently (Marcellino 2006). Also, since the iterating method does not require a separate model to be specified for each forecast horizon, the time paths of the iterated forecasts may be less erratic.

One final aspect of the forecasting model remains to be discussed, namely the time series to use as inputs in $\boldsymbol{y}_t$, the leading indicators. The next section discusses previous empirical research on leading indicators and the underlying economic theory, while the leading indicator candidates selected for this study and the data transformations performed are discussed in the Section 4.

3 The leading indicators

This section begins with some examples of existing composite leading indicators and their component series. Then, leading indicators proposed in empirical literature are studied in greater detail in the form of a literature survey. The leading indicators are separated into two categories, financial and non-financial variables. This is done for practical reasons. Forecasting with non-financial variables is less common in academic studies but their use is prevalent among practitioners. Financial variables as leading indicators have been studied greatly in recent years and have often proved useful predictors of future economic activity. Another reason for the separation is that the euro area has experienced a great deal of structural changes in financial markets over the last few decades and these may be reflected in the predictive abilities of the financial variables. In both categories, the underlying economic theory, upon which the use of these variables as forecasters is based, is also discussed.

3.1 Examples of composite leading indicators

When it comes to the selection of leading variables to use in forecasting models, a number of leading indicators is preferable to a single one because of the mere fact that recessions and expansions can stem from different reasons. The notion, that every business cycle has features that are particular to it, was already suggested by Mitchell and Burns (1938) in their seminal paper, where they screened over 400 variables for leading or coinciding behavior with respect to the business cycle. Variables found to lead business cycle revivals included many measures of industrial production such as passenger car production, total paper production and total railroad operating income. Other leading variables discovered were for example the total liabilities of business failures and the Dow-Jones Index of Industrial Stock Prices. This study and their subsequent work sprouted a myriad of further research digging into the predictive abilities of different variables. As will become evident, the variables they studied in many ways reflect the types of variables still used as leading indicators today.

Before exploring individual leading indicators, Table 1 (p. 9) lists four existing composite leading indicators and the component variables they include. The first one is The Conference Board Leading Economic Index, a popular and closely followed composite leading indicator, forecasting U.S. economic activity. The second and third are the OECD composite leading indicators for the U.S. and Finland respectively. The OECD computes a composite leading indicator of growth cycles for all of its member countries as well as a few other ones. The reference series for most countries, including Finland and the U.S., is industrial production, which is used on the grounds of it being cyclically sensitive and available on a monthly basis.

The fourth CLI in Table 1 is the model-based composite leading indicator of the U.S. economy developed by Stock and Watson (1989). They use it to forecast the composite coincident indicator they also develop. Their CLI was created by screening hundreds of potential candidate variables in their time span of 1961 to 1988. Seven with the best predictive ability ended up in their final CLI.

Table 1 provides an overview of the types of variables used to forecast the future of the economy and can also be use as a point of reference in the literature survey that follows. With an understanding of typical variables incorporated in composite leading indicators, the following two sub-sections now review individual leading indicators, and the variable categories they fall into, more thoroughly.

Table 1: Component series of existing composite leading indicators

Conference Board Leading Economic Index	**OECD CLI US**	**OECD CLI Finland**	**Stock and Watson (1989) CLI**
Average weekly hours, manufacturing	Average weekly hours, manufacturing	Production tendency in manufacturing in-dex	Part-time work in non-agricultural industries
Building permits, new private housing units	Number of dwelling started	Finished goods stocks	New housing authori-zations
Manufacturers' new orders, consumer goods and materials	Net new orders for dur-able goods	Consumer price index of all items, inverted	Manufacturers' un-filled orders for dura-ble goods
Index of supplier deli-veries in manufactur-ing, vendor perfor-	Purchasing managers index	Producer price index	Change in the 10-year treasury bond yield
The S&P 500 stock index	NYSE Composite index	HEX all share index	Difference between a six-month commercial bill rate and a six
Index of consumer expectations	Consumer sentiment indicator	Consumer confidence indicator	Average weekly initial claims for unemploy-ment insurance
Interest rate spread, ten-year Treasury bonds less federal	Spread of interest rates	Spread of interest rates	The spread between a 10-year and a 1-year Treasury bond
Inflation-adjusted money supply (M2)			
Average weekly initial claims for unemploy-ment insurance			
Manufacturers' new orders, non-defense capital goods			

Conference Board (2009), OECD (2009a), Stock and Watson (1989).

3.2 Non-financial variables

Employment data

Variables that affect future employment figures are often used as leading indicators. To the extent that firms adjust labor input in addition to inventories in expectations of rises or falls in the future of the economy, these should have forecasting power. Variables such as new claims for unemployment insurance or average weekly hours of production are both included in the Conference Board CLI (see Table 1). Stock and Watson (1989) use part-time work in nonagricultural industries, which according to them measures slack work, i.e. involuntary part-time work. They find that using this variable contributes to more accurate forecasts than the two aforementioned ones used by the Conference Board.

Sales and consumption

The Permanent Income Hypothesis states that consumers determine their consumption based on expectations of their long-term income. Thus short-term fluctuations in income ought to have little or no effect on consumption. Changes in expectations of future income, due to issues such as anticipations of an economic slowdown or uncertainty about future employment opportunities, could, however, affect current consumption.

In light of this theory, sales or consumption of consumer durables might prove good leading indicators. The number of new passenger car registration, for example, is a statistic available in several countries, and the OECD in fact includes it in its composite leading indicators of some economies, such as France (OECD 2009a). The Stock and Watson (1989) CLI includes a measure of new housing authorizations. They argue that in addition to housing being the most durable consumer good, it indicates changes in the future activity of the construction sector. Of the several measures of sales and consumption studied by Stock and Watson (1989), none improved forecasting performance beyond this series. They suggest this might be due to housing starts and interest rates capturing the predictive content of consumption.

Orders and inventories

Measures of inventories seem a logical indicator of future economic activity, since they contain expectational components. They are also reflections of consumer demand since theoretically inventories ought to be sensitive to changes in expectation of future demands and future production costs.

Changes in orders received and order stock in industry also seem plausible leading indicator candidates. An increasing order stock in industry can indicate an expanding economy and vice versa. Furthermore, such variables measure investments. The theoretical link between investment and output can be found in simple Keynesian models, according to which investments have a multiplier effect on output. Since long term profit expectations are determinants of investments, investments themselves are likely indicators of future economic activity.

The value of manufacturers' new orders for example is a very common CLI component series, included in the Conference Board CLI and the OECD CLI for the U.S. For the euro area, Forni et al. (2001) produce a composite coincident indicator and a composite leading indicator to forecast it. They find that in their sample from 1986 to 2000, the change of the amount of orders in manufacturing industries is a significant leading indicator of the GDPs of all countries included, as well as of the aggregate euro area level GDP. Furthermore, their study finds that the rate of capacity utilization leads the business cycle both on the euro area aggregate level and in the cases of most member countries.

Measures of investments are also screened by Stock and Watson (1989) but they find that their addition to the composite leading indicator provides no additional predictive content. Instead they find that manufacturers' unfilled orders, as a type of negative inventory measure, contain significant predictive ability. Beyond this addition no other inventory or order stock variable provides additional predictive content.

Survey data

Forecasting future economic activity with survey data is also common. Consumer or producer confidence indicators are often used as components of composite leading indicators. Consumer or business surveys of the type, where expectations of future production are evaluated, are of particular interest. If expectations are rational, realizations ought to be close to the expectations of these economic agents.

Mourougane and Roma (2002), for example, find that the European Commission Economic Sentiment Indicator is a useful predictor of short-term GDP growth for five of the six largest euro area countries. In the case of Finland, business survey data has been used in forecasting by e.g. Kauppi, Lassila and Teräsvirta (1996). They use the quarterly business surveys conducted by the Confederation of Finnish Industry and Employers to forecast short-term industrial output by branch during the 1990–1993 recession. Their results indicate that using the business surveys improves forecasting for some branches in manufacturing.

Prices

If a Philips relation is believed to hold and there is a negative relation between prices and unemployment, price and wage data could be used to forecast output growth, which is positively related with employment. While in a U.S. context, the negative relation between inflation and unemployment has been deemed weak, a clear and stable negative relationship between their business cycle components has been discovered (e.g. Stock and Watson 1999).

Furthermore, empirical results have been obtained suggesting that in integrated markets, such as the EMU, inflation differences between member countries can have real economic effects and amplify regional business cycles (Arnold and Lemmen 2008). Although inflation differences between countries have rarely been used in forecasting economic growth, price data is used as components in OECD CLIs for several countries, including Finland.

3.3 Financial variables

While an index of stock prices as a leading indicator was already proposed by Mitchell and Burns (1938), financial variables have received increasing attention in more recent decades. Financial variables have a lot of appeal from a statistical point of view. Being market data, they are exactly measurable, available quickly and not revised at later dates. As is evident from Table 1, using financial variables as predictors of future economic growth is very common. The spread between a long-term and a short-term interest rate appears in all of the CLIs in Table 1. This variable, known as the term spread, has become the most common financial variable used to forecast economic activity and, generally, its performance seems to be quite good.

Interest rates and interest rate spreads

The expectations hypothesis tells us that long-term interest rates depend on expected future short-term interest rates. Formally, letting the annual interest rate on a long-term asset of maturity n be i_L and the expected short-term rates for periods $t = 1 \dots n$ be i_t^e,

$$(1 + i_L)^n = (1 + i_1)(1 + i_2^e) \dots (1 + i_n^e). \tag{3}$$

That is, investing in a long-term asset ought to give the same expected return as investing repeatedly into a shorter term asset. Thus an increase in the long-term rate is indication of an expected increase in short-term rates in the future.

Stock and Watson (1989) for example find that the interest rate on a 10-year government bond is a useful predictor when incorporated into their CLI. However, recent studies increasingly apply interest rate spreads for economic forecasting.

A curve, known as the yield curve, can be drawn to relate the interest rate of a bond to its maturity. The term spread, the difference between interest rates on long-term and short-term government bonds, is a measure of the slope of the yield curve. It has long been acknowledged that a declining yield curve, i.e. a negative term spread, has been a signal of a future slowdown of economic activity. The term spread for U.S. government bonds for instance, has historically become negative approximately one year before almost all U.S. business cycle peaks and, correspondingly, it has become positive again between six months and a year preceding a trough (Stock and Watson 1989).

Following the expectations hypothesis, if investors expect future short-term rates to be lower than current ones, the yield curve would invert because investors would be willing to settle for a lower interest rate on long-term bonds. However, the theoretical connection between the term spread and economic activity is somewhat debated. Stock and Watson (1989) suggest that the forecasting power of the term spread can be due simply to effective monetary policy: rising interest rates decrease inflation, whereas output is positively related to inflation as in a Philips relation. Nevertheless, Estrella and Hardouvelis (1991) show that in a regression that includes the fed rate, the main monetary policy instrument in the U.S., the predictive power of the term spread remains intact. This, according to them, is an indication of the predictive content of the term spread being mostly due to reasons other than monetary policy.

Regardless of this contradicting result, the prevailing interpretation in the context of the U.S. economy seems to be that the term spread indicates effective monetary policy (e.g. Stock and Watson 2003b). Contractionary monetary policy causes short-term interest rates to rise, and the yield curve flattens, or even inverts, as a result. High current interest rates then cause postponements of investments and a decline in economic activity.

It has been shown that the pure expectation hypothesis (3) does not completely explain the yield curve behavior (e.g. Dai and Singleton 2002). Longer term securities, while portrayals of expected future rates, can also be deemed riskier due to the longer time of having to hold the asset. Thus investors may require a premium to compensate for this extra risk.

Rosenberg and Maurer (2008) decompose the term spread of U.S. government bonds into an estimated expectations component and a term premium component. Both components are then used to forecast U.S. business cycle recessions from 1962 to 2007. According to their empirical results, the expectations component is a leading indicator, while the term premium is not. Nonetheless, their findings show that the term spread signaled an imminent recession at the end of their data set, whereas the expectations component did not. Rosenberg and Maurer (2008) interpret this as a sign of the poorer performance of the term premium. However, the end of their time span was in May 2007, while the current recession in the U.S. began in December 2007, according to the National Bureau of Economic Research Business Cycle Dating Committee. Thus their term premium signaled the latest recession, a sign that it may have predictive content as well. Hamilton and Kim (2002) have achieved a similar result, showing that both components of the term spread have predictive content.

Another proposed leading indicator, also a component of the Stock and Watson (1989) CLI, is the paper-bill spread, i.e. the spread between a commercial paper and a Treasury bill of same maturities. Stock and Watson (1989) view the forecasting ability of the paper-bill spread as being due to it being a measure of default risk. A high discrepancy between riskless assets and corporate bonds would indicate growing anticipations of business bankruptcies. Bernanke (1990) in his forecasts of U.S. economic activity finds that among the financial variables analyzed, including a term spread, the best forecasts were given by a spread between a six-month corporate paper and a six-month Treasury bill. On the other hand, Bernanke (1990) argues that like the term spread its predictive ability is mostly due to its ability to capture the effects of monetary policy.

Another fact that favors this latter argument is that the performance of both the paper-bill spread and the term spread has deteriorated since the late 1980s and they both failed to predict the 1990-1991 U.S recession. The main reason offered in literature is that, unlike previous U.S recessions, this one was not preceded by monetary tightening, whereas monetary policy is what the spreads mostly measure. (e.g. Stock and Watson 2003b and Friedman and Kuttner 1998).

The discussion so far has focused a great deal on interest rate spreads in prediction of the U.S economy. However, bearing in mind that a composite leading indicator for Finland is the goal of this thesis, a review of the performance of these variables in a European context is in order. Another reason for this is that the markets for private debt are often less developed in European countries. Thus forecasting with variables based on corporate paper interest rates is not always possible. Still, forecasts using the term spread have been conducted in Europe as well. Generally, the results have been quite encouraging.

Ivanova, Lahiri and Seitz (2000) use a Markov-switch model to predict recessions in Germany. They find that the term spread on bank bonds, the term spread on public bonds and the spread between a ten-year bond and the call rate predicted all recessions in their study from 1973 to 1998. Interestingly, they find that a spread based on the Bundesbank Lombard rate, the rate at which Germany's central bank has lent funds to commercial banks, performs significantly worse. Since this spread ought to measure mostly monetary policy, Ivanova et al. (2000) interpret that the information in market rate spreads is mostly due to factors not related to monetary policy, such as various economy-wide shocks, contradicting the conclusions of the aforementioned U.S. studies.

Moneta (2003) applies a probit model to predict recessions in the euro area as a whole and a few member countries using the term spread. When included in the model, the term spread is found to improve forecasts of euro area recessions statistically significantly in the sample from 1970 to 2001. Although, as in some of the U.S. studies above, the forecasting power of the term spread seems to have deteriorated and the 1990's recession is missed by this model as well. Additionally, the predictive ability of the term spread does not appear to be very robust to model specification, as different specifications provide the best forecasts depending on the country in question.

This non-robustness of the predictive ability of the term spread is also shown by Davis and Fagan (1997). They employ several different interest rate spreads to forecast the GDP growth of nine EU economies from the 1970s to 1992. Statistically significant instability is found in the relation and generally poor out-of-sample performance. Again, this might be due to the inability to predict the 1990s recession. Nonetheless, of all the different interest rate spreads studied, the term spread remained useful and stable in forecasts of three EU countries, namely Denmark, Belgium and the UK (Davis and Fagan 1997).

Monetary aggregates

In addition to interest rates, other financial variables have been used as leading indicators as well. In less recent studies, monetary aggregates were often considered leading indicators. Monetary aggregates measure the supply of money of different broadness. The ECB for example reports M1, M2 and M3 monetary aggregates, where M1 includes merely the currency in circulation and overnight deposits. The broader aggregates expand the definition of money, containing also deposit of longer maturity and e.g. money market fund shares in M3. Thus, they are numbered also according to the effect monetary policy can have on them. Historically they have been used to measure monetary intervention. However, financial deregulation having reduced, even eliminated, the role of money as an indicator of output growth, the short term interest rate now seems to be the proper measure of monetary policy (Bernanke and Blindner 1992). Furthermore, Stock and Watson (1989) find that including monetary aggregates in their composite leading indicator greatly worsened its out-of-sample forecast performance. Note that monetary aggregates are nonetheless used in some CLIs still today, e.g. the Conference Board CLI.

Stocks

Stock price indices are another type financial variable that has been used as forecasters for decades. According to the present value theory of stock prices, the price of the stock should reflect the expectation of the future earnings of the corporation. Thus stock price indices ought to be an indicator of the growth potential of publicly traded corporations and, more broadly, the entire economy. As can be seen from Table 1 (p. 9), they are very common in composite leading indicators, appearing in three of the four ones presented.

Stock and Watson (1989) do not include stock price indices in their CLI, finding that the marginal predictive content of including stock prices to the forecasting of their CCI is modest. In their opinion, this indicates that the expectational role of stock prices is already captured by other financial variables.

Using the variance of stock returns as a leading indicator for U.S. GDP has been suggested by Campbell et al. (2001), who find that high volatility signals low growth in the subsequent quarter. They interpret this as depicting increased doubts about short-term economic prospects. Nonetheless, this predictive power of stock market volatility is substantially weakened in out-of-sample analyses (Guo 2002).

In Europe, Andersson and D’Agostino (2008) forecast euro area real economic growth using sectoral stock prices, arguing that some stock market sectors ought to be more closely linked to the business cycle than others. In their sample from 1973 to 2006 they find that, while the term spread has been the best predictor overall, sectoral stock prices often improve forecasting accuracy on longer, over one year, horizons. A further finding they make is that all financial variables studied predict more accurately after the adoption of the common currency. Andersson and D’Agostino (2008) suggest that this might be due to the lower risk premia induced by the adopting of the euro. As the premia shrunk the relative information content in financial variables grew.

Another interesting stock market variable to forecast the economy with would be dividend yields. Taipalus (2006), for example, creates an indicator based on dividend yields to track periods when stock prices have deviated from their fundamental levels. This so called bubble indicator has some forecasting ability in detecting asset price bubbles in Finland and the U.S making it a potential leading indicator, since in the past bursting bubbles have had destabilizing effects on the financial system, and consequently real activity.

International data

One of the most recent methods for forecasting with financial variables is the use of economic tracking portfolios. Junttila (2007) develops four portfolios, consisting of financial assets, to track industrial production respectively in the U.S. and three euro zone countries. This is done by regressing country specific industrial production growth on stock returns, currency returns and interest rates in the different countries. The regression coefficients determine the weight each asset gets in the portfolio. The resulting portfolio is then used to forecast future output growth. Junttila (2007) finds that the use of such an open economy data set improves forecast accuracy substantially. The inclusion of currency returns in the portfolios, to capture the effect of international linkages, is particularly influential.

An important finding in the above study is the impact of the use of international data. Since any country can be expected to be affected by changes in other economies through trade or financial markets, foreign exchange rates and interest rate differences between countries can be significant leading indicators.

Assuming the exchange rate reflects the upcoming rise or fall in foreign investor demand of domestic goods, the depreciation of a currency relative to other currencies ought to be associated with an increase in net demand for domestically produced goods relative to foreign goods. Stock and Watson (1989) for instance find that a trade-weighted nominal exchange rate between the U.S and its major trade partners makes a small positive contribution to their composite leading indicator.

Furthermore, Osborn and Sensier (2002) find that in predicting business cycle phases international linkages are important. They forecast expansion and contraction probabilities of U.S. and EU country business cycles from 1970 to 2002 using financial data from the different countries. Their empirical results indicate that U.S. financial variables are influential for Germany and the UK, and German variables are significant for forecasts of France and Italy. For a small economy that depends greatly on exports, such as Finland, these types of foreign financial variables, are certainly interesting leading indicator candidates.

3.4 Discussion

In the above literature, a wide selection of proposed leading variables has been addressed. The matter of which variables are leading indicators and which ones do not contain information is debated and seems to be time and country-dependent. Stock and Watson (2003b), for example, research the forecasting power of various asset prices on GDP growth in the U.S., Canada, Japan and four E.U. countries. In their data set ranging from 1971 to 1999, they find that several financial variables have been useful for some countries and in some periods but no single asset price is a reliable predictor of output growth over multiple decades. According to the literature reviewed here, the term spread seems to come closest to this though, being a useful predictor in most studies mentioned, even though its predictive ability in the 1990s and since has been contested.

International linkages also seem to be important. Particularly this might be the case when forecasting a relatively small economy. Foreign country variables, both financial and non-financial ones, may then contain predictive ability.

Non-financial variables, while more common among practitioners, nonetheless appear in the composite leading indicators presented as examples. Economic theory that advocates the use of these variables can be found for all of the non-financial leading indicator categories explored above. Yet it would seem, in light of the literature surveyed, that financial variables provide the most accurate predictions of the future — in the euro area more so after adopting the euro. Being market data and not subject to revisions is another advantage of financial variables over other indicators, although there are some non-financial variables that are unrevised as well.

A contradicting result on the usefulness of financial variables in economic forecasting has however been obtained. Forni et al. (2003) forecast euro area industrial production with a time span of 1987–2001. Using their large data set of over 400 monthly time series, they find that the inclusion of financial variables in their forecast models provides no additional information over the non-financial ones.

Clearly, the evidence on the performance of financial variables as leading indicators is somewhat mixed as well. A point worth noting, however, is that all of these studies use different models to perform their forecasts. Since, in some cases, results differ considerably, model selection seems to be a particularly relevant issue. Forni et al. (2003) reassert this view of empirical evidence not being robust to model specification, sample choice or forecast horizons, calling it a puzzle worth investigating for economic theory.

The variable selection process applied in this thesis to create the composite leading indicator is described in Section 5, while the performance of the CLI itself comes under evaluation in Section 7. The next section, however, discusses the variables used in this study. These candidate leading indicators were chosen based on the myriad of research and empirical results surveyed above.

4 The data

A general problem with the selection of potential leading indicators for this study is the availability of data. Time series on several of the variable types discussed in the above literature survey, do not exist for the Finnish economy, or the time series are not of sufficient length. The data used in this study span from January 1988 to December 2009. All time series used are monthly.

4.1 The candidate variables

As international linkages appeared important in the above literature, whenever available, this thesis also evaluates EU, Euro area or even U.S variables in each of the variable categories. The fact that the rather small and export-intensive economy of Finland depends greatly on the economic situation of its main export countries, further suggests that foreign country data may contain relevant information. Even if a direct connection to the Finnish economy were not to exist through exports, if a particular variable is a leading indicator for its own economy, and the economy of Finland lags behind the global economy, such a variable ought to be of some use for forecasting the Finnish economy as well.

For a comprehensive list of the data used in this study, the reader is referred to Tables A1 and A2 in the Data Appendix. The following briefly describes the choice of candidate variables in the categories corresponding to the above literature survey.

Employment data

Employment statistics are somewhat problematic as typically one would expect employment to lag economic activity instead of lead it. In light of the literature surveyed, the employment data included in this study are mainly short-term employment market statistics that have been used to some success in previous research. Whether employment data contains relevant information with respect to future Finnish economic activity is assessed by including employment service statistics from the Finnish Ministry of Employment and the Economy in the pool of candidate variables. Series included are monthly new vacancies, unfilled vacancies, the amount of unemployed workforce and the unemployment rate.

Orders, inventories and survey data

Data on the monthly value of new orders in manufacturing is a common forecaster in literature. However, for forecasts of the Finnish economy, the existing statistic on new orders in manufacturing is too short, having been published in its current form only since 2005. The time series on inventories, while longer, are quarterly, whereas this thesis works with monthly data. To be able to use monthly time series of sufficient length to measure orders and inventories, this thesis uses survey based data instead.

The time series on the balances of the individual questions of the monthly business surveys conducted by the Confederation of Finnish Industry and Employers for both the industry and construction sectors are used. These questions include assessments on order-book levels, export order-book levels, stocks of finished products, as well as production, sell price and employment expectations. The composite indices based on these questions are also used but in many cases provide worse forecasts than the balances of certain individual questions. In addition to these Finnish surveys, the predictive abilities of the European Commission produced EU and euro area business surveys, with respect to Finnish industrial production, are studied. These EU and euro area level surveys contain identical questions to the Finnish business surveys.

The forecasting performance of the Economic Sentiment Indicator, found useful for some euro area countries' GDP forecasts by Mourougane and Roma (2002), is evaluated in this thesis as well; again the indicators of Finland, the EU and the euro area are all studied as leading indicator candidates.

The consumer confidence indicator of Finland, while definitely a valid candidate leading indicator, has only been published since 1995 in its present form. Nevertheless, most of the questions in the survey have remained unchanged since the late 1980s. The time series of the balances of the individual questions are thus used as candidate leading indicators in this study, in order to obtain series of longer length.

Sales and consumption

As was discussed in the previous section, economic theory and to some extent empirical research suggests time series measuring sales or consumption of durable goods may be of particular interest. The ones evaluated in study include new passenger car registrations and granted construction permits. Data on the retail sales of automotive fuel and motor vehicles are also studied, as well as the retail sale of watches and jewelry, which, as types of luxury items, might capture consumer behavior similar to the consumption of durable goods.

Price data

The OECD composite leading indicator for Finland incorporates both the producer and consumer price indices. In this study, the consumer and producer price indices, as well as the wholesale price index and some individual components of the consumer price index, such as the energy and the food price indexes, are included.

Energy prices seem a particularly valid leading indicator candidate because energy is often a key input in Finnish industrial production, which is particularly energy intensive. In addition, Forni et al. (2001), when constructing their CLI for the euro area, found the value of energy consumption to be a leading indicator for several euro area countries. Food prices on the other hand could have real effects through consumers' free disposable incomes.

Financial variables and international linkages

The financial data used in this study include three stock market indices: the S&P500, the index of European shares Dow Jones Eurostoxx 50 and the all-share price index of the Helsinki Stock Exchange. In addition, several interest rates and interest rate spreads from the U.S., Finland and Germany are used.

This thesis also evaluates whether the spreads between yields on different governments' bonds provide some predictive ability. The spread between interest rates on U.S. or German government bonds and a domestic government bond is included. These portray the risk premium investors require for holding the domestic bonds. Uncertainty in domestic financial markets and expectations of the future of the domestic economy could then be reflected in this risk premium.

The list of candidate variables also includes open economy variables such as the effective real exchange rate of the euro and Finnish real competitiveness indices. In addition, the FIM/EUR – USD nominal exchange rate is used. The time series on all interest rates, stock price indices and the nominal exchange rate are computed as monthly averages of daily closing prices or equivalent values.

4.2 Data transformations

Most studies on leading indicators typically find both the dependent and the leading variables to be $I(1)$ and work with first-order differences. Other methods include the use of different kinds of frequency-domain filters. The difference between the two involves a broader discussion on what one interprets as the business cycle.

Two varying definitions of a business cycle exist in literature. The classical cycle measures absolute changes in the dependent variable, whereas the growth cycle is viewed as fluctuations of the target variable around its long-term growth trend (Artis et al 2004). Thus dealing with the classical cycle corresponds to taking first-order differences of the series, and a classical cycle recession would be indicated by a decline in the target variable in absolute terms. A growth cycle recession on the other hand need not be exhibited by a decline in absolute terms as long as growth slows down below its long-term trend. The growth cycle can be extracted using detrending methods such as frequency-domain filters.

A comprehensive discussion on such filters is beyond the scope of this thesis but essentially their purpose is to filter out the components that are not of interest from the time series, that is the long term fluctuations in the series and preferably also the high-frequency irregular fluctuations that are not in line with the duration requirements one has set for a full cycle. The OECD, for example, filters the variable they forecast as well as the component series of its composite leading indicators. They define that the duration of a business cycle can be anywhere between 12 and 120 month and thus apply a frequency domain filter that allows for the time series to contain variation only between these periodicities (OECD 2009b). Consequently, their CLIs attempt specifically to forecast growth cycles of the countries in question.

When it comes to forecasting, the studying of growth cycles can be more informative, even though it complicates the analysis due the filtering methods necessary for the extraction of the cycle. The forehand knowledge of whether economic growth will be slowing down in the future is important for policy makers, even if no actual contraction of the economy in absolute terms is in sight.

However, an issue of concern when using filtering methods, in the case of composite leading indicators, is the selection of which of the leading indicators to filter as well. While it has intuitive appeal to filter the leading series when the target variable is filtered, it may be questionable to do so with some variables, for example financial ones. Another problem with filtering methods is in interpreting the dependent variable. When forecasting a filtered variable, one is essentially forecasting a business cycle component of that variable, which is a far more abstract concept.

For these reasons, this thesis opts to use unfiltered series instead. The six-month growth rate of industrial production is the dependent variable being forecast, as was described in Section 2.

The time series in this thesis span from January 1988 to December 2009, apart from data on construction permits, which are available from 1990 forth, and employment statistics, available from 1991. The series of the FIM/EUR–USD exchange rate also begins in 1991. All of the data used in this study is monthly and, apart from financial data, seasonally adjusted. In most cases, readily available seasonally adjusted series were used. The Tramo/Seats procedure was adopted to adjust the few series for which only original series were available. This group mainly consists of sales data.

DF-GLS tests (see Elliot, Rothenberg and Stock 1996) were run on the candidate variables. Most variables could not reject the null hypothesis of having a unit root, in which case first order differences were taken. For the few variables that appeared stationary in levels, rejecting the null hypothesis of a unit root, both levels and differences were studied. For a large majority of the variables, treated as $I(1)$, to approximate monthly growth rates, the first order differences were taken on the natural logarithms of the variables. In the cases of some series, mainly financial or survey data, the first-order difference was imposed on the original series. All differenced series rejected the null hypothesis of a unit root and no further transformations were conducted.

To develop composite leading indicators from this rather large list of potential component series, it had to be narrowed down in several ways. The next section describes the elimination process eventually leading to the construction of the final CLIs.

5 The variable selection process

The selection of the leading indicators is conducted in two steps. First, the large pool of potential leading indicators is narrowed down to a base set of variables. In the second step, the variables with the best predictive ability in this base set are chosen using a 3-stage pseudo out-of-sample forecasting procedure, described in the latter of the following two sub-sections.

A good property of the method of variable selection applied in this thesis is that it requires little subjective judgment once the candidate series have been gathered. The final CLI is obtained in a rather straightforward and algorithmic manner following the procedure outlined below.

5.1 Step 1: Forming the base set of variables

The first step in the variable selection procedure applies in-sample and out-of-sample measures to the pool of candidate variables, then conducts a small amount of further screening and finally results in a base set of relevant variables for further evaluation. Figure 1 provides an overview of this process, which is described below in greater detail.

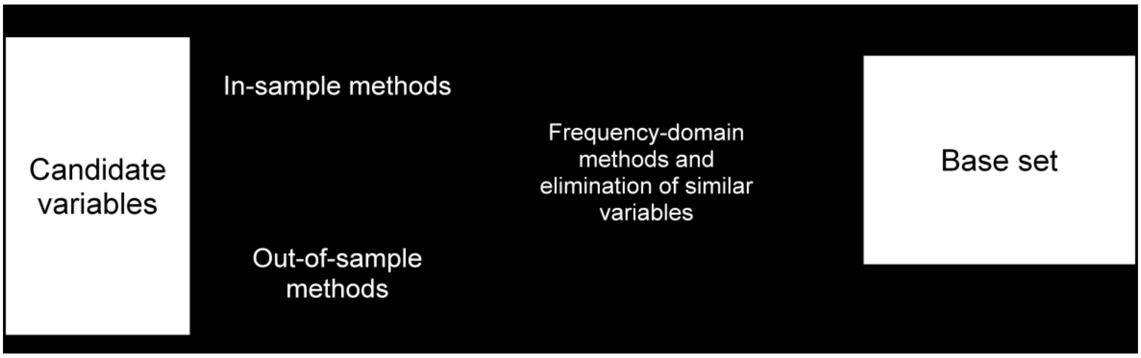

Figure 1: Step 1 of the variable selection procedure.

Initially, in-sample measures are used on the pool of candidate variables. This entails estimating regressions of the type (2) using the full sample from January 1988 to December 2009. The model is repeated below for convenience[1]:

$$x_{t+6|t} = c + e(L)x_t + f(L)y_t + u_t. \qquad (2)$$

Going through each candidate variable, one at a time, the six-month growth rate of industrial production is regressed on its lags and a leading indicator candidate and its lags. The lag length of the variables is determined using the Schwarz information criterion.

The statistical significance of the regression coefficients of the candidate variable and its lags in these bivariate regression models is then evaluated by means of an F-test, using Newey-West standard errors. Those leading indicator candi-

[1] The vector notation is dropped as y_t is now a single variable in these bivariate models that include only one candidate variable, the dependent variable and their lags.

dates, whose effect on industrial production is statistically significant are added to the base set of variables.

The second criterion used in determining whether a variable should be added to the base set is the marginal forecasting ability of each candidate variable beyond the lags of the dependent variable. To do this, pseudo out-of-sample forecasts of models of the type (2) are conducted, inputting a single variable at a time as y. Because the out-of-sample performance of a variable is used as a selection criterion, these forecasts are calculated using only a sub-sample of the data from January 1988 to December 2006. This is done in order to be able to perform out-of-sample forecasts of the final constructed CLI in a sample that has not been used in variable selection. Section 6.2.1 assesses the performance of the developed CLI in this true out-of-sample setting.

The pseudo out-of-sample forecast method used throughout this study is a rolling window scheme, where the number of previous observations used to compute the following pseudo out-of-sample forecast is constant. A fixed 11-year estimation window, which corresponds to half the size of the full sample rounded to full years, is used. These forecasts are conducted by initially estimating the model (2) for the first 11 years and producing a forecast of the dependent variable six months ahead. The squared error of this forecast is calculated as the squared difference between the forecast and the actual value of the dependent variable that occurred

$$\hat{\varepsilon}^2_{t+6} = (x_{t+6} - \hat{x}_{t+6|t})^2. \tag{3}$$

The estimation window is then moved one period ahead; the next observations in the sample are added and the oldest ones are dropped. Again, forecasts using the model are performed and the squared errors are calculated. This procedure is repeated until the end of the sub-sample in December 2006. The lag-length used for the variables is re-determined recursively at each step using the Schwarz information criterion.[2] Advancing in this manner, through the sample until December 2006, produces a sequence of squared forecast errors. The

[2] The variable selection method seems rather robust to the choice of the lag-selection criterion; the entire procedure was repeated using the Akaike information criterion, resulting in a composite leading indicator with nearly all the same variables and equal forecasting power. For brevity, only the results of the SIC procedure are presented.

mean of these squared forecast errors (MSFE) is computed and compared to the mean squared forecast error produced by an autoregressive benchmark model of the type

$$x_{t+6|t} = c + e(L)x_t + u_t, \tag{4}$$

where the lag length is again determined recursively at each step.

The criterion measuring marginal predictive ability is then simply the ratio of these two MSFEs. Formally

$$\frac{MSFE}{MSFE_0} = \frac{\frac{1}{T_2 - T_1 - 6 + 1}\sum_{t=T_1}^{T_2-6} \hat{\varepsilon}_{t+6}^2}{\frac{1}{T_2 - T_1 - 6 + 1}\sum_{t=T_1}^{T_2-6} \hat{\varepsilon}_{0,t+6}^2}, \tag{5}$$

where T_1 and $T_2 - 6$ are the first and last period over which the pseudo out-of-sample forecasts are computed, $\hat{\varepsilon}_{t+6}^2$ and $\hat{\varepsilon}_{0,t+6}^2$ are the squared forecast errors of the pseudo out-of-sample forecast, of period $t + 6$, as in (3), for the candidate model and the benchmark model respectively. (Stock and Watson 2003a,b). Whenever this ratio is less than one, the candidate forecast model is estimated to have performed better than the benchmark. This relative MSFE is calculated for all candidate variables, that is the procedure is repeated for all of the variables inputting a single variable at a time as y in (2). All variables that improve upon the AR-forecasts are added to the base set.

Performing the pseudo out-of-sample forecasts of bivariate models for each candidate leading indicator in this way yields a list of variables that improve forecasts upon the AR-benchmark. This list is augmented with the variables that were selected using in-sample criteria.

The pseudo out-of-sample forecast scheme used here has advantages when, for example, large structural changes have taken place in the economy towards the beginning of the sample period, or for other reasons, one simply does not want the older observations to be overly emphasized. An alternative pseudo out-of-sample forecast method would be to use an expanding estimation window, where all previous data is used in performing the forecasts. This alternative was explored but generally resulted in less accurate forecasts of industrial production growth. The rolling window scheme is thus used throughout this thesis.

Finally, the set of variables obtained using in-sample and out-of-sample methods is rendered smaller by evaluating the coherence and phase leads between the predictor variables and industrial production growth; time series exhibiting very low coherence and no phase leads are eliminated from the set. Particular attention is paid to variables from the same categories, measuring similar concepts. Those with the best predictive ability are retained. This shortened list of variables is then finally defined as the base set.

5.2 Step 2: The three-stage variable selection procedure

The second step involves selecting those variables from the base set that, when combined, produce the most accurate forecasts. Stock and Watson (1989), in the creation of their CLI, use a modified stepwise regression procedure, to select the variables from a similar set of leading indicator candidates. They develop several different leading indexes by evaluating the full-sample R^2 and a subsample R^2. The variables appearing most often in the indexes are selected to be the components of their CLI.

A method of this type was attempted, but resulted in a selection of variables with poor out-of-sample predictive ability. This is not surprising as typically in-sample measures in no way guarantee good out-of-sample performance (e.g. Zarnowitz and Braun 1989 and Marcellino 2006).

However, using merely the out-of-sample performance of the bivariate forecast models as a criterion in this type of elimination process would not seem to result in the best list of predictor variables either. This is due to the fact that while a single variable may be poor at forecasting individually, it may improve forecasts when combined with the use of another variable. In other words, it may still have some marginal predictive ability.

For these reasons, the method undertaken here is similar to the Stock and Watson (1989) method, but differs in that it emphasizes the pseudo out-of-sample forecasting performance, instead of the in-sample performance, of candidate leading indicators jointly.

Initially, the three variables that produced the most accurate pseudo out-of-sample forecasts in step 1, as measured by the relative MSFE, are selected. For each of these three bivariate forecast models, an additional predictor is added from the base set one at a time and the relative MSFEs against the same AR benchmark are calculated. Formally this entails producing pseudo out-of-sample forecasts as in the previous step but now using models of the type

$$x_{t+6|t} = c + e(L)x_t + \boldsymbol{F}(\boldsymbol{L})'\boldsymbol{y}_t + u_t, \tag{7}$$

where $\boldsymbol{y}_t = [y_1 \quad y_2]'$, y_1 is one of the three most accurate predictors in the base set and y_2 goes through all other variables in the base set.

The sample used in these forecasts is the same as in the previous section - the last three years are removed. The forecasting method also remains the same; an 11-year rolling estimation window is used. The lag-length of the variables is again determined recursively at each step of the pseudo out-of-sample forecasts using the Schwarz information criterion.

For each of these three variables initially chosen, the three trivariate models, with two predictors in each, producing the most accurate forecasts, as measured by the relative MSFE, are then selected. Finally, for each of these nine trivariate models, one additional predictor variable is added from the base set. All variables in the base set are again gone through and the relative MSFEs of the pseudo out-of-sample forecasts are calculated. The vector $\boldsymbol{y}_t$ in the forecasting model is now defined as $\boldsymbol{y}_t = [y_1 \quad y_2 \quad y_3]'$, where y_1 and y_2 come from the nine trivariate models and y_3 goes through all other variables in the base set.

Again, the best three of these three-predictor models for each of the nine trivariate models are chosen, resulting in a maximum of 27 models with three predictor variables. The actual number of models may be less as it is possible that less than three variables are such that they provide additional predictive ability when added. In these cases only the models that actually improve forecasts are used.

Figure 2 summarizes the procedure so far: Starting from the base set, the three variables producing the most accurate forecasts are selected. The models are augmented by adding one additional predictor variable from the base set. Again, the three best models for each of the three initially chosen variables are selected. A variable from the base set is again added to all of these models and finally the three best models for all of the nine two-variable combinations are selected.

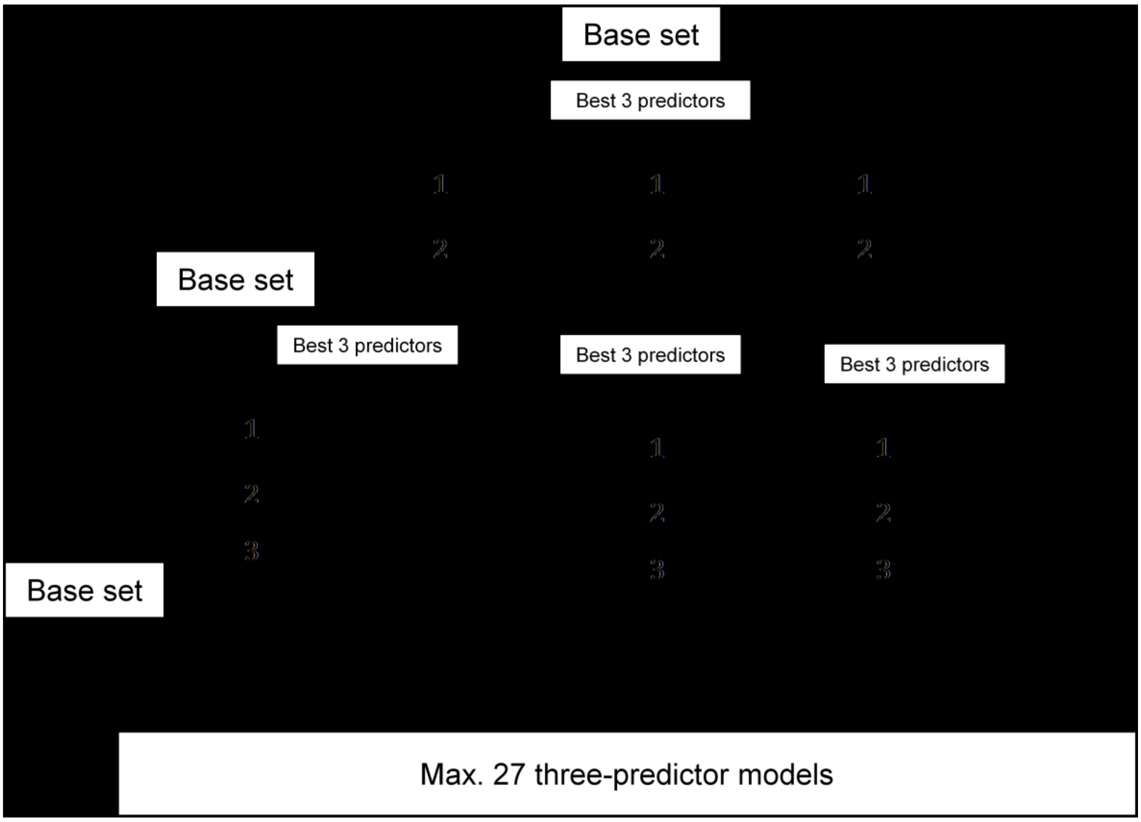

Figure 2: Step 2. At each of the three stages, the three variables that improve pseudo out-of-sample forecasts the most, when added to the models, are selected resulting in a maximum of 27 models with three predictor variables.

Finally, the variables appearing most often in these three-predictor models, and performing best when combined, are used as the component series of the final composite leading indicators.

Note that the method undertaken in this study has its downfalls as well. Interpreting that the forecast performance of a variable improves upon the benchmark when the relative MSFE is less than one, is somewhat questionable, as this can merely be due to sampling variability (Stock and Watson 2003b). This naturally means that the variable selected may not actually forecast statistically significantly better than the benchmark. To determine whether the relative MSFE is statistically significantly less than one requires testing the null hypothesis of equal predictive ability. When the benchmark model is nested – a special case of the candidate model – as is the case here, this can be done for example using methods developed by Giacomini and White (2006)[3]. In this thesis, the forecast errors of the final CLI model are tested by applying the Giacomini and White (2006) test of conditional predictive ability, which evaluates whether the forecasts of the CLI model will be more accurate than the benchmark model in the future, given its past performance.

6 Results

6.1 The composite leading indicator

Conducting the first step of the procedure, outlined in the previous section, on the candidate leading indicators yields a base set of variables. This set contains a total of 54 variables from most variable categories.

The second step of the elimination procedure, summarized in figure 2 (p. 32), results in 24 three-predictor models composed of 16 different variables. The

[3] Another test often applied in literature to assess whether the obtained forecast errors are statistically significantly smaller, when the benchmark model is nested as is the case here, is the one developed by Clark and McCracken (2001). Advantages of the Giacomini and White (2006) method are easy computation and the fact that its limiting distributions are not context-specific.

variables in these models, as well as the relative MSFEs the pseudo out-of-sample forecasts produce are presented in Table 2 on the next page. The three best forecasters from the bivariate forecasts are the interest rate on a 10-year U.S. treasury bill, the Finnish construction sector survey question reflecting the share of answerers stating the shortage of materials as the main hindrance to production and the amount of unemployed jobseekers.

At the next level, the three variables that improve the forecasts the most for each of the three bivariate models are presented on the rows labeled $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. At the third level, the three variables providing the most marginal predictive ability for each of the 9 trivariate models are tabled on the rows labeled $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Note that in the case of some models, no variable from the base set improves the forecasts when added.

Table 2. The component series of the models of the elimination process, relative MSFEs of the forecasts in parentheses.

$\boldsymbol{y} = [y_1]$	**Yield on 10-year U.S. government bonds (0,893)**		
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$	FIN consumer confidence survey: Ability to make major purchases (0,815)	Granted construction permits (cubic volume) (0,846)	Consumer price index (0,851)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Total jobseekers (0,704)	FIN consumer confidence survey: Ability to make major purchases (0,760)	Total jobseekers (0,763)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Unfilled vacancies at the end of the month(0,675)	Unfilled vacancies at the end of the month (0,705)	Unfilled vacancies at the end of the month (0,724)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Total unemployed (0,716)	FIN construction sector survey: % stating material shortage as main hindrance to production (0,765)	Total unemployed(0,788)
$\boldsymbol{y} = [y_1]$	**FIN construction sector survey: % stating material shortage as main hindrance to production (0,914)**		
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$	EU construction sector survey: Evolution of orders (0,853)	Total jobseekers (0,861)	FIN consumer confidence survey: Ability to make major purchases (0,877)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Fuel sales, value index (0,788)	Yield on 10-year U.S. government bonds (0,783)	Fuel sales, value index (0,788)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	New passenger car registrations (0,795)	EU construction sector survey: Evolution of orders (0,840)	EU construction sector survey: Evolution of orders (0,805)
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	FIN consumer confidence survey: Ability to make major purchases(0,805)	U.S. 3-month interest rate (0,851)	Difference between yields on 10-year U.S. and Finnish government bonds (0,785)
$\boldsymbol{y} = [y_1]$	**Unemployed jobseekers (0,917)**		
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$	EU business survey: order stock in industry (0,915)	FIN construction sector survey: % stating material shortage as main hindrance to production (0,896)	
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	FIN consumer confidence survey: Ability to make major purchases (0,766)	FIN consumer confidence survey: Ability to make major purchases (0,840)	
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Yield on 10-year U.S. government bonds (0,831)	Yield on 10-year U.S. government bonds (0,757)	
$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$	Euro area consumer confidence indicator: ability to make major purchases in next 12	U.S. 3-month interest rate (0,831)	

Clearly, the Finnish consumer confidence indicator question on the ability to make major purchases is an important indicator, appearing six times in the models and improving the forecasts quite substantially when used. This is not surprising as it ought to be a reflection of consumers' economic situations, free disposable income and, consequently, future consumer demand.

Two variables measuring sales appear in Table 2, fuel sales and new passenger car registrations. Neither of these improves forecasts when the abovementioned measure of consumers' disposable income is utilized. This would indicate that perhaps this survey question already captures the effect of consumer demand on output growth.

Another observation to be made from Table 2 is that several of these leading indicators are U.S. interest rates. This finding suggests that perhaps the Finnish economy lags its American counterpart, of which these interest rates may be leading indicators. Moreover, the forecast power of these U.S. interest rates may reflect the fact that, historically, several global recessions have started from the U.S. economy — or even U.S. financial markets. The yield on a 10-year U.S. government bond in particular seems to provide predictive content. While the term spread is a more common forecaster, this long rate has also been found a good forecaster in past studies. It is in fact a component of the Stock and Watson (1989) CLI.

The difference between U.S. and Finnish long-term interest rates appears in several models of Table 2 as well. This risk premium on Finnish government bonds has marginal predictive ability beyond the 10 year U.S. rate as well. Both of these financial variables end up in the final CLI.

It is important to note that even though both the U.S. long rate and the difference between the U.S. long rate and the Finnish long rate are used as forecasters, and the model is a linear regression model, this is not equivalent to simply forecasting using the Finnish long rate. As the Schwarz information criterion is applied to set the lag length, the U.S. long rate gets between 0-1 lags in the forecast models, while 3-4 lags of the interest rate difference are used throughout the recursive pseudo out-of-sample forecasts. Consequently, applying the final CLI model is in fact equivalent to forecasting using the Finnish long rate and lagged values of the U.S.-Finnish long rate difference.

In addition to interest rates, short term employment market data seem to be important leading indicators as well. Employment situations affect consumers' economic situations through their free disposable income and consequently indicate consumer demand. Furthermore, employers may vary the amount of labor to employ in expectation of the future economic situation. The amount of unfilled vacancies at the end of the month for example improves several forecasts in Table 2. Variables measuring the amount of jobseekers also appear, however their predictive ability is negligible when vacancies are included in the regression. Nevertheless, the variable measuring unfilled vacancies remains a component of the final CLI. As it is more affected by employer decisions than the amount of jobseekers, the latter of the two channels of effect might be the stronger one.

Construction sector variables are likely leading indicators, as the industry is particularly volatile and cyclical. They may also reflect consumer demand or consumers' economic situations, as large construction projects employ workers for long periods of time, and housing construction may provide some indication of consumers' demand for housing. Two construction sector variables are included in the final CLI, namely the cubic volume of granted construction permits and the percentage of answerers in the Finnish construction sector surveys stating material shortage as the main hindrance of production. The EU construction survey variables appearing in Table 2 do not improve forecast when these two Finnish time series are used.

The fact that the EU-level order stock of industry has forecasting power could, in addition to indicating future production, be an indication of the EU economy leading the Finnish one. Yet a contradicting result to this has been obtained by Forni et al. (2001), who find that, in fact, the Finnish GDP and Finnish investments lead their euro area counterparts. Although, they state that the phase leads manifesting in their data may be largely due to the exceptional recession that occurred in Finland in 1991–1993.

The final CLI combined from these variables takes the form presented in Table 3. After the addition of these variables, none of the other ones improve the forecasting performance and, correspondingly, the omitting of any variable worsens the forecasts. The relative MSFE of pseudo out-of-sample forecasts using this final CLI model is 0,5800 with a Giacomini and White (2006) test p-value of 0,0006.

Table 3. The components of the final CLI.

Component series	Data transformation
Difference between yields on U.S. and Finnish 10 year government bonds	Monthly change
EU business survey: order stock in industry.	Level
FIN construction sector survey: % stating material shortage as main hindrance to production	Level
FIN consumer confidence survey: Ability to make major purchases	Monthly change
Granted construction permits (cubic volume)	Monthly growth rate
Unfilled vacancies at the end of the month	Monthly change
Yield on 10 year U.S. government bonds	Monthly change

6.2 Evaluating the performance of the composite leading indicator

The empirical results of the various studies on leading indicators presented in Section 3 gave the impression that the evidence on the predictive abilities of different variables is somewhat mixed and depends on the study and model at hand. It is therefore obvious that certain robustness evaluations on the forecast performance of the obtained indicator are in order.

Because variable selection was conducted using the data until 2006, it is now possible to study the out-of-sample performance of the CLI model, with the seven component series listed in Table 3 as inputs, in a sample that has not been used for variable selection. This is the issue studied in the first of the following subsections.

A second matter of some concern, relevant whenever conducting economic forecasting, is data revisions. Macroeconomic data is periodically revised as newer data becomes available. The forecasting power of the constructed CLI on real-time data of industrial production is evaluated in the latter of the following subsections. This analysis assesses, whether the composite leading indicator would have been useful, had it been used in the past.

The final two subsections address the stability of the CLI model and its ability to forecast a composite coincident indicator of the Finnish economy, respectively.

6.2.1 Forecasting in an unused sample and the latest recession

Since the variable selection procedure in this thesis, is conducted using also the out-of-sample forecast performance of the variables, evaluating the performance of the constructed CLI in a sample that has not been used for variable selection is necessary. The last three years of the full sample, 2007–2009 are used to this end. An advantage of this particular sub-sample is that it allows the assessment of the forecast performance of the composite leading indicator during latest recession.

The chosen sub-sample is quite particular in that it contains a period of rather high growth and a steep fall in production occurring in a very short time. This can be seen in Figure 3, depicting the natural logarithm of industrial production, as well as in Figure 4, where the six-month growth rate of industrial production is plotted.

Figure 3: The logarithm of monthly Finnish industrial production 2007-2009.

Figure 4: The six-month growth rate of industrial production standardized to annual percentage growth.

To evaluate the forecasting performance of the CLI in this sample from 2007 to 2009, the composite leading indicator model, constructed above, is applied to forecast the six-month growth rate of industrial production from 2007–2009. As before, the relative MSFE is calculated against a benchmark AR forecast but now for this three-year sample. Again, an 11-year rolling window scheme is used and the lag-length of each variable is determined recursively using the Schwarz information criterion.

In addition to using the full model, the component series of the CLI are also used individually to forecast the dependent variable in this sub-sample, in order to assess whether the performance of the CLI can be contributed to some select variables. Another benefit of this analysis of the forecasting power of the component series is in obtaining information on which variables forecasted most accurately during the latest recession.

Table 4 lists the results of these forecasts. As can be seen, judging by the relative MSFEs, the composite leading indicator is not able to provide more accurate forecasts than a simple AR model in this sub-sample. While, this may be an indication of the relatively poor forecast performance of the CLI in general, it ought to be borne in mind that the sample in which it is now evaluated is quite particular. Looking at the forecasts of the components of the CLI, it is evident that their individual forecast performance also deteriorates greatly in this sub-sample. Of the seven components it would seem only one, namely the volume of granted construction permits, improves upon the AR forecasts. Even this result would not seem statistically significant. The p-value of the Giacomini and White (2006) test of conditional predictive ability is 0,0921.

Table 4: Relative MSFEs produce by forecasts using the full model and its individual components series over the 2007 - 2009 sample.

Forecaster	Relative MSFE
Composite leading indicator	3,222
Difference between yields on U.S. and Finnish 10 year government bonds	1,001
EU business survey: order stock in industry	1,100
FIN construction sector survey: % stating material shortage as main hindrance to production	1,469
FIN consumer confidence survey: Ability to make major purchases	1,050
Granted construction permits (cubic volume)	0,975
Unfilled vacancies at the end of the month	1,031
Yield on 10 year U.S. government bonds	1,173

To assess whether failure to forecast the single steep drop in production at the end of 2008 is the reason for this poor performance, the squared forecast errors of the forecasts using the individual components of the CLI are plotted in Figure 5. The squared errors have been scaled by dividing them with the MSFE of the AR model forecasts, which is described by the horizontal line at 1.

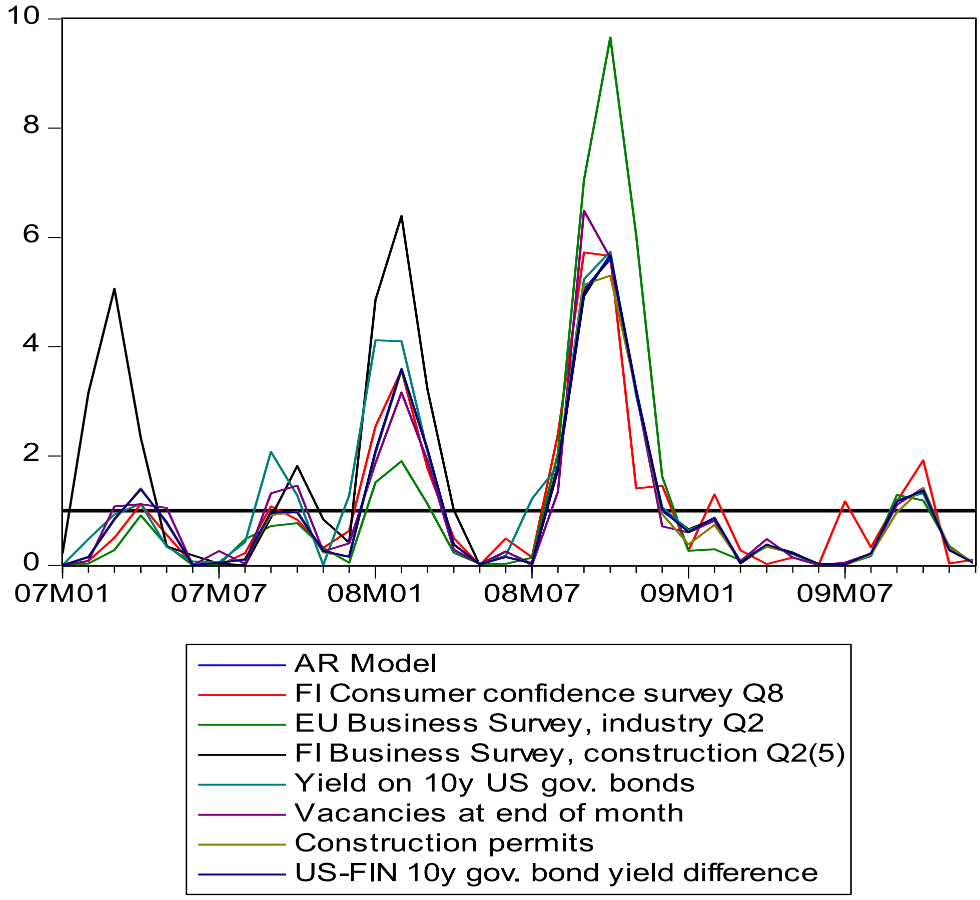

Figure 5: Squared forecast errors 2007 - 2009 produced by forecasts using the individual components of the CLI. The horizontal line indicates the mean of the squared forecast errors of the AR model.

It would indeed seem that a relatively small number of individual large forecast errors contribute to most of the MSFEs. These in fact do seem to coincide with the largest spikes in the growth rate of industrial production in figure 4 and, especially, the substantial drop in production towards the end of 2008. Apparently, the poor forecast performance of the components of the CLI, as measured by the mean of these forecast errors, would thus seem to be, at least to a great extent, due to the inability to forecast the few large spikes in industrial production growth.

The construction permits variable, found to improve slightly upon the AR benchmark in Table 4, is the only one of the component variables which is revised data. Macroeconomic data is typically periodically revised as newer information becomes available. From the point of view of forecasting this poses problems, as the revised data used in the analysis is quite different from the first observations that were available at the time. To verify, whether the CLI or some of its individual components would have been of use had they been used in the past, a forecasting exercise using real time data is performed in the following sub-section.

6.2.2 Forecasts using a real-time data set

All the analysis conducted so far in this thesis has been done using the data that was available in December 2009. An important evaluation criterion of any macroeconomic forecast method, however, is its performance with real-time data. Due to data revisions, the initially published numbers of, e.g. industrial production are updated to render them more accurate. These revisions can often be substantial. From the point of view of forecasting this is problematic because the most recent observations are the least accurate ones. Furthermore, pseudo out-of-sample forecast measures used so far do not exactly tell us how well a model would have performed, had it been used in the past, if the data used to calculate them are the final revised numbers available today. Hence, as another method of evaluation, whenever possible, the latest figures that were available at each time ought to be used in the estimation of the pseudo out-of-sample forecast errors, instead of the most accurate, revised data available currently. Since all actual forecasting is done in real time, strictly speaking, it cannot be claimed that a composite leading indicator would have been of use in the past, if this assessment is based on final revised figures.

In order to assess how useful a predictor the constructed CLI would have been, had it been used in the past, it is thus used to forecast the real-time data set of Finnish industrial production published by the OECD. The data set consists of 124 time series, or vintages of data. Each series represents the first published data available each month from September 1999 to December 2009 respectively. The first observations included in the sample are still from 1988 as before.

For the following analysis, the series on the total volume of granted construction permits is left out of the CLI as it is revised data and no unrevised series were available. Keeping the unrevised series in the CLI model, one would not be able to simulate real time forecasting. The other component variables are unrevised survey, financial or employment data and consequently pose no problems. [4]

Out of sample forecasts over September 1999–December 2009 are calculated using an 11-year rolling window scheme as before, apart from the fact that now the squared forecast errors are calculated as

$$\hat{\varepsilon}_{t+6}^2 = (x_{t+6|\tau+6} - \hat{x}_{t+6|\tau})^2, \tag{8}$$

where τ indicates the data vintage, where the first value for period t is published. That is, the forecast error is the difference between a forecast of the industrial production growth rate at period $t+6$, calculated using the time series on industrial production that were published at period t, and the industrial production growth rate that occurred at period $t+6$ according to the first published numbers of industrial production for that period (the first time series available that include an observation for $t+6$).

For example, the first forecast in this real time forecasting simulation uses the CLI model to forecast industrial production growth, using the time series on industrial production available in September 1999. A six-month-ahead forecast yields a value for March 2000. This is the value one would have obtained had one used the CLI model developed in this thesis for forecasting in September 1999 using the most recent data available then. The forecast error is then computed as the difference between this forecasted value and the first published data on March 2000 industrial production. This method tells us how the CLI model would have performed, had it been used in real time and had its performance been evaluated in real time.

[4] As the time series of the non-financial variables are seasonally adjusted, they are in fact subject to slight revisions. These revisions arising from the seasonal adjustment methods are assumed sufficiently small to not have a substantial effect on the forecasts the variables produce.

The mean of these squared forecast errors is taken and compared to the mean of the squared forecast errors of an AR model, where the errors are calculated in a same manner. The measure for evaluating the forecasts using real-time data is thus the relative MSFE as before but the squared errors are obtained in a different way.

Again, the real time performance of both the full CLI model and its individual components are evaluated. The results of these forecasts using real time data are listed in Table 5.

Table 5: Relative MSFEs of forecasts using the full model and its individual components under real time data.

Forecaster	**Relative MSFE**
Composite leading indicator	2,158
Difference between yields on U.S. and Finnish 10 year government bonds	0,996
EU business survey: order stock in industry	1,253
FIN construction sector survey: % stating material shortage as main hindrance to production	1,653
FIN consumer confidence survey: Ability to make major purchases	1,025
Unfilled vacancies at the end of the month	1,021
Yield on 10 year U.S. government bonds	1,027

The full CLI model again fails to provide any additional information in forecasting. Of the individual components, only the spread between U.S. and Finnish long-term interest rates improves slightly upon the AR-benchmark, although not statistically significantly. A Giacomini and White (2006) test of conditional predictive ability on the forecasts using this component series yields a p-value of 0,5996.

Figure 6 plots the squared forecast errors of the real time forecasts using the components of the CLI. The errors are again scaled to the MSFE of the benchmark AR model.

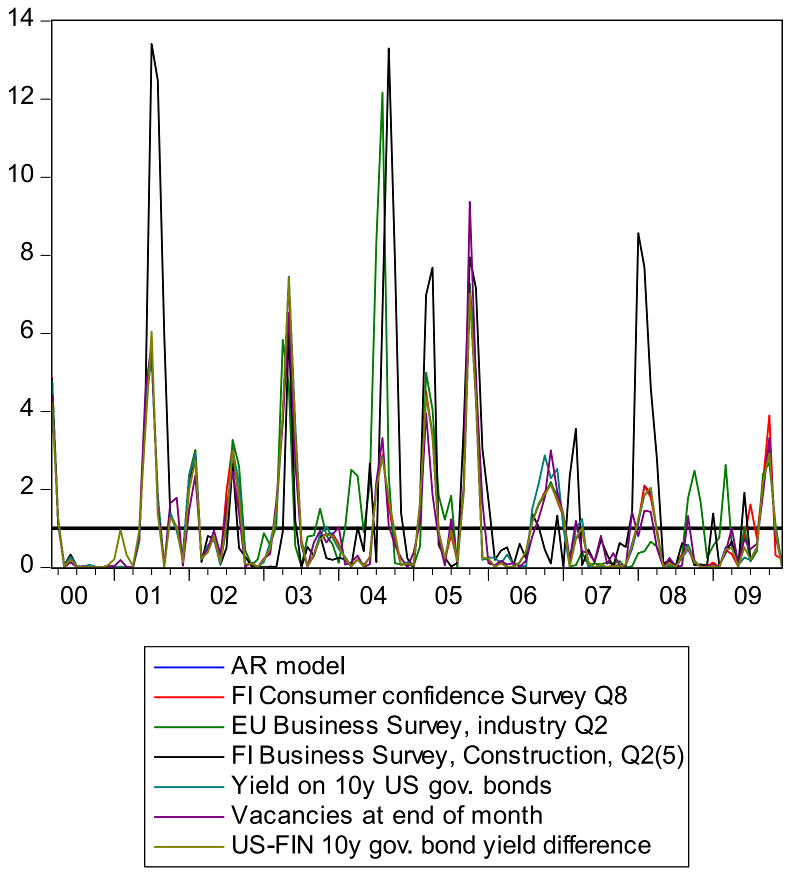

Figure 6: Squared forecast errors of the real time forecasts using the components of the CLI model. The errors are scaled to the mean of the squared errors produced by the AR model, indicated by the horizontal line.

The high MSFE of the forecasts using the component series of the CLI again seems to be due to few very large forecast errors that render the means of the errors large. This manifests as the spikes appearing in the graph in Figure 6. Yet now, the spikes are several and not concentrated to a particular time period or variable. Moreover, in real time forecasting, much larger forecast errors have occurred in the past than the once appearing in the forecasts of the latest recession.

While the composite leading indicator or its component series would not seem to be very helpful, had they been used in the past, the effect individual large forecast errors does appear quite significant. As a similar problem appeared in forecasting the latest recession in the previous subsection, this suggests that perhaps a different loss function for the measurement of the out-of-sample errors instead of the MSFE would be more appropriate, as the mere fact that the errors are squared emphasizes the effect of a single poor forecast.

6.2.3 Forecasting a composite coincident indicator

As a further evaluation of the forecasting ability of the CLI, an attempt to forecast a composite coincident indicator with it is made. As was discussed in Section 2, industrial production is not the best variable for describing the business cycle, since it only reflects a part of the economy. In order to assess how well the CLI performs in forecasting a broader measure of the business cycle, its predictive ability with respect to a composite coincident indicator of the Finnish economy is evaluated.

The CLI model is used to forecast the composite coincident indicator developed by Lanne and Nyberg (2009). The relative MSFE against a benchmark AR model, using again an 11-year rolling estimation window, gets a value of 1,774. The CLI model thus fails in forecasting this alternative measure of the business cycle.

This result may be due to the fact, that the variables chosen mainly lead industrial production and not the other component series of the CCI. Another reason may be the fact that the CCI itself seems to lead industrial production slightly, particularly in the middle of the sample. Yet a third possible explanation is that the CLI created and industrial production both contain a great deal of high-frequency volatility, whereas the composite coincident indicator is a much smoother series. To the extent that this volatility is idiosyncratic, unrelated to the business cycle, further smoothing of the component series, for example by means of frequency domain filters discussed in Section 4, might provide more accurate forecasts in this case.

6.2.4 Stability of the CLI model

Another issue of some concern in this type of study is whether the predictive relationship is stable over time, as several studies referred to in Section 3 found significant instability in the predictive relations between leading indicators and economic activity. Since Finland has experienced a great deal of structural changes between the 22 year time span studied, the matter of stability is of particular importance.

As a simple way of evaluating the stability of the full CLI model, a sup-Wald stability test (see Andrews 1993) is applied to it using the full sample from 1988 to 2009. A 15 % trimming is used. That is, the test looks for a break date in the central 70 % of the sample.

The most probable break point is found in March 2007. Still, the null hypothesis of no break points is not rejected, as the maximum likelihood ratio F-statistic at that date is 6,946 with a p-value of 1,000.

7 Conclusions

On the basis of the previous empirical research, a set of over a hundred candidate variables to forecast Finnish industrial production with was assembled. Using a procedure involving both in-sample and out-of-sample methods, these were formed into a composite leading indicator. The types of variables obtained are in accordance with previous research and existing CLIs; interest rates, short-term employment data, construction permits, consumer survey data and orders in manufacturing industries are all variables appearing in some form in the CLI developed in this thesis as well as most of the CLIs used as examples in the literature review of Section 3.

Two main conclusions arise from the literature surveyed. First, financial variables, particularly interest rate spreads seem to come closest to variables being useful predictors in every economy. This is found in the CLI developed in this thesis as well; the component series include the yield on 10-year U.S. government bonds and the difference between it and a Finnish 10-year government bond. Particularly the latter one seems to be a good predictor of Finnish industrial production growth, producing better forecasts than a benchmark AR model in a real-time forecast simulation. Although this improvement upon the benchmark model is not statistically significant, this result does warrant further research into the predictive abilities of this long-rate spread.

A second conclusion rising to be made from the literature reviewed in this thesis is that, while several leading indicators have been discovered in studies conducted on the U.S. and some European economies, the variables that have predictive ability depend on the sample and economy studied. Furthermore, the results are not robust to the forecast model or forecast horizon used and significant instability has been discovered in the predictive relations.

In a sense, this same pattern is apparent in the results of this thesis as well. While a good forecasting ability for the CLI model was achieved in the sample from 1988 to 2006 and the model passes the parameter stability test applied, the forecasts conducted outside this sample indicate a substantial deterioration in the predictive ability of the CLI. However, the inaptitude of the model or any of its component series in forecasting the steep decline in industrial production that occurred in 2008 seems to contribute greatly to the means of the forecast errors in this sample. Nevertheless, using real-time data also result in poor forecasts of the CLI and its component series.

This non-robustness of the empirical results, found to a greater extent in the literature reviewed, suggests two alternative conclusions: Either there are no reliable predictive relations between leading indicators and economic activity, or the econometric models and procedures applied are simply not equipped to produce accurate forecasts. As economic theory suggests the predictive relations ought to exist, the latter conclusion seems the plausible one. This is also the argument proposed by Stock and Watson (2003) based on their findings of very little consistence in the forecasts using different leading indicators, samples and forecast horizons.

The finding of the forecasting power of leading indicators depending on the forecast horizon is a common one (e.g. Davis and Fagan 1997 and Forni et al. 2001). In this thesis it is of particular relevance as the forecast model used is horizon-specific, forecasting six months ahead. The variable selection procedure was repeated for 12 and 24 month forecast horizons on the same base set of variables. Very similar results were obtained on the forecasting power of the composite leading indicators and these results are omitted from this thesis. A relevant aspect of forecasts on these longer horizons is, however, that the most of the component series of the constructed CLIs were different for all forecast horizons. The finding in previous empirical literature that different variables produce accurate forecasts, depending on the horizon, is thus essentially replicated in these results. Hence, the use of a forecast model that is not horizon-specific would be an interesting topic for future research.

This domain of research is fairly novel in the context of the Finnish economy and further research topics to develop the ideas presented in this thesis are several. An interesting development, left for future research, is the use of an asymmetric loss function instead of the mean squared forecast error used in this study. Such loss functions give their users the opportunity to penalize forecasts that overestimate the actual occurring value, or vice versa. This can be useful in many cases, where, for example, overly optimistic economic forecasts are more damaging.

Another interesting forecast method for further research would be the pooling of individual forecasts. It has long been acknowledged that pooling several forecasts, by e.g. simply taking an arithmetic mean of several forecasts, can improve accuracy. The logic behind it is to use more information than the individual forecasts and consequently produce better forecast results. Pooling, also allows the forecaster to, in a sense, hedge against misspecification or instability in the models (see Timmerman 2006 for a thorough overview on pooling methods in forecasting).

A final suggestion for further studies in this domain would be the creation of a type of automatic composite leading indicator. Such an indicator has been proposed for few euro area countries by Camba-Mendez et al. (1999). The indicator forecasts by automatically reselecting its component series from a large pool of candidate variables at each period.

REFERENCES

Andersson, M. & D'Agostino, A. 2008: Are sectoral stock prices useful for predicting euro area GDP? ECB Working Paper no. 876.

Andrews, Donald W.K. 1993: Test for Parameter Instability and Structural Change with Unknown Change Point. Econometrica 614:821-856.

Arnold, I. J. M. & Lemmen, J. J. G 2008: Inflation Expectations and Inflation Uncertainty in the Eurozone: Evidence from Survey Data. Review of World Economics 144, 2: 325-346.

Artis, M. J., Marcellino, M. & Proietti, T. 2004: Dating Business Cycles: A Methodological Contribution with an Application to the Euro Area. Oxford Bulletin of Economics and Statistics, 66, 4: 537-565.

Bernanke, B. 1990: On the Predictive Power of Interest Rates and Interest Rate Spreads. NBER Working Papers 3486.

Bernanke, B. & Blinder, A. S. 1992: The Federal Funds Rate and the Channels of Monetary Transmission. American Economic Review, 82, 4: 901-921.

Burns, A. F. & Mitchell, W. C. 1946: Measuring business cycles. NBER Studies in Business cycles no. 2, New York.

Camba-Mendez, G., Kapetanios, G., Smith, R., & Weale, M. R. 2001: An Automatic Leading Indicator of Economic Activity: Forecasting GDP growth for European Countries. The Econometrics Journal 4, 1: 56-90.

Campbell, J.Y., Lettau, M., Malkiel, B.G. & Xu, Y. 2001: Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. Journal of Finance 56, 1: 1-43.

Conference Board 2009: The Conference Board Leading Economic Index (LEI) for the United States and Related Composite Economic Indexes for October 2009. The Conference Board.

Dai, Q. & Singleton, K. J. 2002: Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics 63: 415-441.

Davis, E. P. & Fagan, G. 1997: Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries? Journal of Applied Econometrics 12: 701-714.

Elliot, G. Rothenberg, T. J. & Stock, J. H. 1996: Efficient Tests for an Autoregressive Unit Root. Econometrica 64(4): 813-836.

Estrella, A. & Hardouvelis, G. 1991: The Term Structure as a Predictor of Real Economic Activity. The Journal of Finance, 46, 2.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. 2001: Coincident and leading indicators for the Euro area. The Economic Journal 111(May):C62-C85.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. 2003: Do financial variables help forecasting inflation and real activity in the euro area? Journal of Monetary Economics 50: 1243-1255.

Friedman, B. M. & Kuttner, K. M. 1998: Indicator Properties of the Paper-Bill Spread: Lessons from Recent Experience. Review of Economics and Statistics 80: 34-44.

Giacomini, R. & White, H. 2006: Tests of conditional predictive ability. Econometrica, 74, 6: 1545-1578.

Guo, H. 2002: Stock Market Returns, Volatility, and Future Output. Federal Reserve Bank of St. Louis Review 84, 5: 75-85.

Hamilton, J. D. & Kim, D. H. 2002: A Re-Examination of the Predictability of Economic Activity using the Yield Spread. Journal of Money, Credit and Banking 34: 340-360.

Ivanova, D., Lahiri, K. & Seitz, F. 2000: Interest rate spreads as predictors of German inflation and business cycles. International Journal of Forecasting 16: 39-58.

Junttila, J. 2007: Forecasting the macroeconomy with contemporaneous financial market information: Europe and the United States. Review of Financial Economics 16: 149-175.

Kauppi, E., Lassila, J. & Teräsvirta, T. 1996: Short-term forecasting of industrial production with business survey data: Experience from Finland's great depression 1990-1993. The Research Institute of the Finnish Economy Discussion Paper No. 546.

Lanne, M. & Nyberg, H, 2009: Suomen kansantalouden suhdanneindeksi. Kansantaloudellinen aikakauskirja 4/2009: 422-432.

Macroeconomic Advisers 2008: Macroeconomic Advisers' Measure of Monthly GDP. Macroeconomic Advisers' Macro Focus Volume 3, Number 3.

Marcellino, M. 2006: Leading indicators". In Elliott, G., Granger, C. W. J., & Timmerman, A.: Handbook of Economic Forecasting. North-Holland.

Marcellino, M. Stock, J. H. & Watson, M. W. 2003: Macroeconomic forecasting in the Euro area: country specific versus Euro wide information. European Economic Review, 47: 1-18.

Mitchell, W. & Burns, A. F. 1938: Statistical indicators of cyclical revivals. Reprinted in: Moore, G.H. (ed.) 1961: Business cycle indicators. Princeton University Press, Princeton.

Moneta, F. 2003: Does the yield spread predict recessions in the Euro area? ECB Working Paper no. 294.

Mourogane, A. & Roma, M. 2002: Can confidence indicators be useful to predict short term real GDP growth? ECB Working Paper no. 133.

OECD 2009a: Reference Turning Points & Component Series of the Composite Leading Indicators. Retrieved 2.12.2009. www.oecd.org/std/cli

OECD 2009b: Glossary for OECD Composite Leading Indicators. Retrieved 2.12.2009. www.oecd.org/std/cli/glossary

Osborn, D. & Sensier, M. 2002: The Prediction of Business Cycle Phases: Financial Variables and International Linkages. National Institute Economic Review, 182 1: 96-105.

Rosenberg, J. V. & Maurer, S. 2008: Signal or Noise? Implications of the Term Premium for Recession Forecasting. Federal Reserve Bank of New York - Economic Policy Review, Vol. 14, No. 1, July 2008.

Stock, J. H. & Watson, M. W. 1989: New indexes of coincident and leading economic indicators. In: Blanchard, O., & Fischer, S. (eds.): NBER Macroeconomics Annual. MIT Press, Cambridge, MA.

Stock, J. H. & Watson, M. W. 1999: Business cycle fluctuations in US macroeconomic time series. In: Taylor, J. B. & Woodford, M.: Handbook of Macroeconomics, vol. IA. North-Holland, Amsterdam.

Stock, J. H. & Watson, M. W. 2003a [2007 revised edition]: Introduction to econometrics, Pearson International Edition. Pearson Education, Inc., Boston, MA.

Stock, J. H. & Watson, M. W. 2003b: Forecasting Output and Inflation: The Role of Asset Prices. Journal of Economic Literature, vol. 16: 788-829.

Taipalus, K. 2006: Bubbles in the Finnish and US Equities markets. Bank of Finland Studies E: 35 - 2006.

Timmermann 2006: Forecast combinations. In Elliott, G., Granger, C. W. J., & Timmerman, A.: Handbook of Economic Forecasting. North-Holland.

Zarnowitz, V. and Braun, P. 1989: Comment to “New indexes of coincident and leading economic indicators”. In: Blanchard, O., & Fischer, S. (eds.): NBER Macroeconomics Annual. MIT Press, Cambridge, MA.

Data Appendix

Table A1: The time series used and the data source.

VARIABLE	SOURCE
Orders and inventories	
Construction sector - evolution of overall order books (EU, business survey question 4)	European commission
Construction sector - evolution of overall order books (euro area, business survey question 4)	European Commission
Construction sector - evolution of overall order books (Finland, business survey question 4)	European Commission
Evaluation of raw material stock (US business survey)	OECD
Industry - Export order book levels (EU, business survey question 2)	European Commission
Industry - Export order book levels (Euro area, business survey question 2)	European Commission
Industry - Order book levels (EU, business survey question 2)	European Commission
Industry - Order book levels (Euro area, business survey question 2)	European Commission
Industry - Order book levels (Finland, business survey question 2)	European Commission
Industry - Stock of finished products (EU, business survey question 4)	European Commission
Industry - Stock of finished products (Euro area, business survey question 4)	European Commission
Order inflow in industry (US)	OECD
Business surveys (see Table 3 for questions that compose the surveys)	
Construction business survey - Composite index and individual questions (Euro area)	European Commission
Construction business survey - Composite index and individual questions (EU)	European Commission
Construction business survey - Composite index and individual questions (Finland)	European Commission
Euro Area Business Climate Indicator	European Commission
Industry business survey - Composite index and individual questions (euro area)	European Commission
Industry business survey - Composite index and individual questions (euro area)	European Commission

US Business Survey - Industry - Composite index	OECD
US Business Survey - Industry - employment expectations	OECD
US Business Survey - Industry - production expectations	OECD
Consumer surveys (see Table 3 for questions that compose the surveys)	
Consumer Confidence indicator - composite index and individual questions (Euro area)	European Commission
Consumer Confidence Indicator - EU	European Commission
Consumer Confidence Indicator - Finland, Individual questions only	
EU Economic Sentiment Indicator	European Commission
Euro area Economic Sentiment Indicator	European Commission
Finland Economic Sentiment Indicator	European Commission
Price indices	
Consumer price index	Statistics Finland
Consumer price index - Energy	OECD
Consumer price index - Food	OECD
Consumer price index - Housing	OECD
Consumer price index without food or energy	OECD
Export Price index	OECD
Import price index	OECD
Producer price index - domestic	OECD
Producer price index of export	OECD
Total Producer price index	OECD
Total wholesale price index	OECD
Wholesale price index - domestic	OECD
Wholesale price index - imports	OECD

Consumption and sales	
Granted construction permits, housing (amount)	Statistics Finland
Granted construction permits, housing (cubic volume)	Statistics Finland
Granted construction permits, all (amount)	Statistics Finland
Granted construction permits, all (cubic volume)	Statistics Finland
New passenger cars registered	Statistics Finland
Retail sale of automotive fuel, value index	Statistics Finland
Retail sale of automotive fuel, volume index	Statistics Finland
Retail sale of motor vehicles, value index	Statistics Finland
Retail sale of motor vehicles, volume index	Statistics Finland
Retail sale of watches and jewelry, value index	Statistics Finland
Retail sale of watches and jewelry, volume index	Statistics Finland
Wholesale value index	Statistics Finland
Wholesale volume index	Statistics Finland
Financial variables	
Difference between US and Finnish short rates	
Difference between US and Finnish term spreads	
Difference between yields on 10y US and German government bonds	
Difference between yields on German and Finnish 10y government bonds	
Difference between yields on US and Finnish 10y government bonds	
Dow Jones Euro Stoxx 50 Stock Index	ECB
German 3-month interest rate	OECD
Helsinki Stock Exchange all share index	OECD
Term spread (fin)	

Term spread (ger)	
Term spread (us)	
The S&P 500 Stock Price Index	ECB
US 3-month interest rate	OECD
Yield on a 10-year Finnish government bond	OECD
Yield on a 10-year German government bond	OECD
Yield on a 10-year US government bond	OECD
Exchange rates and competitiveness indices	
Euro effective (trade weighted) exchange rate	Statistics Finland
Finnish Euro area augmented real competitiveness indicator	Statistics Finland
Finnish narrow real competitiveness indicator	Statistics Finland
Narrow Euro Real exchange rate	European Commission
Nominal USD-FIM/EUR exchange rate (1990M1-)	Bank of Finland
Employment statistics	
Total jobseekers	Ministry of Employment
Individual lay-offs	Ministry of Employment
Total vacancies	Ministry of Employment
New vacancies	Ministry of Employment
Unfilled vacancies at the end of the month	Ministry of Employment
Unemployment rate	Ministry of Employment
Total unemployed	Ministry of Employment
Unemployed jobseekers	Ministry of Employment

Table A2: The questions of the business and consumer surveys.

Construction Business Survey (EU, Euro area and Finland)		
#	**Question**	**Details**
1	Building activity development over the past 3 months	Only FIN & EA
2	Main factors currently limiting your building activity (below)	Only FIN & EA
2(1)	None (%)	
2(2)	Insufficient demand (%)	
2(3)	Weather conditions (%)	Omitted
2(4)	Shortage of labor force (%)	
2(5)	Shortage of material and/or equipment (%)	
2(6)	Other factors (%)	Omitted
2(7)	Financial constraints (%)	Omitted
3	Evolution of your current overall order books	
4	Employment expectations over the next 3 months	
5	Prices expectations over the next 3 months	
The composite index is calculated as (Q3+Q4)/2		

Industry Business Survey (EU, Euro area and Finland)		
#	**Question**	**Details**
1	Production trend observed in recent months	
2	Assessment of order-book levels	
3	Assessment of export order-book levels	Only EU & EA
4	Assessment of stocks of finished products	Only EU & EA
5	Production expectations for the months ahead	
6	Selling price expectations for the months ahead	
7	Employment expectations for the months ahead	
The composite index is calculated as (Q2-Q4+Q5)/3		

Consumer Confidence Survey (EU, Euro area and Finland)		
#	**Question**	**Details**
1	Financial situation over last 12 months	
2	Financial situation over next 12 months	
3	General economic situation over last 12 months	
4	General economic situation over next 12 months	
5	Price trends over last 12 months	Only EU& EA
6	Price trends over next 12 months	Only EU
7	Unemployment expectations over next 12 months	
8	Major purchases at present	
9	Major purchases over next 12 months	Only EU& EA
10	Savings at present	
11	Savings over next 12 months	Only EU & EA
12	Statement on financial situation of household	Only EU & EA
The composite index is calculated as (Q2+Q4-Q7+Q11)/4		