Macroeconomic Forecasting
using
Business Cycle Leading Indicators

Ard H.J. den Reijer

This book was typeset by the author using L^AT_EX

Cover Design: Maryland

Printed by US-AB Stockholm

ISBN 978-90-9024994-0

Copyright ©2010 Ard den Reijer

All rights reserved. No part of this publication may be recorded or transmitted in any form by any means electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the author.

Macroeconomic Forecasting using Business Cycle Leading Indicators

Macro-economisch voorspellen op basis van voorlopende
conjunctuurindicatoren

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. Mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 26 maart 2010 om 16.00 uur.

door

Adrianus Hendrikus Johannes den Reijer

Maastricht University

Promotoren:

Prof. dr. F.C. Palm

Prof. dr. L.H. Hoogduin (Universiteit van Amsterdam)

Beoordelingscommissie:

Prof. dr. J.-P. Urbain (Voorzitter)

Prof. dr. B. Candelon

Prof. dr. S.J. Koopman (Vrije Universiteit Amsterdam)

Acknowledgements (in Dutch)

Het proefschrift is voltooid! Zoals geldt voor zoveel zaken zou ik dit proefschrift nooit hebben kunnen voltooien zonder de begeleiding, steun en aanmoediging van mijn omgeving. Het uiten van dankbaarheid aan alle vrienden, familie en collega's, die direct of indirect hebben bijgedragen aan dit proefschrift, is een welhaast onmogelijke taak, omdat zovelen in aanmerking komen voor een welgemeende dankbetuiging. Zonder afbreuk te willen doen aan ieders individuele bijdrage, wil ik bij deze toch een aantal mensen expliciet bedanken.

Elk groot project begint met de nodige inspiratie, waarvoor ik André Hoogstrate wil bedanken. Het werd mij snel duidelijk dat ik de academische en de financiële wereld wilde combineren. Dit proefschrift is geschreven in de periode dat ik werkzaam was bij De Nederlandsche Bank. Ik ben mijn promotoren Franz Palm en Lex Hoogduin erkentelijk voor de geboden mogelijkheid. Ik dank beide promotoren voor hun optimisme, het gestelde vertrouwen en de geboden motivatie. De discussies, vele kritische en constructieve commentaren en suggesties hielden mij scherp en bouwden mee aan de brug tussen theorie en praktijk.

Ik wil bovendien de leden van de beoordelingscommissie danken voor het lezen van de conceptversie van de dissertatie en voor hun vriendelijke en bruikbare commentaar. Ik voel me vereerd met de bereidheid van hooggevalueerde experts met een indrukwekkende staat van dienst om zitting te nemen in mijn commissie.

Een speciaal woord van dank gaat uit naar de coauteurs met wie de intensieve samenwerking nuttig en leerzaam was en bij vlogen zelfs bijzonder vreugdevol. Veel dank Peter Vlaar voor de coöperatie en het kamergevoel. Je liet me zien hoe het is om elke situatie en vraagstuk soeverein en onbevreesd in al z'n ins en outs helemaal te doordenken. Een memorabel

moment is jouw verdediging van hoofdstuk 2 in de lobby van het Frankfurter Hof hotel vlak voor aanvang van een conferentiediner.

Veel dank ook aan Jan Jacobs en Pieter Otter met wie ik gezamenlijk als drie musketiers optrok. Pieter, ik vond het bijzonder dat we elkaar leerden kennen door samen hoofdstuk 7 te schrijven voordat we elkaar voor het eerst in levende lijve ontmoetten. Jan, jij hebt mij in verschillende wereldsteden bijna alle hoofdstukken van dit proefschrift zien presenteren. Dank voor je aanwezigheid, je feedback, je inhoudelijke commentaar, je coaching, onze discussies, je directe en indirecte uitnodigingen en dat je de eerste versie van hoofdstuk 7 afrondde slechts enkele uren voordat Ofelia geboren werd. Dat ook Elvira ter wereld kwam voordat deze dissertatie is afgerond lijkt me toch een duidelijk signaal dat jullie heel bijzondere kinderen zijn. Het is aan jullie dat ik dit proefschrift opdraag. Moge het als bron van inspiratie dienen voor jullie om datgene te voltooien dat jullie op jullie pad tegenkomen, gelijk in betekenis zoals opa en oma dat voor mij vormen.

Ik wil mijn ouders en zus bedanken voor hun constante ondersteuning en betrokkenheid gedurende mijn hele educatieve periode, welke symbolisch door mijn vader tot uitdrukking komt in zijn rol als paranimf. Een speciaal dankwoord gaat uit naar Caspar Pors vanwege zijn bereidheid, zelfs voordat de eerste letter op papier stond, op te treden als paranimf en daarmee op mijn *viva voce* zijn brandweeruniform te verruilen voor een rokkostuum.

Till odödliga älskade Cecilia, min ängel, mitt allt, mitt hela själv. Bara några ord idag och det så i en avhandling!

Ard den Reijer, Stockholm, februari 2010

Contents

1	Introduction	1
1.1	Macroeconomic forecasting using business cycle indicators . . .	2
1.1.1	Macroeconomic forecasting of key variables	3
1.2	Econometrics of business cycle indicators	6
1.2.1	Explicit selection: small collection of targeted indicators	6
1.2.2	Implicit selection: weighted average across a large col- lection of available indicators	7
1.3	Outline of the thesis	8
2	Forecasting Inflation: an Art as well as a Science!	11
2.1	Introduction	11
2.2	Model Selection Procedure	13
2.2.1	Preliminary Data Analysis	13
2.2.2	Statistical Criteria	14
2.2.3	Economic Evaluation	16
2.3	The Empirically Optimal Models	17
2.4	Forecast Uncertainty	19
2.5	Model Evaluation	22
2.6	Conclusion	28
2.A	Appendix	30
3	The Dutch Business Cycle: a Finite Sample Approximation of Selec- ted Leading Indicators	35
3.1	Introduction	36
3.2	The business cycle	37
3.2.1	The Hodrick-Prescott filter	39
3.3	Measuring the business cycle	41

3.3.1	The composite reference index	41
3.4	The leading business cycle indicator	43
3.4.1	Turning point prediction	54
3.5	Indicator revision analysis	55
3.6	Indicator based GDP forecasting	58
3.7	Conclusion	60
3.A	Appendix	62
3.A.1	Coincident indicators	62
3.A.2	Leading indicators	63
4	Deviation Cycles in Manufacturing: Business Cycle Measurement and Leading Indicators	67
4.1	Introduction	67
4.2	Measuring business cycles	69
4.3	Detecting and dating Dutch deviation cycles	71
4.4	The empirics of international deviation cycles in manufacturing	75
4.5	Leading indicators	82
4.6	Conclusion	91
4.A	Appendix	93
4.A.1	A technical note on testing for duration dependence . . .	93
5	Forecasting Dutch GDP using Alternative Factor Model Specifications Based on Large and Small Datasets	103
5.1	Introduction	104
5.2	The factor model	105
5.2.1	Factor model representation	105
5.2.2	Estimating the factors	106
5.2.3	Factor forecasting	108
5.3	Forecasting Dutch GDP	110
5.3.1	Real-time forecast simulation design	110
5.3.2	Factor model diagnostics	110
5.3.3	Dutch data	111
5.4	Empirical results	112
5.4.1	Forecast accuracy of various model specifications	114
5.4.2	Forecast accuracy of various data configurations	116
5.5	Conclusions	119
5.A	Appendix	121
5.A.1	The estimator	121
5.A.2	Dutch data set	123

6	Regional and Sectoral Dynamics of the Dutch Staffing Labour Cycle.	129
6.1	Introduction	129
6.2	Staffing agency work	130
6.2.1	Data	131
6.3	Dynamic factor model	132
6.3.1	Aggregate and aggregated staffing employment	134
6.4	The empirics of staffing employment	137
6.5	Forecasting aggregate staffing employment	142
6.5.1	Results	143
6.6	Conclusion	147
6.A	Appendix	148
6.A.1	Data	148
6.A.2	On aggregation	149
6.A.3	The Estimator	150
6.A.4	Empirical results	153
7	Information, data dimension and factor structure	157
7.1	Introduction	157
7.2	Information in data	159
7.2.1	Information measure based on eigenvalues	159
7.2.2	Kullback-Leibler numbers and information	160
7.2.3	Relative information measure Inf_N^R in the approximate factor model	161
7.2.4	A test procedure	162
7.2.5	MSE-prediction	163
7.3	Application	163
7.3.1	The data set	164
7.3.2	Information in the data set	164
7.3.3	Allowing for pure leads and leads and lags	168
7.4	Conclusion	169
7.A	Appendix: The U.S. macroeconomic data set	172
8	Summary and Conclusions	175
8.1	Inflation	175
8.2	Business cycle	176
8.2.1	Business cycle measurement	176
8.2.2	Composite leading business cycle indicator	177
8.3	GDP	178
8.4	Staffing employment	179
8.5	Forecasting the aggregate using disaggregates	180
8.5.1	Inflation	181

8.5.2	GDP	181
8.5.3	Staffing employment	182
8.6	Data dimension and factor structure	182
8.7	Concluding remarks	184
Bibliography		187
Summary in Dutch		197

List of Tables

2.1	The Netherlands: HICP (sub)indices	19
2.2	Euro area: HICP (sub)indices	20
2.3	The Netherlands: Recursive root mean squared forecast error 1998-2002, 1 to 18 months ahead.	24
2.4	Euro area: Recursive root mean squared forecast error 1998- 2002, 1 to 18 months ahead.	25
A.2.5	Data, notation and source code	33
3.1	The coincident business cycle indicators that constitute the ref- erence index	42
3.2	Statistical criteria of potential leading indicators	46
3.3	Turning point dating of GDP, the composite coincident index and leading indicators	48
3.4	Composition of business cycle indicators for the Netherlands by different institutions	52
3.5	The size of adjustment of the indicators from their mid-sample estimates.	58
3.6	Forecasting performance of the leading indicators.	60
4.1	Dutch deviation cycle turning points of industrial production and GDP	76
4.2	Summary statistics for the deviation cycle indicators of the man- ufacturing industry	80
4.3	Statistics of duration dependence of the deviation cycle in the manufacturing industry	81
4.4	Index of concordance and correlation between indicators of de- viation cycles in the manufacturing industry	83

4.5	Results of selection leading indicators	86
4.6	Dating of turning points	89
A.4.7	Data source	99
5.1	Diagnostics for different factor model specifications	113
5.2	Forecast accuracy of different model specifications; one quarter ahead forecast horizon	115
5.3	Forecast accuracy of different model specifications; average over forecast horizons	116
5.4	Forecasting performance of different data sets: one quarter fore- cast horizon	117
5.5	Forecasting performance of different data sets: average over forecast horizons	118
A.5.6	Description of data set	125
6.1	The empirical results of the staffing labour cycle at the regional level.	138
6.1	The empirical results of the staffing labour cycle at the sectoral level.	139
6.3	The relative forecasting performance of different model spe- cifications	144
6.4	The relative forecasting performance of the dynamic factor spe- cification	146
A.6.5	Forecasting performance of different models using recursive win- dows	154
A.6.6	Forecasting performance of different models using rolling win- dows	155
7.1	Ranking of series according to relative information criteria . . .	165
7.2	Ranking of series according to relative information criterion: pure leads, and leads and lags	170
A.7.3	Description of the Stock and Watson data set	173

List of Figures

2.1	Realisation and forecast of HICP inflation for the Netherlands and the euro area	21
2.2	Root mean squared error Dutch HICP inflation 1 to 15 months ahead	27
A.2.3	HICP (sub)indices in original, monthly and annual inflation format for the Netherlands	30
A.2.4	HICP (sub)indices in original, monthly and annual inflation format for the euro area	31
A.2.5	Exogenous variables	32
A.2.6	Endogenous variables	32
3.1	Cyclical motion of GDP and respectively the reference index, industrial production, household consumption and staffing employment	44
3.2	Stance of the business cycle in an historical and a multivariate perspective	55
3.3	Consecutive updates of the coincident and leading indices . . .	57
A.3.4	Cyclical motions of leading indicator variables	64
4.1	Deviation cycles of the Netherlands	74
A.4.2	International deviation cycles in the manufacturing industry and composite leading indicators	96
6.1	Year-on-year growth rate of total employment, staffing employment and turnover	135
6.2	Invoiced staffing hours and its model model decomposition . .	136

7.1	Relative information (dashed line) and relative eigenvalues information (solid line) of ordered data set	166
7.2	Does an additional variable add information?	168
7.3	Comparison of relative information	171

Chapter 1

Introduction

The analysis of interaction between separate entities, individuals, organizations, groups of agents over time, i.e. seasons, years, business cycles, life cycles, intergenerational and space, i.e. cities, agglomerations, countries, currency areas, continents is a key objective of the economics science. For this purpose, huge collections of data are almost instantaneously available at relatively low costs. Data on the National Accounts for virtually all countries, prices for thousands of products, assets and services, purchases by individual consumers, for instance available through scanner data from supermarket and warehouse chains, unemployment data for hundreds of types of jobs, for example turnover data on flexible staffing labour, tick-by-tick observed transaction data of various financial markets, environmental data on CO₂-emissions and climate change, on health care, education, etcetera.

Big-Data refers to a data-rich environment that is characterized by the explosion in quantity and quality of potential relevant data, mainly due to progress in recording and storage technology. In the information age, statistical samples are no longer measured by the number of observations, but in, say, gigabytes. In the near future, even more data will become available also at higher sampling frequencies, months, weeks, days, minutes, while traded share prices are nowadays only seconds away from being publicly available in real-time.

The information age posed central banks with the challenge of executing monetary policy in a data-rich environment (cf. Bernanke and Boivin, 2003). An accurate assessment of the current and near future macroeconomic outlook is crucial information in the decision process of policy makers and businesses. The sheer fact that central banks take the effort of collecting and analyzing a large variety of data series at least suggests that policymakers and applied forecasters are keen on extracting information from many series describing economic activity at a more disaggregate level. Since monetary policy can be a function of the forecasts, as with inflation targeting (cf. Svensson, 1997), the choice of the forecasting model is important. In the practice of central banking, macroeconomic data analysis is characterized by employing dynamic stochastic general equilibrium models, macroeconomic structural mod-

els, small-scale statistical models and heuristic, judgemental weighing of anecdotal evidence stemming from varying and diverse sources. In the modern day data-rich environment, this thesis explores whether, and if so how, Big Data can fruitfully be employed in measuring and forecasting macroeconomic key variables of interest.

1.1 Macroeconomic forecasting using business cycle indicators

The Maastricht Treaty formulates a clear mandate for the monetary authority of the Eurozone being the collective of European Union member states that have adopted the euro as their sole currency. The Eurosystem consists of the European Central Bank and the National Central Banks of the member states of the euro area. The mandate is to maintain price stability and is given a quantitative content by the inflation rate, which is the rise in the general level of prices of goods and services in an economy over a period of time. The goal is an inflation rate for the euro area as a whole that should be close to, but below 2% in the medium term. In order to determine the nature and extent of risks to price stability, a two-pillar monetary policy strategy is executed by thoroughly analyzing economic and, respectively, monetary developments and organising, evaluating and cross-checking the information according to these two complementary analytical perspectives. The economic analysis assesses the short to medium-term determinants of price developments, with a focus on real activity and financial conditions in the economy, that are largely determined by the interplay of supply and demand in goods, services and factor markets. The monetary analysis exploits the long-run link between money and prices. The monetary analysis mainly serves as a means of cross-checking the indications for monetary policy coming from the economic analysis.

So, under the second pillar of economic analysis, forecasting of macroeconomic key variables at a short to medium term horizon is important for both monetary authorities and private agents who try to understand and react to the central bank's behavior. An accurate assessment of the current and future state of the cyclical position of the economy yields valuable information for short to medium term forecasting. The idea is to exploit business cycle leading indicators that either cause cyclical fluctuations or quickly react to positive or negative shocks. As cyclical downturns, i.e. recessions, originate from different sources and possess different characteristics, some indicators possess significant forecasting power during some time periods for some countries,

but it is in general infeasible to identify a single indicator that shows consistently good forecasting performance for all countries and time periods (cf. Stock and Watson, 2003). In order to pick up signals originating from different spheres in the economy, a composition of business cycle leading indicators is assembled that covers the behavior of various macro actors in the economy: households, firms, monetary and fiscal authorities and the foreign sector.

1.1.1 Macroeconomic forecasting of key variables

The key objective of this thesis is to develop and compare short- and medium term forecasting models based on small and large collections of leading indicator variables aimed at, mainly Dutch, key macroeconomic variables such as inflation, gross domestic product (GDP), staffing employment and the business cycle.

Inflation

When the general price level rises, each unit of the functional currency buys fewer goods and services. Consequently, inflation is a decline in the real value of the monetary unit of account in an economy and so a loss of purchasing power of the internal medium of exchange. The inflation rate is the percentage change in a general price index being the Harmonised Index of Consumer Prices (HICP) for the European Union. The purpose of a price index is to measure the prices of a large basket of representative goods and services and render an overall indexed price as the weighted average of all the prices of the items in the basket. The HICP uses data collected by surveying households to determine a basket of weights as the proportion of a typical consumer's spending on specific goods and services. Index prices are typically expressed in relation to a base year, which gets assigned a value of 100.

The HICP is generally divided into 5 subcomponents labelled unprocessed food, processed food, non-energy industrial products, energy and services. A policy maker has an interest in forecasts of the HICP subindices to construct a measure of core inflation defined as total HICP excluding the components unprocessed food and energy. These two volatile components are generally more affected by short run demand and supply conditions in specific markets and thereby less susceptible to monetary policy.

The leading indicator variables that potentially trigger inflationary developments can be classified into: 1) Demand-pull inflation, inflation caused by excess demand and favorable market conditions due to increased public and private consumption and stimulated investment; 2) Cost-push inflation, supply shock inflation caused by increased producer input prices, like for ex-

ample oil, metals and commodities prices; 3) Built-in inflation, induced by expectations of economic agents about future inflation developments that potentially become self-fulfilling, often linked to the price/wage-spiral.

This thesis develops forecasting models to predict Dutch and euro area HICP inflation and their five subcomponents with a forecast horizon of eighteen months ahead. Similar papers describing small scale linear models aimed at forecasting HICP inflation concerning different countries, currency areas and methodologies include Banerjee et al. (2005), Benalal et al. (2004), Bruneau et al. (2007), Bruneau et al. (2003), Fritzer et al. (2002), Hubrich (2005) and Moser et al. (2007).

Business cycle

Business cycles can broadly be defined as oscillating motions of economic activity, which are visible as patterns of fluctuations of macroeconomic variables such as output, production, interest rates, unemployment and prices. A business cycle consists of a peak in economic activity, a period of recession followed by a trough and a period of expansion. Consecutive business cycles are separated by turning points, that is peaks and troughs. The *classical* cycle considers the fluctuations of the level of economic activity, see Harding and Pagan (2002), while the *deviation* cycle considers the fluctuations around some trend. Deviation cycles gained popularity, because classical recessions (periods of negative growth rates) have been exceptional in industrialized countries since the Second World War. So, deviation cycles could more naturally be related to the fluctuations observed in the level of employment and unemployment. Moreover, the concept of a deviation cycle as the gap between actual output and trend output gained policy relevance through the stronger focus on Taylor-rule driven monetary policy and cyclically adjusted government balances.

In establishing the chronology of turning points, the American and European business cycle dating committees typically rely on GDP, industrial production, employment, consumption related to wholesale and retail trade and real income. This thesis aims to construct a deviation cycle indicator for the Dutch economy that consists of a reference index, which represents the current stage of the business cycle, and the indicator, which represents the developments of the cycle in the near future. Related literature documenting the construction of Dutch business cycle indicators are Kranendonk et al. (2005) related to the Netherlands Bureau for Economic Policy Analysis (CPB), van Ruth et al. (2005) related to Statistics Netherlands (CBS), Jacobs (1998) and Fase and Bikker (1985) and Berk and Bikker (1995) related to De Nederlandsche Bank.

Gross domestic product

GDP is the aggregate of all economic activity and consists of the market value of all final goods and services produced within the borders of a national economy in a year. GDP is the basic measure that fulfills the macroeconomic accounting identity as it represents at the aggregate level production, expenditure and income. So, in the National Accounts framework, GDP can be compiled in three ways: 1) expenditure approach, the sum of final consumption, capital formation and exports minus imports; 2) production approach, the sum of value added, i.e. output minus intermediate consumption, at all the intermediate stages of production by all the industries; 3) income approach, the sum of remuneration of employees, gross operating surplus of enterprises, i.e. profits and taxes less subsidies on production and imports.

Even though GDP constitutes the aggregate activity, the cyclical fluctuation of its underlying components do not always move synchronously and, in case of government expenditures, can even move anti-cyclically. This thesis aims at exploiting the leading indicator characteristics of its underlying disaggregates to forecast Dutch GDP growth rates for a forecast horizon up to 4 quarters ahead. Similar papers employing a large data set of disaggregates to forecast GDP related to different countries, currency areas and methodologies include Altissimo et al. (2001), Artis et al. (2002), Banerjee et al. (2005), Grenouilleau (2004), Marcellino et al. (2003), Rünstler et al. (2009), Schneider and Spitzer (2004), Schumacher and Dreger (2004), Schumacher (2007) and Stock and Watson (2002b).

Staffing employment

Flexible staffing agency work is characterized by a triangular relationship between the user firm, the employee and the private labour market intermediary. The staffing agency is a private matchmaker that acts as an intermediary between temporary labour supply and demand. From the perspective of the client firm, flexible staffing labour constitutes a mere variable factor of production. During the last 30 years, temporary employment expanded rapidly prior to macroeconomic upturns, while sharp declines in temporary employment preceded recessions. Hence, fluctuations in staffing employment are timely indicators of broader business cycle motions. Goldschmeding (2003) and Franses and de Groot (2005b) analyse the Dutch staffing labour market developments to monitor and forecast macroeconomic business cycles.

Based on a large data set of observations that are directly obtained from the administrative source of the leading market participant in the Netherlands, i.e. Randstad, this thesis aims to document the cyclical developments of staffing

employment at the disaggregate level and to identify the regions and sectors that show leading properties. The second question is then how the disaggregate information, particularly the identified leading indicators at the sectoral and geographical level, can be exploited to forecast the country aggregate of staffing employment.

1.2 Econometrics of business cycle indicators

The availability of sets of data for a large number of individual entities and the importance of understanding individual behavior and the interaction between individuals have created a demand for the refinement of the existing and the development of a new toolkit of econometric techniques.

1.2.1 Explicit selection: small collection of targeted indicators

A basic modelling strategy is to augment an autoregressive process for the target variable, i.e. GDP, inflation, employment with a single leading indicator as explanatory variable. Allowing for a limited amount of autoregressive terms, this low-order equation strategy generates many single indicator based forecasts for a specific target variable. Faced with a large collection of potentially relevant leading indicator variables, Bates and Granger (1969) propose to combine the different forecasts. Moreover, Palm and Zellner (1992) relate to the relative merit between combination and selection to obtain optimal forecasts. The selection or the combination of the model specifications employing one of the available indicator variables can be based on statistical criteria related to model-fit and forecasting performance. Calculating entities as marginal likelihoods of different model specifications and corresponding predictive distributions are straightforward exercises when considering a Bayesian framework, especially for large data samples given the recent developments in computational optimization procedures.

A related strategy is to improve the estimation of low-order equations by employing panel data techniques and so collect a longitudinal data set that records similar characteristics of different, and by assumption independent, entities, or panels. Based on a multi-country data set consisting of leading indicator variables, Garcia-Ferrer et al. (1987) apply pooling techniques to establish a relationship between annual output growth and indicators such as real stock returns and real money growth (see also Hoogstrate et al., 2000).

An often adopted framework to analyze the mutual interaction between a limited number of time series variables is the vector autoregressive model (VAR). Allowing all variables to be endogenously determined, a VAR can only

be implemented for small systems typically consisting by two up till six variables. The approach of medium scaled multivariate VAR models is most convenient in case out of a large set of candidate variables, only not too many aggregate key variables exhibit significant relationships. In case of many candidate variables each possessing only a marginal significance, then the ratio of estimated parameters to observations and so, the sampling error in the estimated coefficients will be large. In case there are more variables than observations over time, a VAR becomes even infeasible, i.e. the curse of dimensionality. However, including even just a moderate amount of lags already makes a medium scaled VAR system prone to overparametrization. One way of reducing dimensionality is to identify and model common characteristics of the endogenous time series variables. In the terminology of Engle and Kozicki (1993), a feature is a characteristic of a time series such as serial correlation, trend, seasonality, volatility, etc. Common features arise when variables have such features in common in the sense that there exists linear combinations of the time series that fail to have the feature even though each of the time series individually exhibits the specific feature (cf. Vahid and Engle, 1993; Hecq et al., 2002). For non-stationary series exhibiting a trending pattern, a common structure could result from the presence of cointegration, that is from the occurrence of common stochastic trends. In that case, two or more time series variables follow a common trend process, which can be exploited to reduce the dimensionality of the multivariate model specification. The general modelling strategy of imposing consistency between the univariate characteristics and the multivariate model structure is called SEMTSA, i.e. Structural Economic Modeling Time Series Analysis (see part I in Zellner and Palm, 2004).

In this thesis, the outlined considerations will be most directly applied to the inflation forecasting models. The models consist of VARs including exogenous variables whose precise specification is dependent on the patterns of the data regarding seasonality and the common-trend relationship, especially between wages and prices. Moreover, an automated model selection procedure based on goodness-of-fit, parsimony and forecast accuracy is developed and implemented.

1.2.2 Implicit selection: weighted average across a large collection of available indicators

Macroeconomists naturally think of the comovement in economic time series as arising largely from relatively few unobserved structural shocks like productivity, monetary policy, etc. The underlying notion that economic motions are captured by a few driving aggregate forces implies that the information

contained in each available economic key variable at an aggregate level is less informative about macroeconomic behavior than the information contained in the available variables at a disaggregate level. A diffusion index captures the common motion that is most widespread across the large collection of available indicators, thereby representing the unobserved common shock, that is the common factor. Stock and Watson (2002a; 2002b) propose the principal component to estimate the unobserved common factors, which is generalized by Forni et al. (2000; 2001; 2001a; 2004; 2005) to allow for dynamic relationships between the common factors and the individual time series. Modelling the dynamics explicitly is in principle desirable, since macroeconomic variables in general are non-synchronized and leading indicators should be able to play a crucial role in a forecasting context. More recently, the factor model approach is further generalized by Eichler et al. (2009) to allow for non-stationarity and so, accommodates smooth transitions over time in the covariance structure of observed multivariate time series. Other recent research designs panel unit root tests using a common factor structure to model the cross-sectional dependence, (cf. Bai and Ng, 2004) and, inter alia, Gengebach et al. (2006) aims to detect non-stationarity in the data that is driven by a reduced number of common stochastic trends. This thesis applies the factor model approach to forecasting Dutch GDP based on the stationarized, i.e. trend adjusted, collection of leading business cycle indicators consisting of the data underlying the central bank's macroeconomic structural model. In order to empirically determine the importance of the size and the structure of the data set, we generate forecasts for different configurations of the data set. We determine the factor model diagnostics of each specification and data configuration and aim to establish a relationship with the forecasting performance. Boivin and Ng (2006) show that enlarging a big data set not necessarily improves the factor forecasting performance if the additional series are noisy or unrelated to the target variable. The final section of this thesis formalizes the optimal size and composition of a subset of the large collection of available leading indicators with the aim of applying factor analysis related to a target variable. Finally, this thesis applies the factor model to analyse the staffing labour cycle in the Netherlands based on a disaggregate data set that describes the number of hours of staffing employment for 15 different regions and 58 different sectors. The common factor at the country level is extracted and compared to the common cycles at the disaggregate level, which exclude the effects of sector or region specific shocks.

1.3 Outline of the thesis

The aim of the first chapter is to develop forecasting models to predict Dutch and euro area inflation. All models are linear vector autoregressive or error correction models, possibly including exogenous variables. First, we describe a procedure to select an optimal forecasting model. The selected models for the Netherlands and the euro area are reported and employed to generate inflation forecasts and uncertainty bounds surrounding the forecasts. Finally, the forecast results of the models are evaluated by comparing the recursive root mean squared forecast errors to those of random walk models and optimal autoregressive models.

The aim of the second chapter is to construct a business cycle indicator for the Dutch economy following the approach of using leading and coincident indicators that has been developed at the National Bureau of Economic Research (NBER) in the US in the 1930s (cf. Burns and Mitchell, 1946). So, first the concept of a business cycle is defined as waves with lengths longer than 3 years and shorter than 11 years and methods to distinct cycles and trends are discussed. Then we determine the set of broad measures of macroeconomic activity from which we extract the motion of the business cycle. We screen the cyclicity of financial, survey and real activity variables on their replication and leading properties with respect to the business cycle and construct a composite leading indicator. Moreover, we analyse the reliability of the real time cyclical estimates of both composite indices and quantify their convergence as more data becomes available and show the pseudo real-time forecasting performance of the most recent turning point. Finally, the composite leading indicator is employed to generate forecasts for GDP growth rates.

The aim of the third chapter is to measure and analyse the business cycles represented in the manufacturing industry of nine countries. For each country, we provide descriptive statistics like the dating of cyclical turning points, low- and high-growth periods and summary statistics describing features like amplitude, steepness and duration of the cycle. Moreover, we will test for duration dependence of low- and high-growth periods. Duration dependence means that the probability that the high- or low-growth phase ends next period increases with the duration of the ongoing phase. In addition, the international linkages of the cyclical motions in the manufacturing industries among countries can be measured by the fraction of time the cycles are simultaneously in an upturn and/or in a downturn.

The fourth chapter applies the factor model to forecast Dutch GDP up to four quarters ahead. The underlying aggregate forces are modelled as unobserved common factors, which can be estimated by principal components. We

introduce the specifications of different factor models and show the parameter configurations for which the cyclical dynamic factor collapses into a general dynamic factor and a static factor. In order to empirically determine the importance of the size and the structure of the data set, we generate forecasts for different configurations of the data set. We determine the factor model diagnostics of each specification and data configuration and aim to establish a relationship with the forecasting performance.

The same specifications of the factor model are applied in the fifth chapter related to staffing employment. We first describe the data set that is directly obtained from the administrative source of Randstad and is nearly instantaneously available. Then, the factor model is employed to extract the staffing labour cycle from the data at the aggregate country level and to identify leading indicators at the disaggregate level of sectors and regions. Finally, the performance is compared between different factor model specifications that exploit the information at the disaggregate level to forecast the staffing labour developments at the country level.

While the fourth and the fifth chapter show in an empirical analysis to what extent employing disaggregate information contributes to forecasting an aggregate variable, the sixth chapter aims to provide a formal framework. This chapter exploits concepts from information theory, in particular Kullback-Leibler criteria, to analyse information in the data. We propose two relative information measures, one based on eigenvalues, the other based on Gaussian distributed data. Ordering the series of the data set according to these measures enables us to identify a subset of the data set that is most informative to modelling a variable of interest. A test procedure is derived to determine whether an additional variable is sufficiently correlated assuming an approximate factor structure in the data.

Finally, the last chapter summarizes the thesis by presenting the forecasting models for the different target variables. The extent to which disaggregate variables can be employed to forecast an aggregate target variable is related to the optimal size and composition as the subset of the large collection of available leading indicator variables.

Chapter 2

Forecasting Inflation: an Art as well as a Science!

Macroeconomics and reality.

Christopher A. Sims

Abstract

In this study we build two forecasting models to predict inflation (Harmonised Index of Consumer Prices, HICP) for the Netherlands and for the euro area. The models provide point forecasts and prediction intervals for both the components of the HICP and the aggregated HICP-index itself. Both models are small-scale linear time series models allowing for long run equilibrium relationships between HICP components and other variables, notably the hourly wage rate and the import or producer prices. The model for the Netherlands is used to generate the Dutch inflation projections for the eurosystem’s Narrow Inflation Projection Exercise (NIPE). The recursive forecast errors for several forecast horizons are evaluated for all models, and are found to outperform a naive forecast and optimal AR models. Moreover, the same result holds for the Dutch NIPE projections, which have been provided quarterly since 1999. The aggregation method to predict total HICP inflation generally outperforms the direct method, except for long horizons in the case of the Netherlands.

2.1 Introduction

The mandate of the European Central Bank (ECB) is to maintain price stability in the euro area. This goal is given a quantitative content by requiring that the year on year growth of the Harmonised Index of Consumer Prices (HICP) for the euro area as a whole should be close to, but below 2% in the medium

We thank an anonymous referee and participants of the second Conference of the Euro Area Business Cycle Network, the 10th International Conference on Computing in Economics and Finance and seminar participants at De Nederlandsche Bank, Erasmus University Rotterdam and Maastricht University, in particular Filippo Altissimo, Bob Chirinko, Bertrand Candelon, Denise Osborn and Adrian Pagan for useful comments.

term. The ECB is monitoring and forecasting price developments under the first pillar of its monetary policy strategy¹. Therefore, forecasting inflation rates has become important for both monetary authorities and private agents who try to understand and react to the central bank's behaviour. The aim of this paper is to describe the procedures used at De Nederlandsche Bank to predict Dutch HICP inflation and a new model to directly predict overall euro area inflation. The models forecast both the components of the HICP, as requested by the euro system's narrow inflation projection exercise (NIPE), and, for comparison reasons, the total HICP itself. The forecast horizon is eleven to eighteen months ahead. Similar recent papers describing small scale linear models aimed at forecasting inflation include Banerjee et al. (2005), Benalal et al. (2004), Bruneau et al. (2007), Bruneau et al. (2003), Fritzer et al. (2002), Hubrich (2005) and Moser et al. (2007).

One of the issues addressed in this paper is that of disaggregation, which can be regarded in three dimensions, that is across components (sub-indices of an index), across time (higher frequency) and across space (different regions of an economic area). In the European context there is the aggregation of the forecasts of individual countries to a euro area level, see for instance Espasa et al. (2002), Marcellino et al. (2003) and Benalal et al. (2004). In this study we will only address the aggregation of HICP component forecasts². Obviously, there is a clear interest in finding out whether aggregating component forecasts performs better than forecasting the aggregate directly. Aggregating forecasts of component models is potentially beneficial as forecast errors might cancel between components. Moreover, the disaggregate components can be better modelled by choosing a more suitable model for each component separately and by possibly incorporating additional explanatory variables. This argument is indeed apparent in theoretical models, see Lütkepohl (1987) and Hendry and Mizon (2000). The latter authors assume for instance a known and constant data generation process. However, Hubrich (2005) and Benalal et al. (2004) find empirical evidence for euro area data and across various specifications that directly forecasting the aggregate HICP performs better than aggregating the forecasted components, especially for a forecast horizon up to 12 months ahead. Apart from the possible efficiency gain, a policy maker also has an interest in forecasting the HICP components to construct a measure for

¹In practice, each country provides four times a year its own inflation forecast for an horizon of 11-15 months and these forecasts are used to construct an area wide forecast. This periodic procedure is called the Narrow Inflation Projection Exercise (NIPE).

²In contrast, the eurosystem's NIPE creates a euro area forecast by aggregating the individual countries' inflation forecasts. So, the model for the euro area built in this study generates forecasts for the area wide aggregates, while the NIPE aggregates the individual country's forecasts to the euro area level.

core inflation, defined as HICP excluding the components unprocessed food and energy. These two components are generally considered more volatile and less susceptible to monetary policy.

Another important issue in this paper is the model selection procedure, for which a new heuristic method is developed. The method involves three steps. The first step involves the visual inspection of the data, primarily to detect changing seasonal patterns. This step determines the general structure of the small scale model. The second step involves calculating all possible models given this structure, allowing for a small set of exogenous and endogenous variables and variable lag lengths. Optimal statistical models are subsequently selected according to goodness-of-fit, parsimony and/or out-of-sample forecasting accuracy. The final selection is based on the economic evaluation of the statistically selected models. Especially, the long run properties are important in this respect.

This paper is organized as follows. In section 2.2 we describe our procedure to select an optimal forecasting model. The selected models for the Netherlands and the euro area are described in section 2.3. Section 2.4 elaborates on the uncertainty surrounding the forecasts. In section 2.5 the forecast results of the models are evaluated. First, the recursive root mean squared forecast errors are compared to those of random walk models and optimal autoregressive models, both for the component models and the direct HICP models. Then, the Dutch NIPE results are evaluated. Finally, section 6 concludes.

2.2 Model Selection Procedure

The model selection procedure consists of three steps. The first step involves preliminary data analyses. This step is used to select the optimal model structure. That is to say, either a vector autoregressive (VAR) model in first differences or a vector error correction model (VECM) in both first and twelve month differences. The second step involves the computation of all possible models given the possible set of exogenous and endogenous variables and the allowable lag length. The optimal models according to several statistical criteria are subsequently shown. The last step implies the economic evaluation of the statistically selected models in order to select the final optimal model. This involves both an economic interpretation of the coefficients, primarily with respect to the error correction term, and an analysis of the model properties to provide stable long run forecasts.

2.2.1 Preliminary Data Analysis

The main purpose of this first step is to detect time varying seasonal patterns as the HICP data are not seasonally adjusted. They are plotted in the appendix, figure A.2.5 and figure A.2.4 for respectively the Netherlands and the euro area. All series in both cases are plotted in raw format, in annual inflation rates, that is 12 month differences of the log-transformed HICP and in the monthly change of the log price index, that is monthly inflation. The figures show that especially the (log) sub-indices for non-energy industrial goods (P^i) and services (P^s) as well as total HICP (P^{total}) are subject to a changing seasonal pattern. Particularly for the euro area but also for the Netherlands, this change is not concentrated in just one month. Otherwise, a second set of seasonal dummies should be sufficient to remove the seasonality. Instead, the pattern is filtered out of the data by taking 12 month differences. Moreover first differences are taken to eliminate the (near) unit root in inflation. However, lagged annual inflation is also included to prevent overdifferencing. In this way, stationarity of annual inflation, or cointegration of inflation with other variables is allowed for. The same model is used for processed food (P^{pf}) for the Netherlands, for which the seasonal pattern is less clear. Concerning the series unprocessed food (P^{uf}), energy (P^e), and P^{pf} for the euro zone, seasonal dummies are included in the model to capture the seasonal effects. As these variables are clearly non-stationary, they are modelled in first differences. Cointegration is not allowed for these models in first differences as it appears that the economic rationale for a long run relationship among price levels is less obvious than among inflation rates. Moreover, even if such a relationship would exist, it is prone to structural breaks due to for instance indirect tax adjustments.

2.2.2 Statistical Criteria

Given the model structure dictated by the seasonal pattern, all possible models given the set of explanatory variables are computed in an automated selection procedure³. That is to say, a model is computed for every possible combination of explanatory variables and every possible lag structure from zero up to a maximum (usually 12)⁴. Moreover, the 12th lag is analysed separately as this lag is theoretically important for monthly data. For instance, if inflation

³The forecasts of natural gas prices (part of P^e) and housing rents (part of P^s) for the Netherlands are generated outside the model as they are for the time being only adjusted twice respectively once a year, according to some strict rules.

⁴However, some limits are imposed on the total number of variables and/or parameters in the model in order to preserve enough degrees of freedom.

is surprisingly low in a certain month due to an earlier start of summer sales, it can be expected that inflation will be relatively high the next year in the same month (unless the change in sales pattern is permanent). For models in first and twelfth differences, a negative AR(12) coefficient can also partly remedy the permanent base effects of, for instance, indirect tax changes. The same lag structure is used for all the variables included. Moreover, for the VECMs, every possible combination of lagged twelve month differences including inflation itself is added to check for long run relationships (and to correct for possibly incorrectly imposed unit roots in inflation)⁵.

Regarding the explanatory variables, both endogenous and exogenous explanatory variables motivated by economic theory and data availability are potentially included. Some exogenous variables are provided by the ECB for the NIPE-exercise practice, like the future paths for policy variables and the variables regarding the external environment of the EMU⁶. Besides these variables nominal wages are also included exogenously since overlapping contracts and other institutional features make them relatively easy to predict in the short run. Potential endogenous variables include producer prices, import prices, the nominal money stock, industrial production, credit data, retail turnover and business cycle indicators.

The primary selection criteria for the optimal models are modified versions of the classical Information Criteria existing in the VAR literature: Schwarz (SC), Hannan-Quinn (HQ) and Akaike (AIC). These standard criteria are primarily used to determine the optimal lag length of a given VAR system. They need to be modified for our goal as not only the lag length needs to be decided, but also the choice of additional (endogenous or exogenous) variables. Different models can hardly be evaluated on the basis of covariance matrices if models with different endogenous variables are compared, as the residuals of different equations are computed. Therefore, we calculate the information criteria based only on the number of parameters and the residuals of the inflation equation alone. A good fit for the inflation equation alone is not enough however to guarantee a reasonable forecast. If the model includes other endogenous variables that can hardly be forecasted themselves, the inflation forecast will be hampered as well. Consequently, we also apply an alternative measure of fit widely used within the forecasting literature, namely the root

⁵The selection of the long run relationship is based on the same criteria as the variable selection, the lag length and the inclusion of the autoregressive term at lag 12. No formal cointegration rank tests are performed. As the left hand side variable (the monthly change in inflation) is clearly stationary, the twelve month differences will only enter if they are indeed cointegrated or stationary themselves.

⁶With respect to interest rates and exchange rates a no-change path is implemented, whereas futures are used to project oil and other commodity prices.

mean squared forecast error. Too much focus on the out-of-sample performance on the other hand would favour exogenous variables too much as they are included assuming perfect foresight⁷. Moreover, given our relatively short sample, it seems hardly efficient to ignore the in-sample fit⁸.

The modified information criteria are therefore based on the weighted average of the in-sample variance of the inflation equation and the out-of-sample forecast error variance; SC^{mixed} , HQ^{mixed} , AIC^{mixed} . Besides these mixed criteria, the in-sample criteria (SC^{in} , HQ^{in} and AIC^{in}) are also compared as well as the root mean squared (forecast) errors both in-sample ($RMSFE^{in}$)⁹, out-of-sample ($RMSFE^{out}$) and combined ($RMSFE^{mixed}$). These nine different criteria often provide nine different optimal models. In principal, the model selected by AIC^{mixed} is chosen. The relatively low penalty for extra parameters for this criterion is justified as the risk of overfitting is mitigated by the inclusion of the out-of-sample forecast variance. The other models might give important indications with respect to the preferred specification as well, however. The fact that the model selection choice is not robust with respect to the selection criterion puts some doubts on the existence of *the* optimal model. This is further confirmed by a periodic evaluation of the results. One more year of data often leads to different selected models.

2.2.3 Economic Evaluation

Given that different criteria prefer different models and the fact that these choices are not very robust with respect to the addition of more data, it is obvious that statistical criteria alone are hardly sufficient to select the optimal model. Judgmental issues, based on economic criteria are important as well. Here, three issues come to mind. First, does the choice of variables make sense? Second, are the parameter values in the model of the right sign and order of magnitude? Third, does the model include a stable anchor for long run forecasts?

With respect to the undesirable variable selection, the example for services inflation in the Netherlands is typical. Both the within sample and the mixed criteria selected the oil price as an important variable for services inflation. In

⁷Including those variables endogenously instead is not an option as either the forecast is assumed to be conditional on the exogenous variables (provided by the ECB), or because they are included exogenously precisely because they are better forecasted using institutional knowledge than with a statistical model (wages).

⁸Moreover, a purely out-of-sample selection method would select overparameterised models with positive probability (Inoue and Kilian, 2005).

⁹The in-sample variance is hereby computed with a correction for the number of estimated parameters in order to get an unbiased estimate. Otherwise extra variables can only improve the result.

the estimated models this was reflected in a very significant negative contemporaneous coefficient. As there is no economic rationale for such a negative impact, the oil price was not included in our preferred model. The significance was probably due to an incidental correlation of outliers in the past. Indeed, according to model selection criteria based on the current data set, the oil price would no longer be selected by the mixed criteria, but the within sample criteria would still select it. Another example is the short term interest rate. This variable showed up in the estimated models with a significant positive contemporaneous coefficient. Due to the widely acknowledged lag in monetary transmission, the central bank's interest rate actions to fight inflation apparently created a positive short term relation between nominal and real interest rates and inflation. In addition to checking the correct sign of the parameters, most attention regarding the order of magnitude of the coefficients is concentrated on the error correction term. In the automatic model selection procedure, all variables in the error correction term are already checked for their sign. Apart from that, implausible long run elasticities might be remedied by slightly adjusting the model (for instance changing the lag length). Probably, the most important economic criterion in the evaluation of models, is the presence of a stable conditional anchor for long run forecasts. Here, the difference between endogenous and exogenous variables is essential. Including endogenous variables that are themselves hard to predict might lead to the drifting of inflation forecasts to unlikely regions, especially if this endogenous variable is included in the error correction term. Therefore for instance, the variable M3 is not allowed to appear in the cointegration relationship. This endogenous explanatory variable is very difficult to forecast over a longer horizon in this VECM setting, and a bad forecast would imply a severe bias in the long run forecast of inflation. Christoffersen and Diebold (1998) show that error correction terms among endogenous variables alone do not help to produce better long run forecasts as they have expectation zero in the long run. Exogenous variables in an error correction term on the other hand do positively affect long run forecasting as they steer the long run outcome for the endogenous variables. Therefore, the wage development as an exogenous explanatory variable is imposed in the error correction term of the selected model if validated by the data. Wage development is well exogenously predictable due to sluggishness in the wage formation process and can act as an anchor of the model. Another anchor is formed by import prices although they are endogenous. However, the import prices themselves are well predictable from (exogenous) oil price and exchange rate developments.

2.3 The Empirically Optimal Models

We applied the selection criteria on a sample running from 1987(10) and 1990(1) until 2002(8), respectively, for the Netherlands and the euro area. The in-sample errors are calculated from the model based on the sample up until 2000(12) and the forecasting errors are obtained from the sample 2001(1) onwards. The model selection is based on the number of 20 forecasts, which are generated using the realized values for the exogenous variables. The forecast errors of the exogenous variables is therefore excluded in the selection process, since the aim of the models is to produce inflation forecasts conditional on the exogenous explanatory variables¹⁰. As stated before, apart from the components P^{uf} , P^e as well as P^{pf} for the euro area, all models are specified in changes of 12 month differences. So, for most models this differencing implies a loss of 13 observations and a remaining sample size of $T=166$ and $T=139$ for respectively the Netherlands and the euro area. The sample for fitting the model is much larger than the sample for obtaining the out-of-sample forecast residuals. Although these forecast errors are less numerous, they get a weight of 0.4 for the mixed criteria, so as to emphasise the importance of good forecasting performance.

We find in accordance with the literature, Stock and Watson (2003), that the forecasting track record of specific models and leading indicators is not invariant over time. Checking the robustness of the model specification by evaluating out-of-sample forecasts for different time periods is further complicated by small sample availability. Moreover, the different criteria produce different optimal models. Therefore, the model selection procedure is rerun regularly.

The selected models for the 5 components and the HICP-index are presented in table 2.1 for the Netherlands and in table 2.2 for the euro area. The optimal model for unprocessed food turns out to be a univariate random walk for the Netherlands and an AR(1) process for the euro area (both including seasonal dummies). Energy prices depend mainly on oil prices, and in the euro area also on producer prices. The most dominating explanatory variable for the other sub-indices is the wage rate, which is imposed and statistically validated in all cointegration relationships for both the Netherlands and the euro area. Wages tend to be more important for the Netherlands than for the euro area as revealed by the twice as high Dutch long term coefficients. Besides wages, a relatively dominating leading indicator for the Netherlands is the import price index of Germany showing up in the cointegration relation-

¹⁰For this reason, the Eurosystem's BMPE uses the word 'projection' to indicate that the forecast is actually conditional on exogenous assumptions.

ship for all four indices¹¹. For the euro area on the other hand, the producer price index is taking this role. These endogenous variables are themselves primarily driven by the oil price, the Euro/Dollar exchange rate and the commodity prices excluding energy. Finally in the euro area, unprocessed food inflation appears as explanatory variable in the index for processed food. This index of processed food turns out to be important for services, which can be explained by restaurant prices. A cointegrating relationship between the processed food and services prices is also found by Espasa et al. (2002). The (un)processed food prices as explanatory endogenous variables for the models of processed food and services respectively are forecast according to the corresponding model specifications, even though the optimal models for the food prices seem to be different. Using different models for food prices might reduce positive correlation among forecast errors of these three components of HICP.

For both areas, the small number of lags selected for all models is noticeable. In previous specifications lag lengths of up to 12 were included, but it seems that the few significant coefficients with longer lags are not very stable. Over the latest sample, especially the selection criteria with relatively strong penalty for extra parameters suggested at most one lag. For services and the total index, the 12th lag is significant as well.

2.4 Forecast Uncertainty

The constructed models provide conditional forecasts for the inflation rates in the short to medium term. We will supplement the point forecasts generated by the models with prediction intervals that provide a quantitative content for the uncertainty surrounding them. The Bank of England quantifies uncertainty by publishing¹² density forecasts, which is an estimate of the complete probability distribution of the possible future values of a variable, see also Wallis (1999). In this study we will use non-parametric bootstrapping to construct a probability distribution and deduce the corresponding prediction intervals, see Horowitz (2001). We perform a simulation experiment in which the error terms are drawn from the distribution of the residuals of the estimated models. The simulation draws from the multivariate empirical distribution to preserve the contemporaneous interdependence of the residuals of the five models. If the bootstrap procedure would be computed for all five categories separately, the overall HICP confidence band would become nar-

¹¹The German import price index is used as no monthly Dutch import price index is available.

¹²The Bank of England has published a density forecast of inflation in its quarterly *Inflation Report* since February 1996.

Table 2.1: The Netherlands: HICP (sub)indices

HICP-index	P^{uf}	P^{pf}	P^i	P^e	P^s	P^{total}
Exogenous	-	$wages^{NL},$ $e^{e\$}$	$wages^{NL},$ P^{oil}	P^{oil}	$wages^{NL},$ $e^{e\$}$	$wages^{NL},$ $e^{e\$},$ P^{oil}
Endogenous	-	Pim^{GE}	Pim^{GE}	-	Pim^{GE}	Pim^{GE}
EC term	-	P^{pf}	P^i	-	P^s	P^{total}
		$wages^{NL}$ Pim^{GE}	$wages^{NL}$ Pim^{GE}		$wages^{NL},$ Pim^{GE}	$wages^{NL}$ Pim^{GE}
Lags included	0	1	0	0	1, 12	12
Specification	Δ_1	$\Delta_1 \Delta_{12}$	$\Delta_1 \Delta_{12}$	Δ_1	$\Delta_1 \Delta_{12}$	$\Delta_1 \Delta_{12}$

Notes:

Δ_x is defined as the x month difference of the variable. The error-correction (EC-)term is specified in annual inflation rates. The models in first differences include seasonal dummies.

Table 2.2: Euro area: HICP (sub)indices

HICP-index	P^{uf}	P^{pf}	P^i	P^e	P^s	P^{total}
Exogenous	-	$wages^{NL},$ $e^{e\$}$	$wages^{NL},$ P^{oil}	P^{oil}	$wages^{NL},$ $e^{e\$}$	$wages^{NL},$ $e^{e\$},$ P^{oil}
Endogenous	-	Pim^{GE}	Pim^{GE}	-	Pim^{GE}	Pim^{GE}
EC term	-	P^{pf}	P^i	-	P^s	P^{total}
		$wages^{NL}$ Pim^{GE}	$wages^{NL}$ Pim^{GE}		$wages^{NL},$ Pim^{GE}	$wages^{NL}$ Pim^{GE}
Lags included	0	1	0	0	1, 12	12
Specification	Δ_1	$\Delta_1 \Delta_{12}$	$\Delta_1 \Delta_{12}$	Δ_1	$\Delta_1 \Delta_{12}$	$\Delta_1 \Delta_{12}$

Notes:

Δ_x is defined as the x month difference of the variable. The error-correction (EC-)term is specified in annual inflation rates. The models in first differences include seasonal dummies.

rower due to the neglected positive correlation among components¹³. Wages, the exchange rate and the oil and commodity prices are exogenous explanatory variables in our models by assumption. The uncertainty surrounding these exogenous predictions is not taken into account. The bootstrapped dis-

Figure 2.1: Realisation and forecast of HICP inflation for the Netherlands and the euro area

¹³Moreover, the confidence band is widened further since in the bootstrap procedure adds an additional residual to the prediction in order to reflect the future uncertainty of unforeseeable shocks.

tribution of the inflation forecasts turns out to be fairly symmetric, see Figure 2.1. The graph shows the forecasts for the Netherlands and the euro area over 2002 and the first half of 2003, based on the models given in Table 2.1 respectively 2.2 and the realisations for the exogenous variables. For both areas, two forecasts and corresponding confidence bands are given: one based on aggregation of the models for the components, P^{agg} ¹⁴, and one based on the model for total HICP inflation, P^{total} . Both model forecasts are in the centre of the confidence band. The bootstrap median and mode are almost identical to its mean implying an almost symmetrical bootstrap distribution. In the past, for the Netherlands we sometimes found that the point forecast was not in the middle of the confidence band. The difference between the two was probably due to a bias in the AR-coefficients. The bias correction methodology of Kilian (1998), which implies using the bootstrap twice, might reduce the difference between the two under those circumstances. Although the models are estimated with data up until 2001, this exercise is not fully out-of-sample in the sense that the data for 2002 was previously included to select the optimal models. Despite these facts, the realised inflation in January 2002 for the euro area was clearly above the prediction interval for the total HICP model and just on the border for the aggregate model. The most likely reason for this underestimation of inflation seems to be the cash changeover to the euro. This event is not explicitly incorporated in these forecasts. An attractive feature of both models for both areas is that the forecast accuracy does not deteriorate over the forecast horizon. Probably, the inclusion of the error correction term including exogenous variables guarantees reasonable long-term forecasts. Of course, this is only true as far as we are able to predict the exogenous variables with reasonable accuracy. With respect to the difference between the aggregated and the total HICP model results, for both areas the aggregated results seem better. First of all, they follow the realised inflation more closely. Moreover, the confidence bands for the total HICP models are wider, especially for longer forecast horizons. However, this forecast evaluation is only based on one period. In the next section, we will evaluate the forecast performance of all models more systematically.

2.5 Model Evaluation

In order to systematically evaluate the models, we have computed the root mean squared error of recursive dynamic out-of-sample forecasts (cf. Stock

¹⁴The weights in the aggregation are updated yearly. Over the forecast horizon, they are assumed constant.

and Watson, 1999), for the model specifications given in Tables 2.1 and 2.2. For this purpose, the realisations of the exogenous variables are used. For the Netherlands this includes gas prices, housing rents and radio and television (RTV) licences. The latter is taken as exogenous as the abolition of them in January 2000 had a huge negative impact on services inflation in that year (see figure A.2.5). If no account is taken of this event, the fit of the models deteriorate. Our first out-of-sample evaluation is for the period 1998:1 up until 1999:6 based on data up until 1997:12, whereas the last exercise involves 2003:1 up until 2004:6, leading to 61 recursive forecast errors for each horizon.

Tables 2.3 and 2.4 show the RMSFE for the Netherlands respectively the euro area. The forecast errors are evaluated for the year-on-year percentage change in the respective HICP component. The forecast errors of the estimated models are compared to those of a naive forecast, which sets all the forecasts ahead equal to the latest observed annual inflation rate, and optimal univariate AR models. The AR models are in first differences, with seasonal dummies, and the lag length is chosen based on the Schwarz criterion for the full sample. With respect to the total HICP, both the results of the own model for total HICP and the one based on an aggregate of components are reported.

For both areas, we find that the models outperform the naive forecast almost uniformly. For the Netherlands, this is always the case, whereas for the euro area the up to three periods ahead model forecast for P^i and the up to twelve periods ahead model forecasts for P^{pf} are about equally good as the naive ones. Although, outperforming the naive forecast seems hardly demanding, the results for the optimal AR models for P^i and P^s for the euro area and the ones for P^{total} and P^{agg} show it is far from trivial. Relative to both benchmarks, the models perform very good, although the models for P^{pf} , and P^i are slightly outperformed by the AR models for short horizons.

Table 2.3: The Netherlands: Recursive root mean squared forecast error 1998-2002, 1 to 18 months ahead.

	p_{uf}		p_{pf}		p_i		p_e		p_s		p_{total}		p_{agg}							
h	<u>naive</u>	<u>AR(0)mod</u>	<u>naive</u>	<u>AR(1)mod</u>	<u>naive</u>	<u>AR(12)mod</u>	<u>naive</u>	<u>AR(0)mod</u>	<u>NaiveAR(0)mod</u>	<u>naive</u>	<u>AR(0)mod</u>	<u>AR</u>	<u>mod</u>							
1	1.93	1.48	1.48	0.42	0.30	0.39	0.44	0.40	0.44	1.92	1.62	0.92	0.38	0.35	0.23	0.36	0.37	0.31	0.35	0.28
2	2.75	2.19	2.19	0.68	0.49	0.62	0.61	0.55	0.59	2.75	2.35	1.29	0.56	0.52	0.33	0.51	0.56	0.41	0.53	0.39
3	3.37	2.71	2.71	0.89	0.66	0.76	0.71	0.64	0.69	3.27	2.91	1.49	0.68	0.67	0.40	0.62	0.71	0.50	0.67	0.45
4	3.88	3.12	3.12	1.07	0.84	0.87	0.84	0.74	0.80	3.87	3.35	1.70	0.80	0.78	0.46	0.76	0.85	0.60	0.81	0.55
5	4.27	3.47	3.47	1.23	0.99	0.96	1.01	0.87	0.90	4.70	3.92	1.99	0.90	0.89	0.50	0.90	0.98	0.67	0.93	0.63
6	4.71	3.77	3.77	1.38	1.15	1.07	1.18	0.99	0.99	5.29	4.36	2.09	0.99	0.97	0.51	1.00	1.09	0.72	1.04	0.69
7	5.10	4.02	4.02	1.53	1.31	1.16	1.32	1.13	1.05	5.88	4.77	2.12	1.10	1.04	0.50	1.10	1.19	0.75	1.12	0.73
8	5.47	4.21	4.21	1.66	1.45	1.20	1.46	1.27	1.09	6.48	5.16	2.25	1.20	1.10	0.50	1.21	1.29	0.76	1.21	0.77
9	5.90	4.49	4.49	1.80	1.58	1.23	1.57	1.37	1.13	7.01	5.58	2.40	1.30	1.16	0.49	1.32	1.39	0.76	1.30	0.81
10	6.36	4.73	4.73	1.92	1.71	1.26	1.68	1.47	1.16	7.36	5.84	2.39	1.41	1.23	0.49	1.42	1.48	0.74	1.37	0.85
11	6.86	4.91	4.91	2.06	1.85	1.29	1.80	1.57	1.19	7.77	6.12	2.40	1.51	1.31	0.49	1.52	1.57	0.73	1.45	0.89
12	7.31	5.14	5.14	2.20	1.99	1.32	1.93	1.69	1.21	8.22	6.43	2.46	1.60	1.39	0.52	1.63	1.68	0.72	1.55	0.95
13	7.56	5.11	5.11	2.31	2.05	1.31	2.03	1.74	1.23	8.50	6.39	2.43	1.66	1.39	0.56	1.71	1.69	0.71	1.57	0.98
14	7.68	5.07	5.07	2.43	2.11	1.31	2.11	1.78	1.23	8.77	6.40	2.40	1.73	1.40	0.63	1.79	1.70	0.74	1.60	1.01
15	7.79	5.05	5.05	2.53	2.13	1.30	2.21	1.83	1.24	9.09	6.40	2.37	1.79	1.40	0.69	1.87	1.72	0.77	1.62	1.05
16	7.83	5.05	5.05	2.62	2.14	1.30	2.30	1.86	1.25	9.39	6.42	2.35	1.84	1.41	0.75	1.94	1.73	0.82	1.64	1.09
17	7.81	5.05	5.05	2.71	2.16	1.29	2.37	1.90	1.26	9.57	6.45	2.35	1.89	1.41	0.80	2.00	1.73	0.88	1.66	1.12
18	7.81	5.07	5.07	2.81	2.18	1.28	2.44	1.93	1.25	9.73	6.47	2.35	1.92	1.41	0.87	2.06	1.74	0.96	1.67	1.15

Notes:

The forecast errors for each horizon h are computed over the annual inflation rates. The models for services and total HICP are corrected for the abolition of RTV licences. The lag length p of the AR(p) models is based on the Schwarz criterion using the full sample. The lowest RMSE for each index is printed in bold face, the highest one in *italics*.

Table 2.4: Euro area: Recursive root mean squared forecast error 1998-2002, 1 to 18 months ahead.

h	p^{pf}			p^i			p^e			p^s			p^{total}			p^{agg}		
	naive	AR(1)mod	naive	AR(3)mod	naive	AR(12)mod	naive	AR(12)mod	naive	AR(0)mod	naive	AR(12)mod	naive	AR(12)mod	naive	AR	mod	mod
1	0.80	0.60	0.16	0.13	0.14	0.21	0.24	0.21	1.87	1.40	0.87	0.19	0.22	0.17	0.21	0.20	0.18	0.15
2	1.36	0.99	0.25	0.21	0.23	0.33	0.36	0.33	2.79	2.09	1.07	0.21	0.26	0.19	0.32	0.29	0.27	0.23
3	1.78	1.30	0.36	0.28	0.30	0.38	0.40	0.39	3.56	2.65	1.32	0.26	0.31	0.22	0.39	0.32	0.33	0.27
4	2.14	1.57	0.46	0.37	0.38	0.40	0.42	0.40	4.39	3.30	1.59	0.29	0.33	0.24	0.43	0.33	0.36	0.41
5	2.42	1.80	0.54	0.44	0.46	0.40	0.42	0.39	5.05	3.73	1.79	0.34	0.38	0.28	0.43	0.33	0.38	0.44
6	2.61	1.97	0.62	0.51	0.54	0.40	0.42	0.38	5.68	4.13	2.00	0.39	0.45	0.31	0.43	0.35	0.40	0.34
7	2.81	2.14	0.69	0.59	0.62	0.45	0.51	0.41	6.31	4.56	2.23	0.45	0.51	0.34	0.44	0.40	0.42	0.51
8	3.04	2.29	0.76	0.68	0.69	0.51	0.61	0.46	6.89	4.90	2.41	0.49	0.56	0.36	0.45	0.45	0.45	0.41
9	3.33	2.49	0.82	0.76	0.77	0.55	0.65	0.48	7.48	5.23	2.61	0.54	0.62	0.39	0.48	0.50	0.49	0.45
10	3.65	2.71	0.89	0.85	0.85	0.57	0.67	0.48	8.07	5.59	2.80	0.58	0.66	0.42	0.52	0.56	0.55	0.50
11	3.97	2.91	0.95	0.94	0.94	0.60	0.67	0.47	8.61	5.93	2.92	0.62	0.70	0.46	0.57	0.61	0.60	0.54
12	4.25	3.10	1.02	1.02	1.03	0.62	0.65	0.46	9.20	6.33	3.06	0.67	0.74	0.49	0.63	0.68	0.65	0.58
13	4.43	3.13	1.07	1.07	1.06	0.64	0.68	0.46	9.61	6.29	3.11	0.70	0.78	0.54	0.65	0.71	0.69	0.59
14	4.57	3.13	1.10	1.10	1.08	0.68	0.72	0.46	9.90	6.25	3.12	0.74	0.83	0.58	0.67	0.73	0.73	0.59
15	4.66	3.14	1.14	1.11	1.08	0.70	0.73	0.45	10.21	6.23	3.11	0.77	0.86	0.61	0.68	0.76	0.77	0.60
16	4.71	3.14	1.18	1.14	1.11	0.70	0.74	0.42	10.42	6.23	3.11	0.81	0.91	0.65	0.69	0.79	0.81	0.60
17	4.74	3.14	1.21	1.16	1.13	0.70	0.74	0.39	10.57	6.22	3.09	0.83	0.95	0.67	0.71	0.83	0.86	0.61
18	4.78	3.14	1.24	1.17	1.15	0.71	0.75	0.37	10.72	6.23	3.05	0.86	0.99	0.71	0.73	0.88	0.91	0.61

Notes:

The forecast errors for each horizon h are computed over the annual inflation rates. The lag length p of the AR(p) models is based on the Schwarz criterion using the full sample. The lowest RMSE for each index is printed in bold face, the highest one in *italics*.

Comparing the naive forecast errors for the Netherlands with those for the euro area, it is clear that Dutch inflation is much more volatile than inflation in the total euro zone. Many of the shocks to inflation are country-specific and these shocks partly cancel for the euro area. The exception is energy inflation. For energy, oil prices are most important and these shocks hit all countries at the same time. For the models, the difference in RMSFE between the two areas is less extreme. Consequently, the improvement relative to the naive forecast is bigger for the models for the Netherlands than for those of the euro area. For forecasts 7 or more months ahead, the Dutch model RMSFE for energy is even smaller than the corresponding euro area one. This is probably due to the assumption of exogenous natural gas prices.

With respect to the advantage of splitting up the HICP index to forecast total HICP inflation, the results are somewhat mixed. For the Netherlands, we find that aggregation of components leads to a lower RMSFE for forecasts up to 7 months ahead, whereas for longer forecast horizons the opposite holds. For the euro area, the aggregation method performs best for all forecast horizons. These results imply that the dominance of the indirect approach for 2002 and the first half of 2003, which appeared clear from Figure 2.1, seems to be a general feature. The relative good performance of the aggregation method for all forecast horizons runs counter to the results of Fritzer et al. (2002). For VAR models, they found the direct approach to perform better for horizons up to 9 months ahead, after which the aggregation approach was to be preferred. Hubrich (2005) and Benalal et al. (2004) on the other hand found that aggregation performed especially worse at long horizons. Also with respect to the AR models, no common feature is found. Whereas for the Netherlands the disaggregated approach produces better results, the opposite holds for the euro area. In general, it seems that forecast errors among HICP sub-indices are too positively correlated to be able to gain a lot by aggregating component models.

The relatively good forecast performance of our models does of course depend on our ability to predict exogenous variables correctly. In Tables 2.3 and 2.4, the realisations are used to make forecasts, but obviously these are not available when making really out-of-sample forecasts. For the Netherlands, we do have a way to check the relevance of this objection as the Dutch models have been used for the NIPE since December 1998. Consequently, we have totally out-of-sample forecasts for HICP inflation and its five components for 16 forecasting rounds.

In Figure 2.2 the root mean squared forecast errors of the NIPE projections are shown together with the ones for the naive forecast, the optimal AR models and those generated with the currently selected models using realisations

Figure 2.2: Root mean squared error Dutch HICP inflation 1 to 15 months ahead

for the exogenous variables. The squared forecast errors are averaged over all projections that were made for a certain forecast horizon, that is 16 for 1 to 11 months ahead, 12 for 12 months ahead, 8 for 13 and 14 months ahead and 3 for 15 months ahead. This explains the sudden drop in RMSFE at horizon 15.

For every HICP component, the NIPE projections outperformed the naive and AR benchmarks almost uniformly. Compared to the model cum realised exogenous variables forecasts the results are more mixed. In principle, there are three reasons for differences between the model forecasts and the NIPE results. First, the assumptions regarding the exogenous variables differ. Second, different models were used. Third, the NIPE also includes ‘add-factors’ to account for judgmental issues. Unfortunately, we do not have a

complete track record of the models and assumptions used. Otherwise, we could identify the exact relevance of each of the three factors. Nevertheless, some interesting conclusions can be drawn. Of the three factors, wrong assumptions probably only lead to worse predictions. For the model specification it can go either way, whereas judgement hopefully only improves the results.

The only two sub-indices for which the NIPE performed systematically worse than the model are energy and services. For energy, this is not at all surprising as the development of oil prices is highly unpredictable and very influential on P^e . Moreover, gas prices, which account to almost 40% of Dutch energy budget, also sometimes moved more than expected. With respect to services, the result is partly due to the abolition of RTV licences in January 2000, which was not foreseen in the NIPE projections of 1999. The figure also shows the impact of ignoring this event for our selected model. It is clear that the model falls apart, as the NIPE now outperforms the model. Besides this effect, the housing rents, which accounted for about 30% of the Dutch services budget, were not always perfectly predictable.

Given the relatively big forecast errors for energy and services, it is surprising to see that for the overall index P^{agg} the NIPE performs even slightly better than the model forecast, even though both are based on the sub-indices. Apparently, the correlation among HICP components was higher for models with correct exogenous variables than for the NIPE. Consequently, the disadvantage of not knowing the future values of exogenous variables was compensated enough by the possibility to add judgement.

2.6 Conclusion

This paper describes the procedures we use to predict monthly Dutch and euro area HICP inflation. The HICP prediction is constructed by aggregating forecasts for the five HICP sub-indices unprocessed food, processed food, non-energy industrial goods, energy and services, whereas total HICP is also modelled directly for comparison reasons. All models are linear vector autoregressive or error correction models, possibly including exogenous variables.

In order to select the appropriate models, the first step is a visual inspection of the data. Those price indices which show a clear changing seasonal pattern are modelled in both first and twelve month differences including an error correction term representing long run equilibrium relationships between inflation and other variables (if they have the correct sign and reasonable order of magnitude). Price indices without clear structural breaks in seasonal pattern (unprocessed food and energy) are modelled in first differences. Here,

no error correction term is included. The second step involves the calculation of all potential models, using a small set of exogenous and endogenous variables. We select the best models according to nine different statistical selection criteria, using both in-sample goodness-of-fit, parsimony and out-of-sample forecasting accuracy. In the third step, the optimal models are chosen, based both on the statistical criteria and economic evaluation. Especially, the long run properties are important here. Expected wage developments form a very important anchor in this respect.

Once an appropriate model is chosen, all available data are used to generate forecasts. Foreseeable shocks over the forecast horizon (for instance a change in indirect taxes) are incorporated ex-post. According to a recursive root mean squared forecast error evaluation exercise, all models outperform the naive forecast and optimal AR models on most forecast horizons. The comparison between the errors of forecasting the aggregate directly or aggregating the forecasts of the components shows a clear preference for aggregating in the euro area. For the Netherlands for short forecast horizons aggregating is better, but for longer horizons the direct approach is to be preferred. These evaluations do depend on perfect knowledge of future values for exogenous variables however.

For the Netherlands, a fully out-of-sample exercise is performed by evaluating the first 16 NIPE rounds. The forecast performance of the NIPE projections is even slightly better than the one for the selected models with perfect foresight of exogenous variables. Again, the naive forecast is outperformed on every forecast horizon for every (sub-)index. Apparently, judgement more than compensates for the lack of knowledge on the future values of exogenous variables. Indeed, forecasting inflation seems to be an art as well as a science!

Overall, the robustness of the inflation forecasting models, both with respect to the selection criterion used and over time, is not encouraging though. The optimal model is not likely to exist, making regular evaluation of models and the permanent good use of common sense all the more important.

Figure A.2.3: HICP (sub)indices in original, monthly and annual inflation format for the Netherlands

2.A Appendix

The sample period of the data set is October 1987 respectively January 1990 for the Netherlands and the euro area until August 2002. Table A.2.5 lists all the variables that are currently included. Apart from these selected variables, other variables have been tested but are not selected in the final models.

Figure A.2.4: HICP (sub)indices in original, monthly and annual inflation format for the euro area

Figure A.2.5: Exogenous variables

Figure A.2.6: Endogenous variables

Table A.2.5: Data, notation and source code

Variable	Notation	External source
<i>Harmonised Index of Consumption Prices</i>		
<u>Euro area</u>		
HICP	p_{ea}^{total}	Eurostat
HICP unprocessed food	p_{ea}^{uf}	Eurostat
HICP processed food	p_{ea}^{pf}	Eurostat
HICP industrial production	p_{ea}^i	Eurostat
HICP energy	p_{ea}^e	Eurostat
HICP services	p_{ea}^s	Eurostat
<u>the Netherlands</u>		
HICP	p_{nl}^{total}	Eurostat
HICP unprocessed food	p_{nl}^{uf}	Eurostat
HICP processed food	p_{nl}^{pf}	Eurostat
HICP industrial production	p_{nl}^i	Eurostat
HICP energy	p_{nl}^e	Eurostat
HICP services	p_{nl}^s	Eurostat
<i>Endogenous variables</i>		
Import price index Germany	p_{im}^{ge}	Statistics Germany
Producer prices (euro area)	p^{prod}	BIS
<i>Exogenous variables</i>		
€/ \$-exchange rate	$e^e/\$$	ECB
Oil price (Brent crude) in euro	p^{oil}	Bloomberg ^{a)}
Hourly wages industry, euro area	$wages^{eu}$	^{b)}
Hourly wages private sector, the Netherlands	$wages^{nl}$	CBS
Commodity prices (excl. energy) in euro	wmp^{exe}	HWWA ^{a)}

Notes:

^{a)} Recent data as well as projections for the forecast horizon are obtained from the ECB. The projections are based on futures prices. ^{b)} The euro area hourly wage is an average of the individual country's hourly wage rates, weighted by the GDP-share in 1995.

©OECD. A.H.J. den Reijer – “*The Dutch Business Cycle: a Finite Sample Approximation of Selected Leading Indicators*,” *Journal of Business Cycle Measurement and Analysis*, forthcoming.
Reprinted with kind permission from the Organisation for Economic Co-operation and Development. Monthly updates of the business cycle indicator are being published on the website of the central bank, <http://www.dnb.nl>.

Chapter 3

The Dutch Business Cycle: a Finite Sample Approximation of Selected Leading Indicators

Measurement without theory.

Tjalling C. Koopmans

Abstract

In this study, we construct a business cycle indicator for the Netherlands. We employ Christiano and Fitzgerald's (2003) approximate band-pass filter to isolate the cycle using the definition of business cycle frequencies as waves consisting of lengths longer than 3 years and shorter than 11 years. The coincident business cycle index is based on industrial production, household consumption and staffing employment. These three variables represent key macroeconomic developments, which are also analysed by both the CEPR and NBER dating committees. The composite leading index consists of eleven indicators representing different sectors in the economy: three financial series, four business and consumer surveys and four real activity variables, of which two supply- and two demand-related. The pseudo real-time performance of the composite indicator is illustrated at the most recent turning point. Moreover, the revisions of the composite indicator as more data becomes available are quantified. Finally, the composite leading indicator is employed in a bivariate Vector Autoregressive model to forecast GDP growth rates.

Comments and suggestions by Robert-Paul Berben, Lex Hoogduin, Jan Jacobs, Franz Palm, Ad Stokman and seminar participants of the 4th *Studiedag Conjunctuur* at Nyenrode University, the 28th CIRET conference on "Cyclical Indicators and Economic Policy Decisions" in Rome, the International Workshop on Computational and Financial Econometrics (CFE 2007) in Geneva, the 27th Annual International Symposium on Forecasting titled "Financial Forecasting in a Global Economy" and seminar participants at The Conference Board in New York City, in particular the late Victor Zarnowitz, are gratefully acknowledged.

3.1 Introduction

Various commercial, academic and government institutions use a business cycle indicator as an instrument to measure and predict business cycle developments and turning points. An accurate assessment of the current and future state of the business cycle is valuable information in the decision process of policy makers and businesses. Most institutions who regularly publish business cycle indicators, follow the approach of using leading and coincident indicators that are developed at the National Bureau of Economic Research (NBER) in the US in the 1930s. Within this dominant methodology, the indicators of the various institutions differ in the specific choices regarding variable selection, the identification of the cyclical patterns, determining the leading/lagging properties of the variables and the weighing of the variables into a single index. Handbooks as user guides for the construction of leading indicators are published by The Conference Board (TCB McGuckin, 2001), the Economic Cycle Research Institute (ECRI, Achuthan and Banerji, 2004) and the Organisation for Economic Co-operation and Development (OECD, Nardo et al., 2008). The institutions construct uniform business cycle indicators for 9 countries and the euro area, 18 countries and 7 zone aggregates and 35 countries and 10 zone aggregates respectively.

The OECD also runs an indicator for the Netherlands, as do the Netherlands Bureau for Economic Policy Analysis (CPB, Kranendonk et al., 2005), Rabobank (Wolters, 2006), Statistics Netherlands (CBS, van Ruth and Schouten, 2004), the Centre for Economic Research at the University of Groningen (CCSO, Jacobs, 1998) and the Dutch central bank (DNB, Berk and Bikker, 1995). All these indicators aim at describing and forecasting¹ the Dutch business cycle, but differ in applied methodologies and empirical applications. We document the operational Dutch indicators and confront them with a band-pass filtered cycle that is required to be timely available and more broadly based than solely on manufacturing production.

In section 2, we explore the concept of the business cycle and discuss the method of band-pass filtering to separate the trend and cyclical motions of a variable. In section 3, we construct a composite reference index, which is based on timely available coincident macroeconomic variables that are closely monitored by business cycle dating committees. In a similar way we construct a composite leading indicator, which is based on cyclically leading financial, survey and real activity variables. As a case study, we show the performance of the indicator variables at the most recent turning point. In section 4,

¹An exception is Statistics Netherlands (CBS), whose indicator only aims at describing the Dutch business cycle.

we analyze the extent to which the composite indicators get revised as more data observations become available. Finally, the last section shows how the composite leading indicator can be employed to generate forecasts for GDP growth rates.

3.2 The business cycle

Business cycles can broadly be defined as oscillating motions of economic activity. Consecutive cycles are separated by turning points, that is peaks and troughs. The *classical* cycle considers the fluctuations of the level of economic activity, see Harding and Pagan (2002), while the *deviation* cycle considers the fluctuations around some trend. A third approach to business cycle representation is the so-called growth-rate cycle. Calculating growth rates can however also be interpreted as a detrending device², see for a discussion Harding and Pagan (2005a). The recession of the classical cycle is characterized by an absolute decline in the level of economic activity, that is by negative growth rates. The recession of the deviation cycle is characterized by economic growth rates that are below potential growth. Classical recession phases are always a sub-sample of the recession phases of deviation cycles³. Deviation cycles gained popularity, because periods of negative growth rates have been exceptional in industrialized countries since the Second World War. So, deviation cycles could more naturally be related to the fluctuations observed in the level of employment and unemployment. Moreover, the concept of a deviation cycle as the output gap between actual output and potential output gained policy relevance through the stronger focus on Taylor-rule driven monetary policy and cyclically adjusted government balances.

In this study, we adopt the definition of a business cycle as all the intrinsic cyclical motion visible in macroeconomic data consisting of waves within a specified frequency interval. This interval of business cycle frequencies corresponds with Burns and Mitchell's (1946) taxonomy of business cycles as waves lasting longer than a pre-specified minimum duration and shorter than

²Calculating growth rates is equivalent to applying the first difference filter, which induces an artificial phase shift $\theta(p) = \tan^{-1} \left[\frac{\sin\left(\frac{2\pi}{p}\right)}{1 - \cos\left(\frac{2\pi}{p}\right)} \right] > 0$ with p the duration of the cycle. (see page 275 Hamilton, 1994).

³Assume that a business cycle variable y_t admits the log-additive decomposition, $y_t = \tau_t + \psi_t$, for which τ_t is the trend and ψ_t the deviation cycle. Then the growth rate can be decomposed into trend growth and cyclical change: $\Delta y_t = \Delta \tau_t + \Delta \psi_t$. The recession phase of a deviation cycle $\Delta \psi_t < 0$ implies a lower than trend (or potential) growth rate: $\Delta y_t < \Delta \tau_t$. A classical recession phase, that is $\Delta y_t < 0$ implies a deviation cycle recession (reasonably assuming a non-negative trend growth rate: $\Delta \tau_t \geq 0$).

a maximum duration. The business cycle frequencies can be isolated by an ideal band-pass filter, which is then used as a benchmark for finite sample approximations. Band-pass filters are designed to adequately extract a specified range of periodicities without altering the properties of this extracted component. The assumption of an ideal band-pass filter is implicit in the work of Baxter and King (1999) (BK-filter), Christiano and Fitzgerald (2003) (CF-filter), Pedersen (2001) and to a certain extent Pollock (2000).

In order to isolate the cyclical component ψ_t out of the time series x_t with the period of oscillation between p_l and p_u , where $2 \leq p_l < p_u < \infty$ using a linear filter $\psi_t = B(L) x_t$, where $B(L) = \sum_{j=-\infty}^{\infty} B_j L^j$ and the lag operator L such that $L^j x_t = x_{t-j}$, the weights B_j of the ideal band-pass filter are defined as (see e.g. Priestley, 1982, page 275):

$$\begin{cases} B_j = \frac{\sin(j\omega_l) - \sin(j\omega_u)}{j\pi}, j \neq 0 \\ B_0 = \frac{b-a}{\pi}, \omega_u = \frac{2\pi}{p_u}, \omega_l = \frac{2\pi}{p_l} \end{cases} \quad (3.1)$$

Given the finite amount of T observations $\{x_t\}_{t=1}^T$, then $\hat{\psi}_t$ is a linear projection of ψ_t onto every element in the sample set, x_t , such that there is a different projection problem for each date t . The filter weights can be determined by a least squares minimization criterion:

$$\min_{\hat{B}_j} E \left[\left((B(L) - \hat{B}(L)) x_t \right)^2 \right], \forall t = 1 \dots T \quad (3.2)$$

The quadratic minimization problem (3.2) depends on the observations $\{x_t\}$ and Schleicher (2004) derives closed-form solutions⁴ in case x_t follows an arima (p, d, q) -process for $d = 0, 1$. The optimal approximate filter coefficients are the ideal band-pass filter coefficients weighted by the auto-correlation properties of the time series, which then need to be estimated first. As corollaries under additional restrictions, Schleicher (2004) shows that the approximations of Baxter and King (1999) and Christiano and Fitzgerald (2003) are special cases in which the underlying data generating process is white noise, i.e. arima(0,0,0), respectively a random walk, i.e. arima(0,1,0). The BK-filter coefficients look like $\hat{B}_j^{BK} = B_j + \frac{\Delta}{(1+2k)}, j = 0, \pm 1, \dots, k$, with $\Delta = - \left[B_0 + 2 \sum_{j=1}^k B_j \right]$. Note that BK-filter is a time invariant symmetric filter, i.e. $\hat{B}_j^{BK} = \hat{B}_{-j}^{BK} \forall t$ and does not induce an artificial phase shift⁵.

⁴Wildi's (2008) Direct Filter Approach is another finite sample approximation that extensively elaborates on spectral methods with a focus on the end of the sample.

⁵For the filtered sequence $\hat{\psi}_t = \hat{B}(L) x_t$, the filter induced phase-shift is determined by the

The CF-filter coefficients look like:

$$\begin{cases} \hat{B}_j^{p,f} = B_j, j = -p+1, \dots, f-1 \\ \hat{B}_{-p}^{p,f} = \frac{B_0}{2} - \sum_0^{p-1} B_j, \hat{B}_f^{p,f} = \frac{B_0}{2} - \sum_0^{f-1} B_j \end{cases} \quad (3.3)$$

Christiano and Fitzgerald (2003) question the restrictions of symmetry, i.e. $p = f$, so effectively allowing for filter induced phase shifts, and constancy, i.e. $p = f = k$ in order to always use of the full sample of available observations. Moreover, they state that the underlying Random Walk assumption for x_t in the optimization (3.2) represents Granger's (1966) typical spectral shape of a trending macroeconomic variable sufficiently well in practice. The resulting weights of the CF-approximation $\hat{B}_j^{p,f}$ are sample size dependent (as $T = p + f + 1$), which effectively transmits to the filtered series. However, an important feature of approximate band-pass filters is that $\hat{\psi}_t \rightarrow \psi_t$ as both $\hat{B}_j^{BK}$ and $\hat{B}_j^{p,f}$ converge to their (identical) infinite sample equivalents B_j as the window size increases due to additional observations, i.e. $k \rightarrow \infty$ and $p, f \rightarrow \infty$ respectively and $B_j \rightarrow 0$ for distant observations $j \rightarrow \infty$.

Canova (1999) examines the sensitivity of turning point classification to applying different trend-cycle decomposition procedures to GDP using as a benchmark the official turning point chronology of the NBER's business cycle dating committee. Of all the employed trend estimators, two filters, namely Hodrick and Prescott's (1997) filter (HP-filter) and the bandpass filter prove to be superior in detecting the NBER's turning points in terms of the number of missed turning points and false alarms. Based on a qualitative comparison, Zarnowitz and Ozyildirim (2002) conclude that the HP- and BK-filters provide comparable in-sample results to the NBER's Phase Average Trend (PAT, see Boschan and Ebanks, 1978) procedure, which is a data driven piece-wise linear trendfilter. Kranendonk et al.'s (2005) empirical analysis of Dutch data shows that the in-sample statistical criteria of the CF-filter are comparable to those of the BK- and HP-filter, but that the CF-filter shows the lowest sensitivity to new data. The tail behavior of the filter refers to the extent to which estimates get revised over time as more data becomes available. In an empirical analysis of cycle extraction methods, van Ruth and Schouten (2004) conclude in terms of tail behavior and real-time turning point detection that both the CF-filter and the HP-filter are subject to relatively sizeable revisions. This lit-

cross-covariance function, or cross-spectrum (see Priestley, 1982) $f_{\psi\hat{\psi}}$, between ψ_t and $\hat{\psi}_t$. This is given by:

$$f_{\psi\hat{\psi}}(\omega) = \hat{B}^{p,f}(e^{-i\omega}) B(e^{-i\omega}) f_x(\omega).$$

The phase is related to the complex part of $f_{\psi\hat{\psi}}(\omega)$. This equals zero, given that both $B(e^{-i\omega})$ and $f_x(\omega)$ are real, if and only if $\hat{B}(e^{-i\omega}) = \hat{B}(e^{i\omega})$, i.e., that $\hat{B}_j = \hat{B}_{-j}$.

erature shows that the HP-filter is currently one of the most widely employed macroeconomic trend estimator.

3.2.1 The Hodrick-Prescott filter

Hodrick and Prescott's (1997) HP-filter, which in the actuarial sciences is called the Whittaker-Henderson Type A method (a Whittaker, 1923) and is originally applied on macroeconomic data by Leser (1961), is a linear time-invariant symmetric filter, which can be characterised by its infinite moving average representation (cf. King and Rebelo, 1993)

$$B^{HP}(L) = \frac{\lambda(1-L)^2(1-L^{-1})^2}{1 + \lambda(1-L)^2(1-L^{-1})^2} \quad (3.4)$$

Hodrick and Prescott (1997) propose to minimize the squared deviations of a time series x_t from its trend τ_t subject to the smoothness constraint λ :

$$\min_{\{\tau_t\}_{t=1}^T} \sum_{t=1}^T (x_t - \tau_t)^2 + \lambda \sum_{t=2}^T [(\tau_{t+1} - \tau_t)(\tau_t - \tau_{t-1})]^2 \quad (3.5)$$

Solving the first order conditions for (3.5), the cyclical component can be recovered as $\hat{y}^{HP} = (\mathbf{I}_T - \mathbf{M}^{-1})x$, where $\hat{y}^{HP}$ and x T -vectors, $\mathbf{I}_T$ an identity matrix and $\mathbf{M}$ a Toeplitz matrix with diagonal band $[\lambda, -4\lambda, 1 + 6\lambda, -4\lambda, \lambda]$, initial and end conditions $M_{11} = M_{TT} = 1 + \lambda$, $M_{22} = M_{T-1, T-1} = 1 + 5\lambda$ and $M_{12} = M_{21} = M_{T-1, T} = M_{T, T-1} = -2\lambda$ and zeros elsewhere. The filter coefficients embodied in the term $(\mathbf{I}_T - \mathbf{M}^{-1})$ only depend on the smoothness parameter λ and are independent of the properties of the data generating process of the underlying time series.

On the contrary, Schleicher's (2004) closed-form solutions for (3.3) can be applied to the infinite sample HP-filter coefficients (3.4) and provides the optimal finite sample approximation employing spectral methods. Schleicher (2004) shows that at the end of the sample, when a filter is by definition one-sided, his approximation is less distorting in terms of leakage and phase shifts than the standard implementation $(\mathbf{I}_T - \mathbf{M}^{-1})$, which is employed by for instance Orphanides and van Norden (2002) and van Ruth and Schouten (2004). Moreover, Schleicher (2004) questions the rationale of interpreting the performance of the HP-filter $(\mathbf{I}_T - \mathbf{M}^{-1})$ as an approximation to the bandpass-filter (3.1), which is the implicit notion of for instance Guay and St.-Amant (2005), Nilsson and Gyomai (2008) and Pedersen (2001). In case the HP-filter (3.4) is to be interpreted as a highpass filter, i.e. $p_u = \infty$ in (3.1), then the parameter λ determines the cutoff duration p_l . Using spectral analysis, the

transfer function (see Hamilton, 1994, page 171) of the HP-filter can be obtained by substituting $e^{-i\omega}$ for L in (3.4) for $\omega \in [0, \pi]$, which is monotonically increasing from $B_\lambda^{HP}(0) = 0$ to $B_\lambda^{HP}(\pi) = 1$. Like Maravall and del Rio (2001), determining the stopband ω_l such that $B_\lambda^{HP}(\omega_l) = \frac{1}{2}$ then it can be shown using $p_l = \frac{2\pi}{\omega_l}$ that

$$B_\lambda^{HP}(\omega) = \frac{\lambda (1 - e^{-i\omega})^2 (1 - e^{i\omega})^2}{1 + \lambda (1 - e^{-i\omega})^2 (1 - e^{i\omega})^2} = \frac{1}{2} \Leftrightarrow \lambda = \left[4 \left(1 - \cos \left(\frac{2\pi}{p_l} \right) \right)^2 \right]^{-1}$$

So for instance, van Ruth and Schouten's (2004) value of $\lambda = 1000000$ for monthly data implies a stopband of $p_l = 198.6816$ monthly periods or 16.56 years. Note that Hodrick and Prescott's (1997) original $\lambda = 1600$ for quarterly data implies a stopband of $p_l = 39.69$ quarterly periods or 9.92 years, which equivalently corresponds to a $\lambda = 6.64$ for yearly and $\lambda = 128900$ for monthly data, cf. Ravn and Uhlig (2002)⁶.

Harvey and Jaeger (1993) provide an alternative representation of the HP-filter in the context of an unobservable components structural time series model. In our notation, $y_t^{HP} = [1 - B^{HP}(L)] x_t + B^{HP}(L) x_t = \tau_t + \psi_t$, which turns out to be modelled as $(1 - L)^2 \tau_t \sim N(0, \sigma_\epsilon^2)$ and $\psi_t \sim N(0, \lambda \sigma_\epsilon^2)$ and so λ is the inverse signal-to-noise ratio. If these specific model assumptions are not met in practice, the HP-filter can generate business cycle periodicity even if none is present in the data, (cf Cogley and Nason, 1995; Mohr, 2005).

3.3 Measuring the business cycle

Using the CF-filter to isolate business cycle frequencies requires an *a priori* specification of the minimum and maximum duration of a business cycle. Burns and Mitchell's (1946) taxonomy of business cycle frequencies is that *in duration business cycles vary from more than one year to ten or twelve years*. For U.S. data, Stock and Watson (2000) and Baxter and King (1999) use the convention of a business cycle as the cyclicalitv of GDP with a duration between 6 and 32 quarters. This convention is based on the NBER chronology of business cycle turning points, which shows that the shortest business cycle since 1858 lasted 6 quarters and the longest cycles are almost all shorter than 8 years. For the business cycle of the euro area, Agresti and Mojon (2003) employ an upper bound on the duration of 10 years. The euro area experienced 3 periods of recession during the last 30 years. van Ruth and Schouten (2004) perform an

⁶Ravn and Uhlig (2002) find approximate values of $\lambda = 6.25$ for yearly and $\lambda = 129600$ for monthly data.

empirical analysis on Dutch GDP using different cycle extraction methods and find that the average duration of the cycle is robust and lies between 10 and 11 years. The average length of the downswing from peak-to-trough is 3 years. A similar analysis for industrial production results in an average cyclical duration of just below 9 years and a peak-to-trough downswing of 2.5 years. These averages for industrial production are corrected for minor cycles. Based on this analysis, we use the CF-filter with a lower bound on the duration of 3 years and an upper bound of 11 years. The minimum duration requirement of 3 years aims to filter out the minor cycles in addition to the seasonal ones such that the resulting business cycle matches more closely its benchmark of cyclical GDP.

3.3.1 The composite reference index

While real GDP as the aggregate of all economic activity constitutes an important statistic for business cycle measurement, business cycle dating committees monitor several macroeconomic variables as the cyclical fluctuations of GDP don't always move synchronously to the ones of its underlying components, which can even move anti-cyclically. The Centre for Economic Policy Research's (CEPR) business cycle dating committee defines a recession as the significant decline in the level of economic activity, spread across the economy of the euro area and mostly visible in two or more consecutive quarters of negative growth of GDP, employment and other measures of economic activity. In establishing the chronology, the committee relies on the turning points of GDP, industrial production, employment, consumption and investment in both the aggregate euro area and in the three largest countries. The American NBER's business cycle dating committee relies on the analysis of GDP, real income, employment, industrial production and the wholesale and retail trade.

It is in general infeasible to identify one single variable, which covers a broad range of economic activity, represents the current stage of the business cycle and is available on a monthly frequency. In addition, other important criteria for selecting the variables for the composite reference index are a limited publication delay and minor data revisions. This means that new data releases are published shortly after the period has ended and that the initially published values are subject to only minor revisions during subsequent publications regarding the same period. The coincident reference index will therefore be composed as a composite index of several synchronous indicators. A composite index as a cross section of the cyclical motions of several variables constitutes a more robust standard of business cycle measurement since the idiosyncrasies of the individual indicators are averaged out.

Table 3.1: The coincident business cycle indicators that constitute the reference index

Variable	available since	Pub. delay (in weeks)	Cross corr. (with GDP)
Industrial production	1965	6	0.9
Household consumption	1977	7/8	0.9
Staffing employment	1967	1	0.6

These considerations lead us to include industrial production, household consumption and staffing employment in the composite reference index of the Dutch business cycle, see Table 3.1. The volume index of consumption by households is closely related to retail sales, since it consists only of those expenditures for which households pay themselves. The household consumption is only available from 1995 onwards and the appendix shows how it is extended until 1977 using the private consumption of the National Accounts. Industrial production constitutes approximately 25% of GDP and its share is declining relative to the share of services production. Despite this small share, the OECD indices are solely based on industrial production as its cyclical motion is considered a representative statistic for the business cycles.

Employment is a macroeconomic aggregate, which covers the whole range of economic activity. Both Dutch employment and staffing employment are released with a substantial publication delay, but the appendix shows how to create a time series for the staffing employment with a long history based on representative data from the market leader Randstad Netherlands and a timely available update based on representative data from the sector organisation⁷. Total employment for the Netherlands consists for 5% of staffing employment⁸. This market segment is most sensitive to business cycle motions, because companies can every moment adjust their use of staffing services immediately to changing market conditions. For instance, de Groot and Franses (2005) use Randstad-data as a coincident indicator in an error correction specification to forecast the growth rate of GDP.

The cyclical motions of the three variables from Table 3.1 are graphed in Figure 3.1. The composite coincident reference index is the unweighted average of the three standardised cyclical motions and is represented in the upper part of Figure 3.1. The cross-correlation of the composite coincident reference

⁷see <http://www.abu.nl>, which is the Dutch association of temporary work agencies and represents more than 60% of the market.

⁸Note that the output produced by a staffing employee employed at an industrial or agricultural company is booked as services production in the National Accounts.

index and the cyclical motion of GDP is 0.9 and is (slightly) higher than the cross-correlations of the three single indicators as shown in Table 3.1. The composite index as a cross section of the three variables constitutes therefore a more robust measure of economic activity since the idiosyncrasies of the individual indicators are averaged out.

Figure 3.1: Cyclical motion of GDP and respectively the reference index, industrial production, household consumption and staffing employment

3.4 The leading business cycle indicator

Theory and practice show that recessions originate from different sources and possess different characteristics. Stock and Watson's (2003) overview of the empirical literature on the usefulness of financial indicators for inflation and GDP-growth forecasting concludes that some asset prices possess significant forecasting power for some countries during some time periods, but that it's impossible to identify one single indicator showing a consistently good forecasting performance for all countries during all time periods. A combination of leading indicators in a composite leading index is therefore useful to pick up signals originating from different sectors in the economy.

Nardo et al. (2008) and Moore and Shiskin (1967) specify economic and statistical criteria for the selection of leading indicators. Economic reasoning guides the selection of the variables. The selected leading indicators should make up a balanced representation and cover the relevant sectors of the economy. The composition should ideally be well balanced between business and consumer survey results, financial series and real activity variables measuring economic behavior. Economic plausibility requires that the leading indicators should be supported by economic theory either as possible causes of business cycles or as quickly reacting to positive or negative shocks. The leading indicator should show consistent timing by systematically anticipating peaks and troughs with a rather constant lead time. Moreover, the conformity to the general cycle requires the leading indicator to possess good overall forecasting performance, not only at peaks and troughs. The statistics criteria require the leading indicators to be promptly available without publication delay and be subject to only minor data revisions. The statistics criteria partly explain the increasing use of surveys and financial variables as leading indicators. Both categories of variables are not subject to data revisions and financial data are directly observable.

The leading indicators are selected based on the abovementioned criteria. The most important selection criterion for the leading indicators is the conformity with the business cycle as represented by the composite reference index. This criterion is quantified by the cross correlation ρ of the coincident indicator, CI_t , with the cyclical motion of the suitably transformed leading indicator variable, LI_t . The transformation consists of taking the log-specification of the variables, if feasible⁹, and *a priori* classifying the variable as pro- or countercyclical. The maximum cross-correlation and the corresponding lead time l are then determined by shifting the leading indicator variable over time: $\rho^{\max} = \max_{l=0, \dots, 36} \rho(CI_t, LI_{t-l})$.

Table 3.2 shows the resulting optimal values for the collection of potential leading indicator variables (cf Berk and Bikker, 1995) fulfilling the requirements of timely availability and minor revisions. The lead times are determined for the cyclical parts of the potential leading indicator variables defined as cyclical motions with a duration longer than 3 and shorter than 11 years.

Another criterion is that the turning points of the leading indicators manifest themselves with a more or less constant lead time with respect to the turning points of the composite reference index. Bry and Boschan (1971) formulate a procedure to detect the turning points of the business cycle. Essentially, the

⁹No logarithm is taken in case the variable is scaled as a quotient (for instance interest rates, unemployment rate) or in case the variable can admit negative values or the value zero (for instance trade balance, stock finished products).

Table 3.2: Statistical criteria of potential leading indicators

Description	log	Maximum Correlation*	Lead*	Covered by
IFO future expectations		0.8065	9	
3 months interest rate		0.8728	26	
yieldcurve		0.7205	30	
order arrivals		0.7751	15	
expected business activity		0.8640	9	
stock finished products		0.7129	10	
Consumer confidence		0.8607	11	
Stock exchange index (AEX)	log	0.7283	7	
House Price real	log	0.7669	7	
Registered motor vehicles	log	0.8104	9	
OECD leading indicator country	log	0.7728	18	
United States, trend restored				
10 years interest rate		0.7195	21	yield - curve
M3	log	0.3190	5	
M3-real	log	0.1811	3	
M1	log	0.3901	3	
M1-real	log	0.6101	3	
M1-real growht rate		0.5360	23	
Level order position	log	0.4981	3	
Producer Confidence		0.7313	10	expected business activity
Consumer Survey Economic Climate		0.8804	12	consumer confidence
Willingness-to-buy		0.8204	3	
Bankruptcies	-log	0.9410	4	
Hwwa-commodity prices index	-log	0.4638	24	
Contractual wages (CAO)	-log	0.4687	3	
House Price	log	0.7023	8	real house price
Consumer price index	-log	0.6287	3	
Oil Price	-log	0.4294	30	
Capital account balance		0.5375	6	
Credit supply	log	0.4095	10	

Note: * Statistics are derived with respect to the composite coincident index.

Table continued

Description	log	Maximum Correlation*	Lead* Covered by	
Wages Industry, total	-log	0.5524	24	
Trade balance		0.5399	22	
Investment, businesses	log	0.9272	3	
Savings	-log	0.3231	3	
Exports	log	0.6842	3	
Imports	log	0.9164	3	
Unemployment rate		0.5641	3	
Wages, hourly	-log	0.5214	3	
Industrial Production, Belgium	log	0.6457	3	
Industrial Production, Germany	log	0.6880	3	
Industrial Production, U.S.	log	0.8274	6	OECD leading indicator
exchange rate	-log	0.0608	3	
Producer Prices, domestic	-log	0.5408	7	
Producer Prices, foreign	log	0.1088	3	

Note: * Statistics are derived with respect to the composite coincident index.

procedure determines a local extremum conditional on a minimum duration of the cycle and the requirement that peaks and troughs alternate. Harding and Pagan (2002) mathematically formalise this algorithm and analyse the resulting business cycles without reference to an explicitly specified trend development. This study adopts a dating algorithm for deviation cycles that determines local extrema conditional on alternating peaks and troughs and that a peak (trough) of the deviation cycle is positive (negative). So, a peak turning point exist at time t if $(1 - L) LI_{t-1} > 0$ and $(1 - L) LI_t < 0$ conditional on $LI_t > 0$ and the previous turning point being a trough. A minimum duration requirement is not necessary for the deviation cycle since the parameter specification of the CF-filter implies a minimum duration between two subsequent peaks or troughs of 3 years. Since subcycles with a duration less than 3 years are filtered out, the conditions on the dating algorithm turn out not to be binding as no local peaks (or troughs) materialize in the sense that $(1 - L) LI_{t-1} > 0$ and $(1 - L) LI_t < 0$ while $LI_t < 0$.

Table 3.3: Turning point dating of GDP, the composite coincident index and leading indicators

	GDP	coincident index	leading indicator	if future expectations	Three months interest rate	Term structure	order arrivals	Consumer confidence	expected business activity	Aex
peak	80q2	79m12	80m7	80m3	80m6	78m10	80m7	79m6	80m3	
trough	82q4	82m10	82m12	82m8	83m2	82m11	82m10	84m5	82m9	82m11
peak	85q3	85m10	85m11	85m12	85m12	86m4	85m7	86m11	85m7	86m7
trough	88q1	87m12	88m3	87m12	87m12	88m8	88m2	88m10	88m2	88m12
peak	90q4	90m11	90m7	91m6	90m1	90m7	90m3	90m8	91m1	90m12
[trough]						92m4				92m10
[peak]						93m4				
trough	94q1	93m7	94m2	93m10	93m12	95m3	94m2	93m8	93m9	
[peak]		95m10	96m5	95m10			96m4		95m12	94m5
[trough]		96m12		97m3						96m5
[peak]				98m10	98m1	98m5				
[trough]			00m9	00m6			98m3		00m3	
peak	00q2	00m9	01m12	01m7			01m8	01m8	01m10	01m4
trough	03q4	03m10	03m10	03m6	03m6	03m11	03m7	04m1	03m10	03m11
false signals	2(wrt GDP)	1	2	1	3	1		1	1	1
missed turning points		1			3	3	1	2	1	2
lead time (months)		7	9		26	30	15	11	9	7

Table continued

	GDP	coincident index	leading indicator	Real house price	New car registration	stock finished products	OECD indicator U.S.
peak	80q2	79m12	80m7	79m4	79m11	80m6	79m1
trough	82q4	82m10	82m12	83m1	82m5	82m12	83m9
peak	85q3	85m10	85m11	88m6	87m7	85m4	85m10
trough	88q1	87m12	88m3		89m8	87m6	87m5
peak	90q4	90m11	90m7		91m10	89m10	89m9
[trough]				92m7			
[peak]				94m8			
trough	94q1	93m7	94m2	95m12	94m5	93m8	93m2
[peak]		95m10	96m5	97m4		95m10	95m4
[trough]		96m12				97m7	96m12
[peak]						98m11	
[trough]			00m9	99m1		00m3	
peak	00q2	00m9	01m12	01m5	00m4	01m10	01m3
trough	03q4	03m10	03m10	04m6	03m5	03m10	03m12
false signals		2(wrt GDP)	1	3	0	2	0
missed turning points			1	3	2		0
lead time (months)			7	7	9	10	18

The dating of the turning points of the coincident indicators is graphically shown in Figure 3.1. The turning points dates of GDP, the reference-index and the eleven selected leading indicator variables are shown in Table 3.3. The dating of the turning points of the leading indicators occurred after alignment of the leading indicator with the number of months as shown in the last row under the heading lead time. Three observations result from the Table 3.3. The most important observation is that the nineties show two subcycles, which are visible in almost all selected leading indicators and industrial production, but not in GDP. Secondly, almost all indicators miss turning points and/or show false signals. A missed turning point is a turning point in the reference index, which the leading indicator fails to signal. A false signal occurs when the leading indicator shows a turning point, which eventually does not materialize in the reference index. Thirdly, the lead times of the leading indicator, as determined by the maximum cross correlation $\rho^{\max}$, show a lot of variation amongst the different turning points, but are not systematically biased and hence need no additional adjustment.

The leading indicator variables are primarily selected on meeting the statistical criteria of a minimum lead of 7 months and a minimum cross-correlation of 0.7. As Table 3.2 shows, a couple of variables that did pass the statistical tests are nevertheless not selected because the signal they represent is already accounted for by other leading indicator variables that are closely related. The composite leading business cycle index consists of the unweighted average of eleven leading indicators after standardisation. The lead time of the composite leading indicator is 7 months and the cross correlation between both series is 0.97.

The eleven selected leading indicator variables make up a balanced representation consisting of three financial series, four business and consumer survey results and four real activity variables, of which two supply and two demand related. The three financial variables are the short term interest rate, the yield curve and the stock price index. Low short term interest rates reduces financing costs and will spur investment demand. The AEX stock price index consists of the 25 most actively traded shares in the Netherlands and indicates the expected future corporate profitability and the underlying growth potential. Since the AEX index is dominated by multinational companies, the leading properties of this variable also reflect external economic developments. Moreover, Jansen and Nahuis (2003) show that negative returns on the stock exchange lower consumer wealth and consumer confidence, thereby causing a reduction in consumer spending. The term structure of the interest rate encompasses two transmission mechanisms: the expected future short term interest rates and the term premium for higher risk and/or lower liquid-

ity. Berk (1998) provides a literature review of the information content of the term structure of interest rates. An inverse curve is usually observed at the start of a recession period and acts therefore as a predictor, see for instance Estrella and Mishkin (1998). The four business and consumer survey results are expected business activity, the IFO-indicator of future expectations, domestic consumer confidence and the OECD's leading indicator for the United States. The IFO-indicator represents the economic expectations of producers in the Netherlands' largest trading partner Germany. The OECD's leading indicator for the U.S. reflects the short term outlook of the world's largest economy, whose business cycle is shown to be leading for the G7 countries, (see for instance Osborn et al., 2005). Finally, two supply related real activity variables are the order arrivals and the stock of finished products and the two consumption related real activity variables are the real house price and the registrations of new cars. Housing wealth, and therefore real house price changes, are an important factor in the total wealth of consumers. The registrations of new cars is a variable that quickly reacts to alterations in the business cycle.

Table 3.4 presents the compositions of the business cycle indicators for the Netherlands of different institutions. As a robustness check, Table 3.4 shows that almost all selected indicator variables in this study are also used by the institutions mentioned, even though different concepts of the business cycle and different trend estimators are applied. The only two unique leading indicator variables for the Netherlands are the real house price and the registration of new cars, although this latter variable is used in the OECD's leading indicators for other countries. Using the real house price as a leading indicator is however also unique in an international context.

The main source of difference with respect to the indices presented in Table 3.4 refers to the coincident index, which is more broadly based than just on industrial production, more timely available than GDP releases and excludes business cycle measurement based on sentiment indicators.

Table 3.4: Composition of business cycle indicators for the Netherlands by different institutions

	CFB	DNB	OECD
Since	1990	1985	around 1980
Latest revision	2003	2006	2002
Versions	the Netherlands, short index in 10 components and a long index		
Reference Series	gross domestic product (interpolated)	the Netherlands, U.S., Japan and the largest European countries industrial production	OECD-countries industrial production
Leading indicators	bankruptcies producer confidence (retailers) issued building permits (business, residential) OECD-indicator (Europe, United States) real money aggregate (M3) consumer-confidence (willingness to buy, economic climate) order arrivals (foreign, domestic, industry) 10-years interest rate business activity (expected, realized, construction, occupancy rate industry) IFO-indicator Evaluation stocks	3-months AIBOR/EURIBOR term structure of interest rates (yield curve) IFO-indicator for Germany expected business activity real money aggregate (M1)	Stock market index (AEX) stock finished products level order position expected order arrivals expected production
Horizon	3-4 months	7 months	±6 maanden
Trend filter	CF	CF	PAT / HP

Table continued

Since Latest revision Versions	Reference Series	Leading indicators	Rabobank	CBS*
			1982 2006 the Netherlands industrial production volume retail trade expected business activity processing industry IFO current index 10-years interest rate 3-months AIBOR/EURIBOR growth credit supply of the Rabobank organisation order arrival processing industry IFO-indicator of future expectations in Germany	2005 the Netherlands producer confidence unemployed labour force consumer confidence jobs of employees temporary jobs consumer survey; purchases of durables exports fixed capital formation business survey; orders received gdp total household consumption industrial production vacancies 10-years bond yield bankruptcies HP
Horizon	5 months			
Trend filter	CF			

Notes:

* CBS does not publish forecasts

3.4.1 Turning point prediction

The composite coincident and leading indices are presented in the upper part of Figure 3.2 and describe the cycle as the deviation of economic activity from its trend level. The x-axis of Figure 3.2 corresponds to the trend level and so, a positive (negative) output gap means that the level of economic activity lies above (below) its trend level. Moreover, an upward (downward) sloping output gap means that the level of economic activity is growing quicker (slower) than its trend growth. Following the peak in the business cycle, four subsequent phases can be distinguished as: i) slow-down; the cycle is positive, but downward sloping; ii) recession: the cycle is negative and downward sloping, which ends in a business cycle trough; iii) recovery: the cycle is negative, but upward sloping; iv) boom: the cycle is positive and upward sloping.

Based on van Ruth and Schouten (2004), the lower part of Figure 3.2 represents the clockwise development of the business cycle along the four phases. While the upper part of Figure 3.1 shows the current stance of the business cycle in a historical time perspective, the lower part shows the current stance in a longitudinal dimension as the multivariate presentation of each of the three individual coincident indicator variables, which are marked with a star and the eleven individual leading indicator variables, for which the seven months ahead signals of the synchronized series are displayed. The x-axis of the lower part of Figure 3.2 now represents the size of the increase/decrease of the indicator that is defined over a three month horizon¹⁰. So, the y-axis represents the magnitude of the indicator variable and the x-axis the magnitude of the derivative of the indicator variable. The movement from one period to another through the y-axis reflects a turning point materializing for the individual indicator. A concentrated cloud of points indicates that all individual indicators convey a similar message. Figure 3.2 presents the pseudo real-time snapshot of the indicator based on data up to and including December 2008¹¹, which eventually turned out to materialize as a turning point in the composite coincident index. The exercise is pseudo real-time as it is based on the fully revised vintage of the data available in February 2009 and therefore does not take into account revisions in published data, that is a so-called real-time macroeconomic data set. Both the composite leading index and the cloud of points clearly signal the materialising turning point.

¹⁰In the notation that will be developed below: $(1 - L)^3 y = y_{t|t} - y_{t-3|t}$

¹¹Given the publication delay of eight weeks for the consumption data, this snapshot of the indicator appeared on the website in March 2009.

Figure 3.2: Stance of the business cycle in an historical and a multivariate perspective

3.5 Indicator revision analysis

Since a policy institution uses the business cycle indicator as an instrument to measure and forecast the business cycle, the real-time assessment of the cyclical motion is a key focus. The cyclical indicator in the period t , say $CI_{t|t}$ is based on information available up to and including period t . If new data at period $t + 1$ becomes available, the cycle for period t based on all data $CI_{t|t+1}$ can change, since the real-time allocation of the dynamics to structural and cyclical forces is necessarily uncertain as information on the future path of the economy is missing. Orphanides and van Norden (2002) conclude that the unreliability of the cyclical estimator, that is the difference between the mid-sample $CI_{t|t}$ and the end-of-sample $CI_{t|T}$ estimates, is quite high for a broad class of detrending methods. Schleicher (2004) shows that the CF-filter is the optimal end-of-sample band-pass filter if the underlying time series follows a random walk (cf van Norden, 2002). For the last available observation, the CF-filter is effectively a one-sided filter, which becomes increasingly two-sided as more and more data becomes available. Since the filter weights are highest

for the leads and lags close to the target date, adding new data substantially contributes to the estimate of the cycle. This feature is even more prominent if the underlying series is seasonally unadjusted since then the variability of the newly arriving data transforms directly into volatile tail-behavior. The CF-filter is designed to filter out trending behavior at the cost of induced higher volatility at the business cycle frequencies and relatively pronounced leakage of the higher, that is seasonal, frequencies. So, the vulnerability of the CF-filter for high frequency movements translates into volatile tail behavior, although the filtered cycle will converge ever closer to its true value.

The filtered cycle at the target date is reasonably converged if it is based on the maximum length of a complete cycle, say k , of which half are future observations of the underlying variable. The filter window at the target date then covers at least a complete cycle symmetrically¹². In order to stabilize the volatile behavior of the tail, we average as much as possible over half a cycle. More precisely, let $CI_{t-i|t-i+j}^{CF}$ be the CF-filtered value of the cycle for period $(t-i)$ given data until period $(t-i+j)$, then the estimated cycle $CI_{t-i|t}$ for the period $t-i$ given data until and including period t is the average of the $(i+1)$ available CF-filtered values : $CI_{t-i|t} = \frac{1}{i+1} \sum_{j=0}^i CI_{t-i|t-i+j}^{CF}$ $\forall i = 0, \dots, k$.

In order to assess the reliability of the coincident and leading indices, we calculated their sequential updates starting with the indicator that runs from February 1977 until January 1988 and ending with the one that runs from February 1977 until February 2009, whose December 2008 snapshot is shown in Figure 3.2. The graph in the front of Figure 3.3 represents the collection of sequential updates for the coincident indicator and the graph in the back for the leading indicator. The z-dimension of the graph reveals the maturing of both indicators as for each point in time the end-of-sample estimates converge to their mid-sample equivalents.

Figure 3.3 shows that the standard deviation of the end-of-sample estimates is lower than its mid-sample equivalent. As more data arrives, the peaks and troughs of both indicators become increasingly more pronounced. In order to quantify the size of the adjustments, we calculate for the coincident indicator CI and the leading indicator LI , $Y = \{CI, LI\}$, the mean difference, $\text{mean} \left(Y_{t|t+i-1} - Y_{t|t+66} \right)$ and the mean absolute difference, $\text{mean} \left| Y_{t|t+i-1} - Y_{t|t+66} \right|$, between the mid-sample estimate of the cycle for which 66 months of future observations are available, $Y_{t|t+66}$, and the correspond-

¹²Note that Baxter and King (1999) recommend a symmetric window size of $\frac{1}{2}k = 12$ periods in case an upperbound of 32 periods for quarterly data is imposed and a size of $\frac{1}{2}k = 3$ periods for an upperbound of 8 periods for yearly data.

Figure 3.3: Consecutive updates of the coincident and leading indices

ing estimates based on a smaller number of i available observations, $Y_{t|t+i-1}$, $i = 1, 12, 23, 34, 45, 56$ where $i = 1$ is the end-of-sample estimate, over the sample period $t = 1988.2 - 1999.11$. All estimates from February 1988 onwards are at least based on a full cycle of past observations. Moreover, the estimates for November 1999 can be calculated including up to half the duration of a cycle of future observations. Finally the results, which are presented in Table 3.5, are averages over one complete cycle.

The mean absolute size of the adjustment between the end-of-sample and mid-sample estimates of the coincident indicator, which is a standardized variable, is 0.32. The mean absolute size of the adjustment of the leading indicator is of the same order of magnitude. The larger cross-section of eleven variables of the composite leading indicator, as compared to the three variables that constitute the composite coincident indicator, does not diminish the magnitude of the adjustment. The cyclical estimates of the individual variables get adjusted in the same direction and consequently do not cancel out in the cross-sectional average. The mean size of the adjustment errors of the coincident and leading indicators are also of the same order of magnitude, albeit of different signs. The negative mean adjustment error of the coincident

Table 3.5: The size of adjustment of the indicators from their mid-sample estimates.

variable/observations	i=1	i=12	i=23	i=34	i=45	i=56
CI, absolute	0.32	0.28	0.20	0.12	0.06	0.03
LI, absolute	0.35	0.27	0.18	0.12	0.08	0.05
CI-LI, absolute	0.23	0.20	0.18	0.17	0.17	0.17
CI	-0.12	-0.10	-0.06	-0.03	-0.01	-0.00
LI	0.12	0.08	0.04	0.03	0.02	0.01
CI-LI	-0.16	-0.10	-0.04	0.01	0.04	0.05

Notes:

Mean of the size of the adjustment of the coincident and leading indicators from their mid-sample estimates as more observations become available. The sample period is t=1988.2-1999.11

indicator entails that the upward adjustments around the peak was on average bigger than the downward adjustments around the trough in the sample period. On average both indicators converge towards each other, which implies that the mid-sample correlation is higher than end-of-sample one. The difference between both indicators $(CI_{t|t+i-1} - LI_{t|t+i-1+7})$ at each point in time t evolves as a larger number of observations i becomes available. Note that the end-of-sample estimate of the coincident index ($i = 1$) corresponds for every period with the estimate of the leading index for which the lead of 7 additional Observations are available. Table 3.5 shows that the mean absolute difference between both indicators declines and stabilizes after 24 additional observations.

3.6 Indicator based GDP forecasting

Marcellino (2006b, Chapter 4) reviews the literature and practices on the choice and evaluation of leading indicator models, possibly combined into composite indexes, and the development of more sophisticated methods to relate the indicators to a forecast for a target variable. The objective is to compare the forecasts of quarterly growth rates for GDP growth based on monthly indicator variables with the ones based on the composite indices. We employ the quarterly averages of monthly variables instead of more elaborate Mixed Data Sampling (MIDAS) regression models, cf. Ghysels et al. (2006). Rünstler et al. (2009) performs a large-scale GDP forecasting exercise involving large

data sets for ten European countries and the euro area based and conclude that factor models exploiting large monthly data sets outperforms traditional bridge equations. However, the Dutch results of Rünstler et al.'s (2009) analysis, which are based on the 14 single indicator variables that comprise the composite indices, perform remarkably well.

We replicate the exercise based on the quarterly Vector Autoregressive Model (VAR) that consists of a bivariate VAR containing GDP growth and the quarterly aggregate of a monthly indicator with the number of lags less than 5 quarters and determined by the Schwartz Information Criterion. The resulting forecast is then the unweighted average of the bivariate VARs containing each of the monthly indicator variables. The forecasting performance is confronted with the results of a pure autoregressive process for GDP growth, AR(1) and the naive forecast consisting of the recursive mean. The indicator based forecasts consist of: i) All, VAR forecasts based on all the fourteen single indicator variables¹³; ii) All synch, idem but after synchronisation of the indicator variables with the corresponding leads reported in Table 3.2; iii) Ci, bivariate VAR based on the composite coincident index $CI_{t|T}$ for $t = 1, \dots, T$ and $T = 2009.2$; iv) Li, bivariate VAR based on the composite leading index $LI_{t|T}$; v) Cirt, bivariate VAR based on the pseudo real-time composite coincident index $CI_{t|t}$; vi) Lirt, bivariate VAR based on the pseudo real-time composite leading index $LI_{t|t}$. The exercise starts with the estimation period of 1978.2-1999.4 and evaluation period 2000.1-2008.4 for a forecast horizon $h = -1, 0, 1, 2$. The forecast horizon $h = -1$ represents the "forecast" for the preceding quarter for which no GDP data release is yet available. Assume that the objective is to forecast GDP growth in the second quarter. We start forecasting in January: this forecast refers to next quarter GDP, $h = 1$, or first month one quarter ahead forecast. In moving forward in time we produce a forecast in each month, and with the GDP flash estimate being published in mid-August, run the final forecast on 1 August. We denote the latter as the second month preceding quarter "forecast", which is actually a backcast. This sequence of forecasts is applied to each quarter of our out-of-sample period.

Table 3.6 presents the forecast performance as the Root Mean Squared Error (RMSE) times 100 for different model specifications. The results suggest that the gains from synchronisation as represented by the difference in performance between the All and All synch specifications seem to be limited. Moreover, the composite leading index doesn't improve very substantially over the coincident index. However, the gains from the trend-cycle decomposition as represented by the difference between Ci and Li over All and All

¹³So, this specification is identical to the Dutch one in Rünstler et al. (2009) apart from the sample and forecast period.

Table 3.6: Forecasting performance of the leading indicators.

h	AR(1)	Naive	All	All synch	Ci	Li	Cirt	Lirt
2	150.01	156.64	146.17	142.17	84.35	83.19	120.01	121.81
1	139.90	155.70	127.62	125.53	78.25	78.15	105.02	101.63
0	123.51	154.59	104.05	102.91	71.63	71.13	86.49	81.83
-1	90.50	153.28	84.32	84.09	65.34	64.25	72.83	69.70

Notes:

The forecast performance is presented as RMSE*100 for different model specifications over the sample period 2000.1-2008.4. The model specifications consist of the naive and an autoregressive process for GDP growth. The other specifications consists of averages of bivariate VARs including the fourteen single indicator variables or one of the composite indices, which can be synchronised and/or in pseudo real-time format.

synch are quite substantial even when considering the pseudo real-time equivalents. Finally, Table 3.6 shows a 25% improvement of the forecast performance of the LI prt over the autoregressive AR(1) specification.

3.7 Conclusion

Policy institutions employ a business cycle indicator as an instrument to measure and forecast the business cycle and its turning points. This study constructs a business cycle indicator for the Dutch economy along the lines of the NBER-approach that consists of a reference index, which represents the current stage of the business cycle, and the indicator, which represents the developments of the cycle in the near future. The business cycle indicator is a deviation cycle that can be interpreted as an output gap, which is the fluctuation of economic activity around some trend. Hence, a most crucial choice regarding the deviation cycle is the method to extract the cyclical and trend component out of a time series of observations of an economic variable. In this study, we apply Christiano and Fitzgerald's (2003) approximate band-pass filter, which isolates from a time series all the intrinsic cyclical motion consisting of waves with lengths longer than 3 years and shorter than 11 years lengths. The minimum duration requirement of 3 years aims to filter out minor cycles and seasonal cycles such that the resulting business cycle matches more closely its benchmark of cyclical GDP.

The composite reference index is based on several coincident indicators that measure real economic activity being industrial production, household

consumption and staffing employment. For an indicator to be useful in practice, a timely update and therefore a limited publication delay for new observations of the source data is a crucial condition to be met. Since economic fluctuations originate from different sources, we combine a number of relevant leading indicators into a single composite index. The main selection criterion for the leading indicators are the lead time and the statistical conformity with the business cycle as represented by the composite reference index. The eleven selected leading indicator variables together make up a balanced representation of near future cyclical developments. The composite leading indicator consists of three financial series, four business and consumer survey results and four real activity variables, of which two supply and two demand related. The two composite indices are presented as time series that view the current stance of the business cycle in a historical perspective. Alternatively, the underlying single indicators are also presented as a cloud of points showing a multivariate assessment of the current phase of the cycle.

We analyse the revision of the cyclical estimator, that is the extent to which cyclical estimates are adjusted over time as more data becomes available. This feature emerges since the real-time assessment of the current cyclical position is dependent on the future, and therefore unknown, path of the economy. We show that the composite reference index and the composite leading index converge towards each other on average as more data becomes available. Consequently, the in-sample fit between both indicators is better than the end-of-sample fit. Finally, we compare the forecasts for quarterly growth rates of GDP based on the observed monthly indicator variables with the ones based on the composite indices. The results show that the composite indices perform better than their underlying indicator variables separately. Moreover, the gain in performance results from the trend-cycle decomposition.

3.A Appendix

The appendix describes the construction of the data underlying the reference index. Moreover, the appendix shows standardised cyclical motions of the single leading indicator variables.

3.A.1 Coincident indicators

As described in section 3.3.1, the composite reference index of the revised DNB-business cycle indicator consists of three monthly variables, that is industrial production, the household consumption and the staffing employment. The industrial production series is readily available with a long history. The time series of consumer expenditures and the staffing employment need to be constructed as described below.

Consumer expenditures

The basic series for consumer expenditures consist of the household consumption, which is published each month with a delay of 7 to 8 weeks and dates back to 1995. The variable is the volume index of household consumption, that is consumptive expenditures by households and non-profit institutions operating on behalf of households. The consumption by households only contains those parts of the actual individual consumption for which households pay themselves, which makes the variable more sensitive to business cycles. Moreover, consumption concerns the domestic consumption on Dutch territory and includes the consumption of foreigners. The series is available on a monthly basis since 1995. The series is extended over the period 1977-1995 with linearly interpolated quarterly data of private consumption from the National Accounts.

Staffing employment

The basic series for the staffing employment consists of the number of staffing hours employed through staffing agencies in the Netherlands. The number of staffing hours relate to staffing service agencies belonging to SBI¹⁴ code 74501, whose main activity is the matching of staffing employees of phase A. These employees are paid by the staffing agency, work on a contractual basis and

¹⁴The Dutch statistics office CBS employs a systematic hierarchical classification system for economic activities called the *Standaard Bedrijfsindeling* (SBI) that is compatible with the International Standard Industrial Classification of All Economic Activities (ISIC).

have no regular labour contract. The basic series of staffing employment is published each quarter with a delay of 5 months and dates back to 1999.

In order to construct the history of staffing services and to overcome the publication delay, we make use of the staffing services turnover data from Randstad Netherlands. The number of staffing employees employed through Randstad Netherlands has been recorded from 1960 onwards. Randstad Netherlands occupies since decades a stable market share of around 40%. The available data cover the period 1967q1 until 2003q4 containing the number of persons of staffing employment employed through Randstad Netherlands as published in de Groot and Franses (2005). Moreover, a third series is available from 2005 onwards describing the market turnover as published by the Dutch association of temporary works agencies¹⁵ (ABU), which represents 60% of the market. The data is published with a delay of only a couple of weeks and is sampled in 13 administrative periods of 4 weeks a year. The transformation for each year of 13 periods to 12 months runs with the following algorithm: $\text{month}(i) = (14-i)/13 * \text{period}(i) + i/13 * \text{period}(i+1)$ for $i=1, \dots, 12$.

So, the first part of the final series covers the period 1967-1999 and consists of the number staffing employees employed through Randstad. The second part consists of the total number of staffing hours employed in the Netherlands as published by the CBS and deals with the period 1999 until the most recently covered quarter. The third part consists of the monthly turnover data from the ABU and are connected additively after indexation with the overlapping year as the base.

3.A.2 Leading indicators

Figure A.3.4 shows the standardised cyclical motions for each of the eleven selected leading indicator variables together with the composite coincident indicator.

¹⁵see <http://www.abu.nl>

Figure A.3.4: Cyclical motions of leading indicator variables

©OECD. A.H.J. den Reijer – “Deviation Cycles in Manufacturing: Business Cycle Measurement and Leading Indicators,” Journal of Business Cycle Measurement and Analysis, vol. 3, no. 1, pp. 43–78, 2007.

Reprinted with kind permission from the Organisation for Economic Co-operation and Development

Chapter 4

Deviation Cycles in Manufacturing: Business Cycle Measurement and Leading Indicators

It's the cycle in the economy, stupid.

(inspired by) Bill Clinton

Abstract

The deviation cycles in the manufacturing industry of nine OECD-countries are identified by applying the Christiano-Fitzgerald band-pass filter. Turning points, low- and high-growth phases and other descriptive statistics are derived from these deviation cycles. A regression based test statistic is applied to test for duration dependence. Moreover, the international linkage between the cyclical motions in the manufacturing industry of two countries is investigated by measuring the degree of synchronisation. In addition to measuring the cyclical fluctuation, a composite leading indicator is constructed which replicates and predicts the deviation cycle in the manufacturing industry. This composite leading indicator is a single index composed of economic, financial and expectation variables possessing leading properties.

4.1 Introduction

Under the second pillar of the European Central Bank's (ECB) monetary policy strategy, information on the cyclical position of the economy is an important element in the assessment of the outlook for future price developments. The

I would like to thank Claudia Kerkhoff and Peter Keus for statistical assistance and Bouke Buitenkamp, Leo de Haan, Lex Hoogduin, Marga Peeters, Maarten van Rooij, Ad Stokman and participants at the internal lunch seminar for useful comments and fruitful discussions. Moreover, I would like to thank Bertrand Melenberg, Franz Palm, three anonymous referees, participants at the Econometric Society European Meeting (ESEM) 2003 conference, the January 2004 SOM workshop on Business Cycle Analysis at the university of Groningen, especially Robert Inklaar, and seminar participants at Sveriges Riksbanken.

cyclical instability of aggregate economic behaviour will sooner or later be reflected in similar patterns across various macroeconomic time series. The cyclical motions of the variables are mostly simultaneous or in a rapid succession of one another. This phenomenon of leading and lagging cyclical behaviour over time across macroeconomic variables can be exploited for the measurement and forecasting of the conjunctural position of the economy. The study of business cycles, which is of cyclical fluctuations in economic activity, goes back to the seminal contribution of Burns and Mitchell (1946), one of the earliest writings on the business cycle. As a starting point, it seems worthwhile to recall their definition:

"Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own."

This paper builds on the tradition of business cycle indicator research at De Nederlandsche Bank, see for instance Fase and Bikker (1985) and Berk and Bikker (1995). The first aim of this study is to identify cycles in the manufacturing industry for 9 OECD-countries: the Netherlands, Belgium, Germany, France, Italy, Spain, the United Kingdom, the United States and Japan. We will construct indicators for the nine OECD-countries with special attention to the Netherlands. Moreover, we will provide for each country descriptive statistics like the dating of cyclical turning points, moment properties like the amplitude describing the severity of the recessions and the strength of the booms. Moreover, we will test for duration dependence of low- and high-growth periods. These statistics are derived for each country separately. In addition, the international linkages of the cyclical motions in the manufacturing industries among countries can be measured by the fraction of time the cycles are simultaneously in an upturn and/or in a downturn. The second aim of this study is to forecast the cyclical motions in the manufacturing industries by exploiting the early signals provided by leading indicator variables. The cyclical movement of a broad set of macroeconomic variables will be evaluated on statistical grounds and accordingly selected to form the basis for the composite leading indicator. The selection criteria are a sufficient degree of similarity both in terms of correlation and in terms of turning points and, secondly, leading properties with respect to the cycle in the manufacturing industry.

4.2 Measuring business cycles

Business cycles can be described in general terms as the oscillating behaviour of economic activity. The business cycle consists of a peak in economic activity, a period of recession, followed by a trough and a period of expansion. Subsequent business cycles are therefore separated by the business cycle turning points. A recession of the classical cycle is by definition a period with a decline in the level of economic activity, see Harding and Pagan (2001; 2002; 2005b). In effect, negative growth is a sufficient condition for a classical recession. A recession of the deviation cycle, or a low-growth phase, means that the level of economic activity is below its trend, or potential, level. Given that the potential trend growth is positive, classical cycle recessions are always a subset of deviation cycle recessions¹ and there can be multiple classical cycle recessionary episodes within one single period of deviation cycle recession. The durations of low- and high-growth periods of deviation cycles are symmetric: on average is the economy half of the time in recession. Contrarily, the duration of classical cycle recessions are typically much shorter than expansions. Since periods of declining economic activity were relatively scarce in Western Europe since the 2nd World War, deviation cycles are more in line with the fluctuations in the levels of employment and unemployment. Moreover, deviation cycles, or output gaps, gained much policy relevance for Taylor-rule driven monetary policy and for cyclically adjusted government budget deficits.

As the trend cannot be observed directly, the strategic choice concerning the deviation cycle is the methodology to filter the trend, and implicitly the cycle, out of a time series of observations. Canova (1998) provides an overview of the comparative properties of cyclical components obtained by using different trend estimators on a couple of commonly used macroeconomic time series. He concludes that the existence of a set of business cycle stylized facts that is robust to varying trend estimators is misleading, because different concepts of a business cycle, and correspondingly different trend estimators, generate different economic objects which not necessarily have similar properties. Basically two concepts of a business cycle can be distinguished, which originate from Koopmans's (1947) measurement without theory debate on whether the economic considerations underlying a statistical based decom-

¹If a variable x_t , say GDP, admits the log-additive decomposition: $x_t = u_t + \psi_t$, where u_t denotes the trend and ψ_t denotes the deviation cycle, then growth, Δx_t , can be decomposed into: $\Delta x_t = \Delta u_t + \Delta \psi_t$, that is potential growth Δu_t and cyclical change $\Delta \psi_t$. Deviation cycle recessions, $\Delta \psi_t < 0$, correspond to periods of growth below potential growth: $\Delta x_t < \Delta u_t$, which then encompasses the classical recessions, that is $\Delta x_t < 0$ (given that potential growth is at least positive, that is $\Delta u_t > 0$).

position should be modelled explicitly? For the econometric model based approach, the optimal filter is the signal extracting device, which minimizes the mean squared error to the actual data generating process. Examples of filters based on econometric models are the exponential smoothing filter, the HP-filter and more recently, the Butterworth-filter known from electrical engineering, see Pollock (2000) and Gómez (2001).

The statistical based approach assumes the existence of an ideal band-pass filter, often characterized by isolating a pre-specified frequency interval, which is then used as a benchmark for finite sample approximations. The choice of the ideal band-pass filter as a reference model is justified by the definition of Burns and Mitchell (1946) of business cycles as recurrent waves lasting longer than a minimum and shorter than a maximum duration. This wave length interval corresponds one-to-one with the business cycle frequencies. Examples of filters based on the ideal band-pass filter are Baxter and King (1999) (BK-filter), Christiano and Fitzgerald (2003) (CF-filter) and Pedersen (2001). These papers focus on the optimal approximation to the ideal filter of a finite sample of data that is possibly non-stationary and auto-correlated. Harvey and Trimbur (2003) combine both approaches within a unified framework. They show that the least squares linear estimator of an Unobserved Components model (UC-model, cf. Harvey and Jaeger, 1993) corresponds to a generalized Butterworth filter, which is a specific type of band-pass filter. Conventional econometric specification of the UC-model implies a sub-optimal parameterization of the Butterworth-filter, as the ideal band-pass filter emerges only as the limiting case. Equivalently, the implied cyclical model corresponding to the ideal band-pass filter has no direct economic interpretation. A more general disadvantage of UC-model based filtering is that adding additional new data may result in a higher likelihood for an alternative model specification. Since the decomposition in trend and cycle is model specific, the alternative model can then induce a radical change in the historical shape of the business cycle. This unreliability of the output gap estimates is undesirable from a practical policy making perspective. Orphanides and van Norden (2002) conclude that the reliability, that is the extent to which output gap estimates are revised over time as more information arrives, is quite low for a broad class of detrending methods. As more data are released however, the estimate of the non-model based band-pass filtered cycle will eventually converge to its final value as the approximation converges to the ideal band pass, regardless of the specific choice of the approximation filter. Therefore, this study defines the cycle to be the optimal band-pass filtered component of the observed variable and remains agnostic about the structural model.

The BK-filter is an approximation to the ideal band-pass filter under side

constraints and renders a series stationary without inducing a phase shift. So, it does not shift the filtered cycles forward or backward in time. The CF-filter is a weighted approximation, which takes into account the empirical property of macroeconomic data known as Granger's (1966) typical spectral shape. The approximation of the ideal band-pass filter at the lower frequencies is consequently more important than at the higher frequencies. Schleicher (2004) confirms theoretically and empirically the optimality of the CF-filter, especially concerning the end-of-sample performance for current analysis². For these reasons, this study will use the CF-filter to extract cycles from time series.

A starting point in defining the range of business cycle periodicities is given in the introduction's quote of Burns and Mitchell (1946) saying between one year and ten or twelve years. Stock and Watson (2000) and Baxter and King (1999) apply the BK-filter using the convention of business cycles as variations in GDP between 6 and 32 quarters. This is based on the NBER business cycle reference dates, for which the shortest cycle since 1858 lasted 6 quarters and the longest ones are mostly shorter than 8 years. However, the latest two US-cycles lasted more than eight years. Agresti and Mojon (2003) use for the euro area business cycle the BK-filter with a periodicity band of 1.5 years and 10 years. They arrive at this higher upper bound by stating that the euro area experienced only three recessions over the last thirty years. This study adopts the 1.5-10 years convention of business cycle frequencies³, see also Kranendonk et al. (2005) concerning the Netherlands.

4.3 Detecting and dating Dutch deviation cycles

GDP is the aggregate statistic of all economic activity and provides the broadest coverage of the economy. As the accounting identity holds, GDP is the aggregate of both production and expenditures and of income. Deviation cycles as recurrent periods of above- and below-trend growth are visible in many economic activities and should therefore be visible in GDP as the aggregate of all activity as well. However, not all economic activity is affected

²Even though the CF-filter might show the best end-of-sample performance, this filter exhibits considerable more instability if it is applied to not-seasonally adjusted data. The (seasonal) instability is corrected for by averaging over (maximum) half a cycle (which is 60 months). More precisely: the definitive deviation cycle indicator DCI^{def} for $i = 1, \dots, 60$ for the period $t - i$ given data up until period t , equals to: $DCI_{t-i|t}^{def} = \sum_{j=1}^i DCI_{t-i|t-i+j}$, where $DCI_{t-i|t-i+j}$ is the CF-filtered cyclical component at period $t - i$ that is determined at period $t - i + j$.

³Noteworthy, Ravn and Uhlig (2002) show that the HP-filter is equivalent to a one sided band-pass filter. The standard parameter value of =1600 for quarterly data filters out cycles of duration longer than 8 years.

simultaneously by the deviation cycle. Some macroeconomic variables lead the cycle, a feature that will be exploited in the next section, and other variables are coincident or lag the cycle. The non-synchronized cyclical behaviour of some sub-components partly cancels in the aggregate GDP. For instance, the expenditure component government consumption is well-known to act as an automatic stabilizer of the general cycle due to counter-cyclical and lagging patterns of tax revenues and social security payments. The expenditure component exports tends to pro-cyclically lead the general cycle, particularly so for a small open economy as the Netherlands.

Since the cyclical motion of GDP does not parallel the ones of its underlying components, the business cycle dating committees of the European CEPR⁴ and the American NBER⁵ look at more macroeconomic variables for dating the turning points of the classical cycle. Apart from GDP, both committees analyse industrial production and total employment. The CEPR committee considers moreover investment and consumption while the NBER committee considers instead real disposable income and the wholesale and retail sales, see Hall et al. (2003). Business cycle committees officially establish a turning point of the classical cycle only after a considerable time period has elapsed since its occurrence. Since the cyclical indicator is used for current analysis policy purposes, it is of practical importance to base the indicator on monthly data that are released with a short publication delay and are only to a minor extent subject to revisions.

Stock and Watson (1989) provide a formal statistical framework to construct a coincident indicator for the United States based on the four macroeconomic variables analysed by the NBER dating committee. The indicator closely resembles McGuckin's (2001) (TCB) one, which is a contemporaneous weighted average of the same four variables. TCB does not straightforwardly construct similar indicators for other countries due to data availability and issues of data definitions. The international indicators of TCB rely on proxy data such as surveys for sales and unemployment rates for payroll employment. Moreover, personal income data is notoriously tedious to obtain at a high sampling frequency without a substantial publication delay. Only the industrial production series is an internationally widely available and comparable statistic.

In accordance with practices at the Nardo et al. (2008) and at De Nederlandsche Bank, see Berk and Bikker (1995), this study will use the manufacturing output volume as the reference series from which the cyclical indicator

⁴Centre for Economic Policy Research (CEPR). See <http://www.cepr.org/data/Dating>

⁵National Bureau of Economic Research (NBER). See <http://www.nber.org/cycles.html>

is filtered. Like GDP, industrial production is a main focus variable of both the CEPR and NBER business cycle dating committees. However, industrial production is available on a higher, that is monthly, frequency, the publication lag is considerably shorter and the revision of the numbers after the first publication is smaller than for GDP. The standardized CF-filtered deviation cycles of GDP, services production and industrial production on a quarterly basis for the Netherlands are displayed in Figure 4.1. The industrial production is, with a cross-correlation with cyclical GDP of 0.77, only approximately cyclically representative. Especially the mid-nineties show discrepancy between the cycles of services production and industrial production. These two types of production constitute 70% respectively 25% of GDP and agriculture producing the remaining 5% for the highly developed countries considered in this paper. The periods between a peak and a trough of the deviation cycle of manufacturing industry are the low-growth phases and are indicated in the graphs of Figure 4.1 by the shaded areas. The collection of peaks and troughs, the turning points, are derived from the deviation cycle according to the procedure outlined below. Equivalent to the standard calculus rule for finding extreme values, i.e. setting the first derivative of a function with respect to a variable equal to zero, a local peak in (the log of) a macroeconomic time series variable, x_t , occurs at time t if x_t exceeds values x_s for $t - k < s < t$ and $t + k > s > t$, where k delineates some symmetric window in time around t . This simple idea is the basis of the NBER procedures summarized in the Bry and Boschan (1971) dating algorithm. In practice, the Bry-Boschan algorithm with $k = 5$ for monthly data also applies some censoring rules ensuring a minimum cycle duration of 15 months and alternation of peaks and troughs. This algorithm has been applied on time series data in levels and so produces the dating of classical cycles. Now consider the deviation cycle variable $DCI_t = y_t - g_t$. The variable DCI_t is now the deviation of the log of a macroeconomic time series variable, x_t , from its long run CF-trend, g_t , and a similar turning point algorithm can be applied on this deviation series. The Bry-Boschan algorithm has been adapted by Harding and Pagan (2002) to date growth cycles. We perform our dating in this spirit and apply an algorithm that obeys the following rules: 1) peaks and troughs must alternate; 2) a peak (trough) must represent above (below) trend growth and so the corresponding $DCI_t > 0$ (< 0); 3) the turning point must be a local optimum:

$$\begin{aligned}
 &\text{a peak at } t \text{ if } (DCI_{t-p} \dots DCI_{t-1} < DCI_t > DCI_{t+1} \dots DCI_{t+q}), \\
 &\quad DCI_{t'} > 0 \text{ for } t' \in [-p \dots q] \\
 &\text{a trough at } t \text{ if } (DCI_{t-p} \dots DCI_{t-1} > DCI_t < DCI_{t+1} \dots DCI_{t+q}), \\
 &\quad DCI_{t'} < 0 \text{ for } t' \in [-p \dots q]
 \end{aligned} \tag{4.1}$$

Figure 4.1: Deviation cycles of the Netherlands

Since the CF-filter parameters are set such that the shortest possible cycle lasts at least 18 months, no two peaks or troughs can coexist within this time frame. Moreover, a deviation cycle is approximately symmetric; the cycle is roughly half of the time below trend level. The study therefore sets $p = q = 9$, which defines a symmetric window consistent with the a priori definition of the minimum duration of the deviation cycle of 18 months. The dating procedure provides the collection of dates at which the deviation cycle reaches a peak or a trough. A low-growth (high-growth) period is straightforwardly defined as the period between a peak (trough) and a trough (peak).

Applying the dating rules (4.1) to a couple of different deviation cycles for the Netherlands results in the collection of deviation cycle turning points which are represented in Table 4.1. It shows both the results of filtered GDP on quarterly basis and of filtered industrial production on monthly basis. Moreover, the filters applied are the CF-filter used in this study, the HP-filter and the Phase Average Trend (PAT) filter used by the Nardo et al. (2008). The last column of the table shows the chronology of turning points by the OECD of the Dutch deviation cycle. This chronology results from the comparison of turning points of PAT-filtered GDP and industrial production combined with expert judgement by country specialists. As shown in the last row, the turning points of both the CF- and the HP-filtered industrial production series produces deviation cycle turning points, as defined by (4.1), which closely match the OECD's final chronology. The mean absolute difference of the timing of the turning point dates as compared to this chronology is close to 3 months for all filtered series considered.

4.4 The empirics of international deviation cycles in manufacturing

The indicators of the deviation cycle in the manufacturing industry of the nine countries are calculated by applying the CF-filter and standardising the resulting cyclical series. Subsequently, the dating is performed by applying the dating rules (4.1) to the indicators. The successive above- and below-trend growth periods are shown together with the deviation cycle indicators in the manufacturing industry in Figure A.4.2 in the appendix. The deviation cycles are represented by the bold lines and the low-growth phases by the shaded areas. The construction of the composite leading indicators, which are shown in the graphs by the dashed lines, will be explained in the next section. In order to derive descriptive statistics and measure synchronization among the deviation cycles in the manufacturing industry, we construct a state variable

Table 4.1: Dutch deviation cycle turning points of industrial production and GDP

	Industrial Production index (manufacturing)			GDP		Dating ¹⁾²⁾
	HP ³⁾	CF ⁴⁾	PAT ¹⁾	CF ⁴⁾	PAT ¹⁾	
Through	67m07	67m06	67m05		67q2	67m08
Peak	70m09	69m12	70m03		70q4	70m06
Through	72m06	72m01	71m12		72q3	72m06
Peak	74m04	74m03	74m08		74q3	74m01
Through	75m07	75m07	75m08		75q1	75m07
Peak	76m11	76m08	76m09			[76m09]
Through	78m01	77m08	78m05	79q1		[78m05]
Peak	79m11	79m11	79m11	80q2	79q4	79m10
Through	83m05	83m01	82m11	83q1	82q4	82m12
Peak	85m12	84 m10	85m01	86q2	86q2	84m10
Through			86m05			[86m05]
Peak			87m01			[87m01]
Through	87m07	87m11	88m04	88q4	88q1	87m08
Peak	90m05	90m08	91m02	91q4	90q4	90m12
Through					91q3	
Peak					92q1	
Through	93m08	93m09	93m06		93q4	93m12
Peak	95m06	95m02	95m06		95q4	95m02
Through	96m08	97m03	98m08	97q2	96q4	96m12
Peak	98m03	98m01			98q1	[98m01]
Through	99m01	99m02			99q1	[98m12]
Peak	00m10	00m10	00m06	00q2	99q4	00m06
Through	04m01	03m10		03q4		
MAD ⁵⁾	3.3	2.6	2.7			

Notes: 1) Phase Average Trend (PAT), see OECD (2002); 2) [] = minor cycle; 3) Hodrick-Prescott (HP) with $\lambda=1000000$; 4) Christiano-Fitzgerald (CF) with lower bound and upper bound of 1.5 and 10 years respectively; 5) MAD = mean absolute difference in months of the timing of the turning point dates as compared to the OECD's chronology that is shown in the last column.

S_t , which equals one in high-growth periods and zero otherwise. The information $\{DCI_t, S_t\}$ provides for each country a picture of the characteristics of the cyclical motion and can be used to produce descriptive statistical measures

as duration, amplitude, steepness and sharpness of high- and low-growth phases, see for instance Harding and Pagan (2002). Moreover, the information $\{DCI_t, S_t\}$ can be used to assess synchronization among deviation cycles in the manufacturing industry of different countries, see Harding and Pagan (2005b) and Camacho et al. (2008).

The most commonly used descriptive statistics are the standard deviation, skewness and kurtosis. The standard deviation measures the variability of the indicator. Skewness measures the one-sided fattedness of the probability distribution of the indicator. So, it measures whether high-growth phases are relatively more exuberant than low-growth phases are severe. Kurtosis measures the two-sided fattedness and indicates whether large cyclical movements are relatively more common.

The state variable S_t separates above- and below-trend growth phases. A peak (trough) occurs at time t when $S_t = 1$ ($S_t = 0$) and $S_{t+1} = 0$ ($S_{t+1} = 1$). A high- (low-)growth phase consists of the period beginning the month after the date of a trough (peak) and ending at the date of a peak (trough). High growth phases and TP (trough-to-peak) and low-growth phases and PT (peak-to-trough) will be used interchangeably. The average duration⁶ of high- and low-growth phases⁷ are:

$$DUR_{TP} = \frac{\sum_{t=1}^T S_t}{\sum_{t=1}^{T-1} (1 - S_{t+1}) S_t} \quad (4.2)$$

The numerator in (4.2) measures the total time spent in low-growth phases and the denominator counts the number of peaks. The average amplitude straightforwardly equals:

$$AMP_{TP} = \frac{\sum_{t=1}^T S_t \Delta DCI_t}{\sum_{t=1}^{T-1} (1 - S_{t+1}) S_t}. \quad (4.3)$$

Thinking of a phase as a triangle with as base its duration and as height its amplitude, then the steepness of a phase is naturally measured by the slope of the triangle:

⁶In calculating average of high- and low-growth phases one has to deal with uncompleted phases at both ends of the sample. The statistics will only be based on completed phases and therefore the summation runs from the beginning of the first completed phase until the end of the last one rather than over $1, \dots, T$.

⁷The corresponding formulas for low-growth phases, PT, are obtained by replacing S_t by $(1 - S_t)$.

$$STEER_{TP} = \frac{AMP_{TP}}{DUR_{TP}} \quad (4.4)$$

In addition to looking at the steepness over the phases, one might want to compare the shape of the cycle in the early part of the cycle to that in the later part. A simple way of capturing this shape is to divide a phase into two parts and to compare the average growth rates in the first part to those in the second part. This statistic is called sharpness and essentially measures a degree of convexity of the indicator, which is the relative steepness of the early stage in the cycle. Let i be the i^{th} phase and d_i be the duration of this phase, starting at time t_i . The measure of the shape of a single phase can then be defined as:

$$v_i = \left[\frac{1}{\lfloor d_i/2 \rfloor} \sum_{k=1}^{\lfloor d_i/2 \rfloor} (DCI_{k+t_i} - DCI_{k+t_i-1}) - \sum_{k=\lfloor d_i/2 \rfloor+1}^{\lfloor d_i/2 \rfloor} (DCI_{k+t_i} - DCI_{k+t_i-1}) \right] \quad (4.5)$$

Thus, if in a TP-phase the growth is faster in the first half of its duration than in the second half, then this measure will be positive. If in a PT-phase the decline is slower in the first half of its duration than in the second half, then this measure will be positive. If a cyclical indicator shows these two positive measurements it implies that the peaks are relatively long lasting and rounded and the troughs are relatively short lasting and pointed.

Averaging (4.5) over all similar phases can be defined as:

$$SHARP_{TP} = \sum_i S_{t_i} \left(\frac{d_i}{\sum_t S_t} \right) v_i. \quad (4.6)$$

In equation (4.6), S_t serves as an indicator function which equals one if the i^{th} cyclical phase starting at time t_i is a high-growth phase. Moreover, the equation is a weighted moving average of (4.5) over respectively high- and low-growth phases. The weight for a particular high-growth phase is the duration of that phase divided by the total time spent in TP-phases over the sample. So, this weighting scheme gives more weight to longer phases and the sum of the weights is equal to one. The descriptive statistics for the nine deviation cycle indicators of the manufacturing industry are presented in Table 4.2. Almost all countries tend to display kurtosis indicating that large cyclical movements in the manufacturing industry are relatively common. The Netherlands and to a lesser extent Belgium stand out on stability,

but not on moderateness. Stability is revealed by the volatility of the cyclical behaviour around the trend and by the amplitude. This is the absolute deepness of a cycle decomposed into the ones belonging to low- and high-growth phases. Both the Netherlands and Belgium show the lowest numbers for all nine countries on the standard deviation and the amplitude. Moderateness is revealed by the steepness statistic. Steepness can be interpreted as the motion deepness of a low- or high-growth phase like amplitude can be seen as the absolute deepness of a low- or high-growth phase. For both phase types, the Netherlands and to a lesser extent Belgium show among the highest numbers for the steepness statistic indicating that the deviation cycle of the manufacturing industries for both countries move quickly from one turning point to another. Japan's manufacturing sector acts as the mirror image of the ones of these two countries by showing the highest cyclical volatility and the highest cyclical severity both in terms of amplitude and steepness for both PT- and to a lesser extent TP-phases. The deviation cycle of the manufacturing industry of the U.S. also display high cyclical volatility and severity. The negative skewness indicates contraction phase outliers indicating at relatively severe recessions. The large slowdown sharpness and the large acceleration sharpness for the U.S. reveals a pattern of smooth rounded peaks and pointed deep troughs. Italy's manufacturing industry demonstrates the largest positive skewness and among the highest peaks, that is the average amplitude of the high-growth phase. Both statistics indicate large upward cyclical outliers. The quite moderate sharpness statistic reveals that the Italian cyclical pattern surrounding turning points is particularly smooth. Germany's manufacturing industry shows a cycle with a relatively long TP-phase compared as to the PT-phase. On the other hand the motion of the low-growth phases is relatively fast according to the steepness statistic and relatively deep according to the negative skewness. The sharpness statistic reveals that the German turning points are the most sharply pointed ones of the nine countries. Therefore, the German indicator looks somewhat like a saw tooth.

The descriptive statistics show the idiosyncrasy of the deviation cycle of each individual country's manufacturing industry as compared to a sinusoid cyclical motion. This stylized cyclical pattern implies a deterministic periodicity. The periodicity of the cycle is the amount of time between two peaks, or two troughs. A complete cycle covers therefore two consecutive phases, which is either a PT-TP or a TP-PT period. The theoretical literature considers the cycle as the total sum of sub-cycles identified by theoretically founded durations. Impulses from inventory holdings, (infrastructural) investments and technological breakthrough innovations are the driving forces behind respectively the Kitchin, Juglar, Kuznets and the Kondratieff cycles.

Table 4.2: Summary statistics for the deviation cycle indicators of the manufacturing industry

	NLD	BEL	FRA	DEU	ITA	ESP	JPN	GBR	USA
Standard deviation	1.67	2.16	2.50	2.64	2.67	2.33	3.38	2.47	2.82
Skewness	-0.38	-0.0091	-0.56	-0.25	0.21	-0.68	0.16	-0.13	-0.35
Kurtosis	3.51	4.08	4.78	2.68	3.25	3.69	2.67	3.14	3.55
Duration	TP 26.00	22.00	24.83	24.90	33.40	38.83	26.00	36.75	27.00
	PT 22.00	28.13	22.33	28.00	33.50	30.67	39.00	33.40	36.50
Amplitude	TP ¹⁾ 2.87	2.67	2.44	2.93	3.40	3.02	3.39	3.00	2.74
	PT 2.42	2.38	2.59	2.31	2.83	2.97	3.12	2.39	2.51
Steepness	TP ¹⁾ 0.11	0.12	0.098	0.12	0.10	0.078	0.13	0.098	0.10
	PT 0.11	0.085	0.12	0.083	0.085	0.097	0.080	0.072	0.069
Sharpness	TP 0.015	-0.017	-0.038 ²⁾	-0.023	-0.66 ²⁾	0.053	-0.14 ²⁾	0.142)	0.083
	PT 0.015	-0.012	-0.042 ²⁾	-0.017	-0.55 ²⁾	0.067	-0.097 ²⁾	0.10 ²⁾	0.061

Notes:

The statistics are calculated over the sample 1965:1-2001:9 and represent the deviations from trend series which are not standardised. The mean of the deviation cycle is by construction approximately zero. So, standardising approximately only implies dividing the deviation time series by its standard deviation. PT means the phase Peak-to-Trough and is therefore synonym to low-growth phase. Equivalently, TP is synonym to high-growth phase.

The skewness measure is $\mu_3/\mu_2^{1.5}$ and the kurtosis measure is μ_4/μ_2^2 , where μ_r is the r^{th} central moment. The skewness of a symmetrical distribution, such as the normal distribution, is zero. Similarly, the kurtosis of the normal distribution is 3.

1) The numbers are presented with opposite signs for the sake of comparison with periods of high growth.

2) These numbers are multiplied by 100.

The elapsed time between two consecutive peaks would then be a fixed period T , for which the duration depends on the type of the subcycle. So, if DCI_t is identified as a peak, then $DCI_{t+\tau}$ must also be a peak with fixed duration time of $\tau = T$ periods, whereas in case of weak periodicity τ is considered as a stochastic duration variable with a mean of T . For $\tau > T$, the probability that the cycle switches to the other phase increases then with the ongoing duration of the current phase. Diebold and Rudebusch (2003) refer to the feature that the probability of a turning point is some function of the age of the cycle as duration dependence. Their evidence for the existence of some duration dependence is corroborated by Ohn et al. (2004) using novel discrete-time test statistics. A minimum duration on the complete deviation cycle in the manufacturing industry of 18 months is imposed in this study by the parameter settings of the CF-filter. The appendix shows that Ohn et al.'s (2004) discrete test statistic is still valid to test duration dependence if the cycle is subject to minimum duration requirements. The hypothesis of no duration dependence is for both low- and high-growth phases for almost all countries rejected, according to the reported statistics in Table 4.3.

Table 4.3: Statistics of duration dependence of the deviation cycle in the manufacturing industry

	NLD	BEL	FRA	DEU	ITA	ESP	JPN	GBR	USA
TP	-4.35	-7.98	-1.93	-5.13	-1.25	-3.74	-4.53	-4.48	-7.58
PT	-6.04	-2.75	-1.86	-3.82	-2.71	-2.71	-5.91	-2.8	-4.62

Notes:

The sample period is 1965:1-2004:2. The numbers are the Newey-West t-statistics for the null hypothesis of no duration dependence. PT means the phase Peak-to-Trough and is therefore synonym to low-growth phase. Equivalently, TP is synonym to high-growth phase.

The next step after examining the characteristics of the deviation cycle in the manufacturing industry is to analyze the relationships between them. More precisely, we want to examine how closely the cycles of individual countries are synchronized to one another. Candelon et al. (2009) propose a small sample test procedure for multivariate business cycle synchronization that is based on the pairwise correlation between two countries' boom or recession phases. We follow Harding and Pagan's (2001) index of concordance that measures the fraction of time that two cycles spend in the same phase, be it the boom or recession phase. Let S^i and S^j denote the state variables of the deviation cycle in the manufacturing industry of countries i and j . The index

of concordance is defined as:

$$IC_{ij} = \frac{1}{T} \sum_{t=1}^T \left\{ S_t^i S_t^j + (1 - S_t^i) (1 - S_t^j) \right\}. \quad (4.7)$$

For all nine countries the results for the index of concordance are presented in the upper triangle of Table 4.4. The italic numbers in the lower triangle are the cross correlation coefficients of the two concerning deviation cycles in the manufacturing industry. Six of the nine countries examined in this study are part of the euro area and account for more than 90% of euro area GDP. These six countries can form 15 unique couples⁸ which provides 15 indices of concordance (ICs) measuring euro area synchronization of deviation cycles in the manufacturing industry. The average of these 15 indices is 0.74 and the average correlation is 0.66. As can be calculated from Table 4.3, the manufacturing industries of two euro area countries are on average roughly three quarters of the time in the same conjunctural state. Both the correlation and the IC of the big euro area countries Germany and France are both 0.68 and those between Germany and Italy only 0.32 respectively 0.65. The average IC and correlation of the potential euro area member the United Kingdom and the five euro area countries equal 0.64 respectively 0.56, which is even worse than the average IC of 0.67 between the United States and the euro area countries and slightly better than the average correlation of 0.54 between these two economic blocks. Moreover, the IC and the correlation between the United Kingdom and the United States equal 0.75 and 0.71 respectively. This indicates that the United Kingdom's deviation cycle in the manufacturing industry is more synchronized with the United States' one than with the cycle of the euro area.

4.5 Leading indicators

The deviation cycles provide a signal for the current conjunctural state in the manufacturing industry. Cyclical motion in the economy is reflected in various macroeconomic time series. Depending on the timing behaviour of their cyclicity vis-à-vis the deviation cycle in the manufacturing industry, variables can be classified as leading, coincident or lagging. Leading variables give by nature an early signal on the conjunctural position. We will exploit this feature by selecting a set of leading variables and transform them into a single composite leading indicator that replicates and predicts the deviation cycle in

⁸One could also construct an index of concordance which only measures the simultaneity of more than two countries together. It then measures the fraction of time that all countries are in the same state. For example, the concordance of Germany, France and Italy together is 0.57

Table 4.4: Index of concordance and correlation between indicators of deviation cycles in the manufacturing industry

	NLD	BEL	FRA	DEU	ITA	ESP	JPN	GBR	USA
NLD		0.81	0.68	0.81	0.70	0.67	0.82	0.63	0.63
BEL	<i>0.82</i>		0.77	0.68	0.71	0.8	0.71	0.68	0.71
FRA	<i>0.76</i>	<i>0.86</i>		0.68	0.81	0.86	0.61	0.59	0.65
DEU	<i>0.80</i>	<i>0.66</i>	<i>0.67</i>		0.65	0.69	0.76	0.66	0.67
ITA	<i>0.64</i>	<i>0.64</i>	<i>0.66</i>	<i>0.31</i>		0.76	0.68	0.62	0.63
ESP	<i>0.58</i>	<i>0.67</i>	<i>0.81</i>	<i>0.47</i>	<i>0.61</i>		0.65	0.65	0.74
JPN	<i>0.74</i>	<i>0.76</i>	<i>0.65</i>	<i>0.52</i>	<i>0.65</i>	<i>0.52</i>		0.63	0.65
GBR	<i>0.66</i>	<i>0.72</i>	<i>0.58</i>	<i>0.37</i>	<i>0.47</i>	<i>0.53</i>	<i>0.60</i>		0.75
USA	<i>0.55</i>	<i>0.6</i>	<i>0.59</i>	<i>0.40</i>	<i>0.51</i>	<i>0.57</i>	<i>0.58</i>	<i>0.71</i>	

Notes:

The sample period is 1965:1-2001:9. The indices of concordance are represented in the upper triangle of the table. The correlation coefficients are represented in *italics* in the lower triangle of the table.

the manufacturing industry. The combination of leading indicators is useful to pick up signals from different sectors of the economy. Stock and Watson (2003) conclude in their empirical literature review on the usefulness of financial indicators to forecast GDP growth and inflation that some asset prices predict inflation or output growth in some countries during some periods, but which series predicts what, when, and where is difficult to predict. As the forecasting literature moreover suggests that the simplicity of a model often implies forecasting power, this study aims to predict the deviation cycle in the manufacturing industry in a non-model based framework. Marcellino (2006a) notes the overall good forecasting performance of the simple non-model based leading indicators concerning the latest two recessions in the U.S. in an overview of the different approaches and methods for the construction, use and evaluation of leading indicators, both in the academic literature and in the forecasting practice.

Potential leading variables that can be expected to lead and predict the general business climate are gathered and will be called basic indicators. The set of potential basic indicators is then screened in the spirit of the formalized scoring system of Moore and Shiskin (1967). The collection of basic leading indicators should be a balanced representation of the total economy. Each basic indicator should be economically plausible, so it possibly causes the

deviation cycle in the manufacturing industry or it reacts quicker to shocks. Moreover, the series should be statistically reliable, that is a long history of observations with as few interruptions and definition alterations as possible. The data should be available with a minimum publication delay and not be subject to substantial revisions after the first publication. The basic indicator should conform to the deviation cycle; it has good forecasting properties, not only at peaks and troughs. Finally, the basic indicator should possess a consistent timing as a leading indicator; systematically anticipate peaks and troughs with a rather constant lead time. The cyclical parts of all basic indicators are compared to the deviation cycle indicator of the manufacturing industry in order to determine the numbers of months by which the basic indicators lead the cycle, i.e. the timing. The measure of this timing is the lead in number of months corresponding to the maximum cross correlation coefficient between the deviation cycle indicator and the lagged basic leading indicator (BI):

$$\rho(l) = \frac{\frac{1}{T} \sum (DCI_t - \overline{DCI})(BI_{t+l} - \overline{BI}_l)}{\sqrt{\frac{1}{T} \sum (DCI_t - \overline{DCI})^2 \frac{1}{T} \sum (BI_{t+l} - \overline{BI}_l)^2}}, \quad (4.8)$$

$$l^* = \arg \max_l \rho(l).$$

In equation (4.8) $\overline{DCI}_t$ and $\overline{BI}_l$ are the averages of DCI_t and BI_{t+l} respectively, over t . The magnitude of the correlation coefficient (4.8) measures the degree of cyclical similarity of the deviation cycle indicator of the manufacturing industry and the basic indicator. Moreover, the corresponding l^* represents the optimal lead of the basic indicator. Analogous to Berk and Biker (1995), we screen a set of potential leading variables consisting of about 40 monthly time series for each country on the criteria of a cross correlation $\rho(l^*)$ of at least 0.5 and a corresponding maximum lead l^* of at least 5 months. These boundaries are chosen as high as possible under the restriction that a sufficient number of basic leading indicators are selected. The selection statistics for the selected basic leading indicators are presented in Table 4.5. The first number in each cell is the optimal fine tuned lead l^* and the second number is the maximum cross correlation $\rho(l^*)$ for the concerning leading indicator and country. The selected leading indicators are of an economic, financial and monetary nature and confidence and expectations surveys.

Of the economic flow variables consumption, investment, exports and imports only the consumption in Japan shows a leading cyclical pattern. Consumption can straightforwardly be seen as a demand pull trigger for industrial activity. Inventories are a result of non-matching economic flows and for most countries the storage of final products provides an early signal for production activity. Other types of inventories are those at the start of the produc-

tion process, like the level order positions and issued building permits. Both variables pro-cyclically lead the deviation cycle in the manufacturing industry. Prices and wages are a fundamental part in every economic model describing cyclical fluctuation. Following Stock and Watson (2000), a general pattern emerges of leading, countercyclical price levels and lagging, pro-cyclical rates of inflation. The nominal wage index exhibits a pattern quite similar to consumer prices, probably due to the contractual indexing of wages. Consumer prices, sales prices, total wages and hourly wages emerge in this study as cyclical leaders. Moreover, labour costs per unit product are a combination of wages and productivity and on average leads the cycle countercyclically. The same pattern shows up for the world market prices of commodities, which basically is a cost variable for production. Dominant classes of leading variables concerning business cycles are those specifically related to the future. Survey variables that pro-cyclically lead the general cycle are consumer and producer confidence, expected future industrial production, prospects total economy, judgment of order arrivals and the IFO-indicator for Germany⁹. Financial variables deal with consumption and investment possibilities and therefore also with the choices made by private and public agents. As the sum of discounted expected future cash flows, share prices are, like surveys, future linked variables. Interest rates are a cost of capital and therefore partly determine the budget constraint. Both short term and long term interest rates lead the cycle with positive interest rates associated with cyclical declines in output. The lead of the short term interest rate is approximately 1.5 years for most countries to over two years for Japan, see Table 4.5. The connection between short term interest rate movements and the output gap is well established in the Taylor rule. The spread between the short and long term interest rates has long been recognized as a leading indicator. However, we prohibit taking this variable into the model together with both the short and the long term interest rates.

Monetary aggregates play an important role in the determination of price levels since nominal frictions in the economy can cause movements of real quantities. Following the time inconsistency literature, in the short run a monetary expansion creates a boost in employment and output that causes in the long run only a rise in the price level. We consider three monetary aggregates, if available, namely M1, M2 and M3 and test them both in nominal and real terms and in levels and growth rates. However, we find only for Spain a monetary aggregate functioning as a leading indicator. Stock and Watson (2003) note that the contemporaneous correlation between the cycles of nominal M2 and output has drastically declined since the early eighties for the U.S.

⁹We choose a subindex of the IFO-indicator that refers to future developments.

Table 4.5: Results of selection leading indicators

	NLD	BEL	FRA	DEU	ITA	ESP	JPN	GBR	USA
Consumer confidence			5/0.60	6/0.63			9/0.51	12/0.75	
Consumer price index*	8/0.54	17/0.72			9/0.79		16/0.53	14/0.83	
World market prices commodities*					18/0.61	15/0.55			14/0.84
MI*						22/0.78			
Equity price			14/0.62						9/0.52
Short term interest rate*	19/0.51	18/0.93		12/0.61	15/0.80		27/0.64	18/0.88	18/0.80
Long term interest rate*	18/0.53	16/0.57			12/0.68		25/0.53		15/0.71
IFO	5/0.74	7/0.63		5/0.91	9/0.60	6/0.60			
Expected business activity	6/0.84								
Expected future industrial production								8/0.94	

Notes: The first number in each cell is the optimal fine-tuned lead (9) and the second number is the maximum correlation (8) of the corresponding leading indicator variable with the deviation cycle indicator of the manufacturing industry, *DCL*. The variables marked with a '*' counter-cyclically lead the cycle instead of pro-cyclically. The composite leading indicator, *CLI*, is the average of the synchronized and standardized basic leading indicator variables.

Table continued

	NLD	BEL	FRA	DEU	ITA	ESP	JPN	GBR	USA
Storage final products*	6/0.59		5/0.61	5/0.90		6/0.70	13/0.70	10/0.92	9/0.77
Prospects total economy					10/0.72			11/0.95	
Judgement of order arrivals			5/0.70	6/0.85	15/0.73	6/0.70			
Level order position			5/0.82	5/0.91					
Hourly wage industry*	6/0.57	12/0.78		5/0.88	11/0.89		11/0.56		8/0.73
Domestic sales prices*		9/0.56			15/0.73		23/0.68	13/0.82	7/0.90
Producer confidence			5/0.74						
Consumption							10/0.67		
Total wages*								16/0.77	
Labour costs per unit product*		7/0.92	12/0.63						
Composite leading indicator	6/0.81	9/0.66	3/0.81	7/0.78	6/0.84	7/0.77	7/0.68	5/0.87	9/0.83

Notes: The first number in each cell is the optimal fine-tuned lead (9) and the second number is the maximum correlation (8) of the corresponding leading indicator variable with the deviation cycle indicator of the manufacturing industry, *DCI*. The variables marked with a '*' counter-cyclically lead the cycle instead of pro-cyclically. The composite leading indicator, *CLI*, is the average of the synchronized and standardized basic leading indicator variables.

The aim is to transform the selected basic leading indicators into a single composite index which replicates and predicts the deviation cycle indicator of the manufacturing industry. The replication property is ensured by the minimum correlation requirement. The prediction property is ensured by the criterion of minimum lead. A single composite index is then constructed as an optimal linear combination of normalized and synchronized selected basic leading indicators. The normalization procedure makes indicators exhibiting weak cyclical variation, for instance wages, comparable to indicators exhibiting strong cyclical variation, for instance confidence indicators and prevents the latter ones from dominating the former ones in a single composite index. The selected basic leading indicators are synchronized with the deviation cycle indicator of the manufacturing industry by shifting them forward with their respective leads l^* that are reported in Table 4.5. Shifting a basic leading indicator variable forward by a number of months equal to its lead means that the shifted basic leading indicator variable no longer leads, but instead coincides with the cyclical phase of the manufacturing industry. Because of the shifting, the series extends beyond the observation period and provides a predictive signal for the near future cyclical development. Each selected basic leading indicator variable provides its own early signal. A one-dimensional single signal is obtained by constructing an optimal linear combination of the individual normalized and synchronized basic leading indicator variables by means of principal component analysis. The first principal component reflects the composite leading indicator since it captures the maximum common variation across the set of leading indicators. According to Moore and Shiskin (1967) formalized criteria, the composite leading indicator is not only supposed to reflect the motion of the deviation cycle indicator of the manufacturing industry, but also to provide an early signal of upcoming turning points. As a linear combination of synchronized basic variables, the composite leading indicator's turning points not necessarily coincide optimally with the corresponding ones of the deviation cycle indicator of the manufacturing industry. The lead of the composite leading indicator will therefore be fine tuned by matching the corresponding turning points of both indicators and minimizing a distance metric. The turning points of both indicators are dated using the procedures as stated in (1). The metric is the sum of the differences measured in months between all corresponding peaks and troughs of both indicators. The optimal leads of the synchronized composite leading indicators and their corresponding cross correlations with the deviation cycle indicators of the manufacturing industries are displayed in the last two rows of Table 4.5.

Table 4.6: Dating of turning points

	NLD		BEL		FRA		DEU		ITA
	DCI	CLI	DCI	CLI	DCI	CLI	DCI	CLI	DCI
Trough	67m6		67m5		68m4		67m6		69m11
Peak	70m1		69m8		69m6		69m12		70m8
Trough	72m4		71m11		71m9		72m1	71m7	71m10
Peak	74m3	74m4	74m4	73m12	74m2		73m11		74m1
Trough	75m7	75m12	75m7	76m4	75m7		75m6	75m8	75m8
Peak	76m8		76m9		76m10	76m10		77m1	76m11
Trough	77m7		77m8		77m9	77m10		78m2	77m12
Peak	79m11	79m11	79m11	80m7	79m11	79m12	79m12	80m1	80m3
Trough			81m2						
Peak			81m12						
Trough	83m1	83m4	82m11			80m4	82m12	83m1	83m3
Peak	84m11	84m9	84m1			82m6	85m9	84m7	
Trough	87m8	86m12	86m12	86m3	85m4	85m6	87m11	88m3	
Peak	90m7	90m1	89m5	89m6	89m6	89m8	91m9	91m1	89m7
Trough	93m8	93m11	93m1	92m4	93m7	93m8	93m8	93m10	93m6
Peak	95m1	95m7	95m7	95m7	95m2	95m3		95m7	95m8
Trough	97m1	96m9	96m6	96m6	96m9	96m10		96m11	96m11
Peak	98m1	98m5	98m1	98m4	97m12	98m6	98m4	98m8	97m12
Trough	99m2	99m9	99m4			99m7	99m5	99m11	99m3
Peak	00m8	00m11	00m8			00m9	00m10	01m1	00m12
Trough	03m9	02m5		02m4	03m5	02m3		02m6	
False		0		0		4		4	
Missed		2		8		0		0	
mean		4/14		1/9		1/10		2/10	
stdv		6.4		6.5		5.2		7	

Notes. *DCI* = deviation cycle indicator of the manufacturing industry; *CLI* = composite leading indicator. The absolute mean and standard deviation refer to the differences in timing of turning points in months between *DCI* and the *CLI*. A false signal occurs if the *CLI* shows a cycle and the *DCI* does not. A missed turning point occurs if the *DCI* shows a cycle and the *CLI* does not.

Moreover, both indicators of each country are graphed in Figure A.4.2 in the appendix. The shaded areas in the graphs correspond to the below-trend growth periods. Table 4.6 presents the dating of the cyclical turning points of both indicators for all countries. The composite leading indicator sometimes

Table continued									
	ITA	ESP	JPN			GBR		USA	
	CLI	DCI	CLI	DCI	CLI	DCI	CLI	DCI	CLI
Trough		68m03				68m07		67m07	
Peak		69m07	70m03	70m03		69m08		69m07	
Trough		71m07	71m11	71m12		72m01		70m11	
Peak	73m12	74m02	73m07	73m10	74m5	73m05		74m01	73m04
Trough	75m9	75m10	75m10	75m06	76m4	75m10		75m07	75m11
Peak	76m10		77m06						
Trough	77m11		78m07						
Peak	79m12	79m12	79m10	80m01	80m6	79m09	79m09	79m05	79m05
Trough						81m02	81m04	80m08	
Peak								81m04	
Trough	83m5	82m05	81m04	83m01	82m5			82m11	83m03
Peak		83m05	83m01	84m12		85m02	84m06	84m06	84m08
Trough		85m05	85m07	87m02		86m06	86m10	86m08	85m12
Peak	89m6	89m09	89m09	91m02	89m6	89m01	88m12	88m07	88m02
Trough	93m7	93m04	93m10	93m11	93m4	91m10	91m09	91m07	91m12
Peak	95m3	95m03	95m06			95m01	95m06	95m01	95m01
Trough	96m12	96m06	96m11				96m09	96m04	
Peak	98m6	98m04	98m06	97m06	97m9		98m01		
Trough	99m7		99m09	98m11		99m03	99m06		98m04
Peak	00m8		00m08	00m10		00m12	00m08	00m05	00m03
Trough	02m5	02m02	02m02	02m01	02m9	02m07	02m05	01m12	01m12
False	0		4		0		2		0
Missed	0		0		4		0		2
mean	3/14		4/14		2/8		2/10		3/11
stdv	3		5.3		10.4		3.9		4

Notes. *DCI* = deviation cycle indicator of the manufacturing industry; *CLI* = composite leading indicator. The absolute mean and standard deviation refer to the differences in timing of turning points in months between *DCI* and the *CLI*. A false signal occurs if the *CLI* shows a cycle and the *DCI* does not. A missed turning point occurs if the *DCI* shows a cycle and the *CLI* does not.

misses turning points or sometimes gives false alarms, which means that the leading indicator signals a turning point that is not present in the deviation cycle indicator of the manufacturing industry. Missed signals and false alarms are mostly connected to minor cycles.

The mean and the standard deviation of the differences in months between the timing of the corresponding turning points of the *DCI* and *CLI* are presented in the last rows of Table 4.6. The absolute value of the mean difference is by construction smaller than a half due to the fine tuning, which optimally matches the corresponding turning points of both indicators. While on average the match is optimal for each country, the standard deviation of the differences reveals the variation in the lead/lag time for the turning point dates. The standard deviation is for most countries of the same order of magnitude as the lead of the *CLI*. The lead time of the *CLI* lays therefore grosso modo just outside the one standard deviation confidence interval of the difference in months between the corresponding turning points of both indicators. Therefore, more or less 85% of all turning points are signalled by the *CLI* during its lead time, so before the *DCI* does. Only Italy is a special case, the lead of the *CLI* is larger than two times the standard deviation. The *CLI* is signalling almost all turning points before the *DCI* does. Moreover, the Italian *CLI* never missed a turning point or gave a false alarm.

4.6 Conclusion

The cycles in the manufacturing industry of nine OECD-countries are identified as deviations from trend. We use the Christiano-Fitzgerald band-pass filter to estimate the unobservable trend and the cyclical component of a time series. We adopt the convention that the business cycle frequencies consist of all cycles with duration longer than 18 months and shorter than 10 years. From the deviation cycle in the manufacturing industry, we derive the cyclical turning points, low- and high-growth periods and summary statistics describing features like amplitude, steepness and duration of the cycle for each country.

The manufacturing industries in the Netherlands and, to a lesser extent, Belgium stand out compared to the nine countries on stability by showing a relatively modest cyclical spread around the trend. Their cyclical motions are however the least moderate so that the manufacturing industries of both countries move quickly from low- to high-growth phases and vice versa. Japan acts as the mirror image by showing the largest cyclical swings. The U.S. reveals a pattern of smooth rounded peaks and pointed deep troughs. The hypothesis of no duration dependence is rejected for nearly all deviation cycles in the manufacturing industries for all countries in both high- and low-growth phases. The international linkage between the manufacturing industries is explored by calculating the fraction of time the two countries are both in the same phase. This statistic shows that manufacturing industry in the United

Kingdom is more synchronised with the one in the United States than with the one in the euro area. Moreover, the average synchronisation between the United Kingdom and the euro area countries is lower than the average synchronisation of the euro area countries with one another.

In addition to measuring cycles we constructed for each country a single composite leading indicator, which replicates and predicts the deviation cycle in the manufacturing industry. The composite leading indicator is based on economic, financial and survey variables possessing leading properties. These variables are selected from a set of candidate variables for each country. The variables that have been selected for five or more countries are the short term and long term interest rates, the storage of final products, the hourly wages, the domestic sales prices, the IFO-indicator for Germany and the consumption price index. The lead of the composite leading indicator is determined such that the dates of the turning points match most closely with the corresponding turning point dates of the deviation cycle indicator of the manufacturing industry. Moreover, the majority of the turning points are signalled by the composite leading indicator within its lead time.

4.A Appendix

4.A.1 A technical note on testing for duration dependence

This appendix derives a regression based test statistic for duration. The state variable S_t is obtained by applying the dating rules (4.1) to the indicators of the deviation cycle. Moreover, let d_i be the duration of the i^{th} phase in number of months. Consider the random sample of n observations $(D_1, D_2, \dots, D_n)$ of durations from a discrete distribution with density function $f(d)$. Then the hazard rate function is defined as,

$$h(d) = \frac{f(d)}{1 - F(d)}.$$

For a small Δ , $h(d)\Delta$ is the probability that the phase will terminate during the interval $(d, d + \Delta)$. If there is to be no duration dependence, then the termination probability must be constant, regardless of the duration of time spent in the phase and so does not depend on d , that is

$$H_0 : h(d) = \theta, \text{ for some } \theta > 0 \text{ and all } d > 0. \quad (4.9)$$

In the discrete case, the null hypothesis of no duration dependence (4.9) can only be satisfied by the geometric distribution. The geometric density of D with termination probability p reads as

$$P(D = d) = (1 - p)^d p \text{ for } 0 \leq d < \infty,$$

and its constant hazard function can be derived as:

$$h(d) = \frac{P(D = d)}{1 - P(D \leq d)} = \frac{(1 - p)^d p}{(1 - p)^d} = p.$$

Moreover, it holds in the geometric case that $E(D) = (1 - p)/p$, $Var(D) = (1 - p)/p^2$ and so $Var(D) - [E(D)]^2 - E(D) = 0$.

So, the null hypothesis (4.9) of no duration dependence implies a testable moment condition. However, in this study a minimum duration on the complete cycle of 18 months is imposed by the parameter settings of the CF-filter. So, effectively the distribution of D becomes a delayed geometric one $P(D^{delayed} = D - c)$, or equivalently :

$$P(D^{delayed} = k) = \begin{cases} 0 & \text{otherwise} \\ p(1 - p)^{k-c} & \text{for } k = c, c + 1, \dots \end{cases}.$$

And it holds that

$$E(D^{delayed}) = \frac{(1-p)}{p} + c, \text{ Var}(D^{delayed}) = \frac{(1-p)}{p} = \text{Var}(D),$$

and the testable moment restriction for the delayed geometric distribution reads as:

$$\text{Var}(D) - [E(D)]^2 - (2c+1)E(D) - c(c+1) = 0. \quad (4.10)$$

A regression based test statistic for duration dependence, which is of delayed geometrically distributed durations (4.10), can be constructed as a variable-addition test in a regression equation. The test is a modification of Ohn *et al.* (2004) to allow durations to be delayed in addition to be geometrically distributed. Consider the following regression:

$$S_t = \beta_0 + \beta_1 S_{t-1} + \beta_2 S_{t-1} X_{t-1} + \text{error}, \quad (4.11)$$

where the number of consecutive months spent in a high-growth phase up and through time t is given by:

$$X_t = 1 + (1 - S_t(1 - S_{t-1}))(1 - S_{t-1}(1 - S_t))X_{t-1}.$$

It is possible to derive an analytic form for the estimate of β_2 which is identical to the moment restriction (4.10). Consider:

$$\frac{1}{T} \sum_{t=1}^T S_{t-1} S_t = \bar{S} - \frac{n}{T},$$

where $\bar{S}$ is the sample mean of S_t , T is the sample size and n is the number of TP-phases. Moreover, $\overline{SX} = \frac{1}{T} \sum_{t=1}^T S_t X_t = \frac{1}{T} \sum_{i=1}^n \frac{(d_i^{delayed} + 1) d_i^{delayed}}{2} = \frac{1}{T} \sum_{i=1}^n \frac{(d_i + 1 + c_m)(d_i + c_m)}{2} = \frac{n}{2T} \left[\bar{d}^2 + (2c_m + 1)\bar{d} + (c_m^2 + c_m) \right]$, where d_i is the duration of the i^{th} high-growth phase¹⁰ minus the minimum of c_m periods. Finally,

$$\frac{1}{T} \sum_{t=1}^T S_t S_{t-1} X_{t-1} = \frac{1}{T} \sum_{i=1}^n \frac{(d_i - 1 + c_m)(d_i + c_m)}{2} = \overline{SX} - \bar{S}.$$

Now, the standard linear regression of $S_t - \bar{S}$ on $S_{t-1} - \bar{S}$ and $S_{t-1} X_{t-1} - \overline{SX}$, with Δ the determinant of the scaled (by $1/T$) cross product matrix of the

¹⁰The analysis can be performed for low-growth (PT-) phases as well by defining $S^{\text{low-growth}} = 1 - S^{\text{high-growth}}$.

two regressors, results in:

$$\begin{aligned}
 \Delta \hat{\beta}_2 &= \bar{S} (1 - \bar{S}) \left(\overline{SX} - \bar{S} + \frac{n}{T} c_m - \overline{SX} \bar{S} \right) - \overline{SX} (1 - \bar{S}) \left(\overline{SX} - \bar{S} - \frac{n}{T} \right) \\
 &\Leftrightarrow \frac{\Delta \hat{\beta}_2}{n (1 - \bar{S})} = \overline{SX} + \bar{S} \left(c_m - \frac{T}{n} \bar{S} \right) = \frac{1}{T} \sum_{i=1}^n \frac{(d_i + 1 + c_m) (d_i + c_m)}{2} \\
 &\quad + \frac{n (\bar{d} + c_m)}{T} \left(c_m - \frac{n}{T} \frac{T}{n} (\bar{d} + c_m) \right) \\
 &\Leftrightarrow \frac{2T^2 \Delta \hat{\beta}_2}{n^2 (1 - \bar{S})} = \bar{d}^2 - 2\bar{d}^2 - (2c_m - 1) \bar{d} - c_m (c_m - 1) \\
 &= Var(D) - [E(D)]^2 - (2c_m - 1) E(D) - c_m (c_m - 1). \quad (4.12)
 \end{aligned}$$

Note that for $c_m := c_m + 1$ equations (4.10) and (4.12) are equivalent. The variable addition test statistic $H_0 : \beta_2 = 0$ for equation (4.11) is therefore also effectively testing for no delayed duration dependence. Because of the binary nature of the variables involved one needs to consider the heteroskedastic and autocorrelation consistent t -statistic.

Figure A.4.2: International deviation cycles in the manufacturing industry and composite leading indicators

Table A.4.7: Data source

Description	Code	Source*
The Netherlands		
Industrial production	IRAA04MB	CBS
Consumer price index	IRDE31MA	CBS
Short term interest rate	IKGB01MA	DNB-FM
Long term interest rate	IKGE11MA	DS
IFO, expected business situation	IJDEFDMA	DS
Hourly wage industry	IRDK11MA	CBS
Storage final products	IJNLFAMA	DS
Expected business activity	IRAD41MA	CBS
Belgium		
Industrial production	IIBEAXMB	BIS
Consumer price index	IIBEBAMA	BIS
Short term interest rate	IKGB01MA	DNB-FM
Long term interest rate	IIBELDMA	DS
IFO, expected business situation	DS	
Hourly wage industry	IJBFEFEMA	DNB S&I*
Domestic sales prices	IIBEBCMA	BIS
Labour costs per unit product	IJBEFFMA	DNB S&I*
France		
Industrial production	IJFRFLMH	DS
Consumer confidence	IIFRAMMB	DS
Producer confidence	IIFRAQMB	DS
Equity price	IIFRLXMA	BIS
Labour costs per unit product	IJFRFFMA	DNB S&I
Judgement of order arrivals	IJFRFBMA	DS
Storage final products	IJFRFAMA	DS
Level order position	IJFRFCMA	DS

table (continued)

Description	Code	Source*
Germany		
Industrial production	IJDEFLMH	DNB S&I
Consumer confidence	IIDEAMMB	DS
Short term interest rate	IKGB01MA	DNB-FM
IFO, expected business situation	IJDEFDMA	DS
Hourly wage industry	IIDEADMB	BIS
Judgement of order arrivals	IJDEFBMA	DS
Storage final products	IJDEFAMA	DS
Level order position	IJDEFCMA	DS
Italy		
Industrial production	IJITFLMH	DNB S&I
Consumer price index	IIITBAMA	DS
Short term interest rate	IKGB01MA	DNB-FM
Long term interest rate	IIITLDMA	DS
IFO, expected business situation	IJDEFDMA	DS
Hourly wage industry	IIITADMA	BIS
Domestic sales prices	IIITBCMA	BIS
Producer confidence	IIITAQMB	DS
Prospects total economy	IJITFEMA	DS
Judgement of order arrivals	IJITFBMA	DS
Level order position	IJITFCMA	DS
World market prices commodities	IRDF01MA	HWWA
Spain		
Industrial production	IJESFLMH	DNB S&I
M1	IJESFFMA	DNB S&I
IFO, expected business situation	IJDEFDMA	DS
Storage final products	IJESFAMA	DS
Level order position	IJESFCMA	DS
World market prices commodities	IRDF01MA	HWWA

Table (continued)

Description	Code	Source*
Japan		
Industrial production	IJPAXMB	BIS
Consumer price index	IJPBAMA	BIS
Short term interest rate	IJPLBMA	BIS
Long term interest rate	IJPLDMA	BIS
Consumer confidence	IJJPFBMA	DNB S&I
Hourly wage industry	IJJPADMA	BIS
Domestic sales prices	IJPBCMA	BIS
Storage final products	IJJPFAMA	DS
Consumption	IJJPFCMA	DNB S&I
United Kingdom		
Industrial production	IIGBAXMB	DS
Consumer confidence	IIGBAMMB	DS
Producer confidence	IIGBAQMB	DS
Consumer price index	IIGBBAMA	DS
Short term interest rate	IIGBLBMA	Financial Times
Domestic sales prices	IIGBBCMA	BIS
Total wages	IJGBFFMA	OECD
Expected future production	IJGBFDMA	DS
Storage final products	IJGBFAMA	DS
Level order position	IJGBFCMA	DS
Prospects total economy	IJGBFEMA	DS
United States		
Industrial production	IIUSAXMB	BIS
World market prices commodities	IRDF01MA	HWWA
Storage final products	IJUSFAMA	DS
Hourly wage industry	IIUSADMB	BIS
Equity price (S&P corporate 500)	IIUSLXMA	BIS
Short term interest rate	IIUSLBMA	BIS
Long term interest rate	IIUSLDMA	BIS
Domestic sales prices	IIUSBCMA	BIS

Notes: CBS = Statistics Netherlands, DNB-FM = De Nederlandsche Bank - division Financial Markets, DNB S&I = De Nederlandsche Bank - division Statistics & Information, DS = Datastream, HWWA = Institut für Wirtschaftsforschung.

Chapter 5

Forecasting Dutch GDP using Alternative Factor Model Specifications Based on Large and Small Datasets

Business cycle modeling without pretending to have too much a priori economic theory.

Thomas J. Sargent and Christopher A. Sims

Abstract

We compare the factor forecasting performance of nested specifications of the generalized factor model based on various configurations of a large macroeconomic data set. The forecast simulation design involves in-sample model selection, factor estimation, parameter estimation and, finally, generating factor forecasts and factor augmented autoregressive forecasts. In order to empirically determine the importance of the size and the structure of the data set, we run the forecast simulation design for different configurations of the data set. We compare the factor model diagnostics of each specification and data configuration with the corresponding forecast performance. The results favour the factor structure as the specification that imposes the factor structure to the least extent and, hence, is allowed most flexibility to adapt to the data is significantly being outperformed. Moreover, the results show that size matters as though smaller macroeconomic data sets exhibit stronger coherence, the factors being well fit do, however, generally not show improved forecasting performance.

The author would like to thank Peter Kugler for his introduction of factor models at the workshop of the Oesterreichischen Nationalbank, seminar participants at CCSO of the University of Groningen, Deutsche Bundesbank, De Nederlandsche Bank, Annual meeting of the European Economic Association 2005, Sveriges Riksbanken, Bertrand Melenberg, Jan Jacobs, Lex Hoogduin, Massimiliano Marcellino, Franz Palm, Peter Vlaar, Bas Werker and an anonymous referee for comments and suggestions, Christophe van Nieuwenhuyze for a careful reading of an earlier version and, finally, Claudia Kerkhoff for extensive statistical assistance. The dynamic factors are estimated using code from <http://www.dynfactors.org>

5.1 Introduction

Empirical research on forecasting macroeconomic key variables aims to provide fiscal and monetary policymakers with the most accurate predictions. The univariate and low order vector autoregressive (VAR) models have for a long time been the standard small-scale models for short term macroeconomic forecasting. These models include only a small number of variables while policymakers and applied forecasters are keen to extract information from many more series describing economic activity at a more disaggregate level. The increase in the quantity and quality of readily available economic data stimulates a macroeconometric literature that explicitly incorporates information from a large number of macroeconomic variables into formal statistical models. For example, Garcia-Ferrer et al. (1987) apply pooling techniques to establish a relationship between annual output growth and leading indicators such as real stock return and real money supply growth using a multi-country data set (see also Hoogstrate et al., 2000).

As an alternative strategy to handle large data sets, Bates and Granger (1969) propose to combine the forecasts of many low-order equations exclusively employing one of the available predictors. Moreover, Palm and Zellner (1992) relate to the relative merit between combination and selection to attain optimal forecasts.

The study adopts the notion that the essential characteristics of macroeconomic motions are captured by a few driving aggregate forces and that the information contained in all potentially available economic key variables at an aggregate level are individually less informative about macroeconomic behaviour. The related empirical literature suggest that factor-based forecasts tend to outperform small-scale rival models, although the evidence is not overwhelming, see Eickmeier and Ziegler (2008) for an overview of the empirical macroeconomic factor forecasting literature and for instance Rünstler et al. (2009) for a comparison over data sets for different European countries. Moreover, Boivin and Ng (2006) show that enlarging a big data set not necessarily improves the factor forecasting performance if the additional series are noisy or unrelated to the target variable.

This paper applies the static factor model proposed by Stock and Watson (2002a) and its dynamic equivalent of Forni et al. (2000) to Dutch quarterly data with the aim of forecasting the growth rates of Gross Domestic Product (GDP) for an horizon up to 4 quarters ahead. The data are a subset of the series underlying the Dutch central bank's macroeconomic structural model for the Netherlands (cf. Fase et al., 1992) supplemented with leading indicator variables. The data set consists of 124 series that can be classified into six

categories. In order to empirically determine the importance of the size and the structure of the data set, we generate forecasts for different configurations of the data set. We determine the factor model diagnostics of each specification and data configuration and aim to establish a relationship with the forecasting performance.

This paper is organised as follows. Section 5.2 introduces the factor model and shows the specifications for which the cyclical dynamic factor collapses to a static one. Section 5.3 describes the real-time forecast simulation design, the diagnostics of the factor model fit and the data set. Finally, section 5.4 reports the empirical results, documents the diagnostics of the model specifications and data configurations and, eventually, shows the best performing outcomes.

5.2 The factor model

Factor models are a tool to cope with many variables without running into problems of too little degrees of freedom often faced in regression based analysis. The factor structure decomposes each variable into a common and an orthogonal idiosyncratic component. The common component is a linear combination of the common factors and therefore strongly correlated across the panel. The idiosyncratic component contains the remaining variable specific information and is only weakly correlated across the panel. The different types of factor models contrast according to the properties of the common and the idiosyncratic components. The generalized static factor model of Stock and Watson(2002a; 2002b) relaxes the classical factor model's assumption of the idiosyncratic components being mutual orthogonal and, instead, allows for some mild, but restricted, cross- and autocorrelation. The generalized dynamic factor model of Forni et al. (2000; 2001; 2001a; 2004; 2005) allows for a dynamic relationship between the common factors and the individual time series. Modelling the dynamics explicitly is in principle a desirable, since macroeconomic variables in general are non-synchronized and leading indicators should play a crucial role in a forecasting context.

5.2.1 Factor model representation

Consider a stationary stochastic vector process $\{\mathbf{x}_t = (x_{1t} \dots x_{nt})'\}$ with zero mean and finite second-order moments $\Gamma_{\mathbf{X}}(k) = E[\mathbf{x}_t \mathbf{x}_{t-k}']$. Each variable x_i , $i = 1 \dots n$ can be decomposed as the sum of two mutually orthogonal unobserved components: the common component χ_i and the idiosyncratic component ξ_i . The common components depend on a q -dimensional orthonormal

white noise process $\mathbf{f}_t = (f_{1t} \dots f_{qt})'$ driven by a small number of q dynamic factors f_{it} with $q \ll n$. The factor model reads as

$$\mathbf{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \mathbf{B}_n(L) \mathbf{f}_t + \boldsymbol{\xi}_t, \quad (5.1)$$

where the dynamic loadings are represented by a $(n \times q)$ -polynomial of order m : $\mathbf{B}_n(L) = (b_{i1}(L), \dots, b_{iq}(L))' = \mathbf{B}_{-m}^n L^{-m} + \dots + \mathbf{B}_0^n + \mathbf{B}_m^n L^m$ with the lag operator $L^s x_t = x_{t-s}$. The factors and idiosyncratic disturbances are assumed to be uncorrelated at all leads and lags, that is $E[\mathbf{f}_t \boldsymbol{\xi}_{il}] = 0 \forall i, l$. Clearly, if we let $\mathbf{F}_t = (\mathbf{f}_{t-m}' \dots \mathbf{f}_t' \dots \mathbf{f}_{t+m}')'$ and $\boldsymbol{\Lambda}_n = (\mathbf{B}_{-m}^n \dots \mathbf{B}_0^n \dots \mathbf{B}_m^n)$, the dynamic factors obey the static representation¹: $\boldsymbol{\chi}_t = \mathbf{B}_n(L) \mathbf{f}_t = \boldsymbol{\Lambda}_n \mathbf{F}_t$. A model with q dynamic factors $\mathbf{f}_t$ thus consists of $r = q(2m+1)$ common static factors $\mathbf{F}_t$. The dynamic nature of (5.1) implies that $\mathbf{F}_t$ has a special structure: if $m > 0$ the rank of the spectral density matrix of $\mathbf{F}_t$ (namely q) is smaller than the rank of the covariance matrix of $\mathbf{F}_t$ (namely r).

5.2.2 Estimating the factors

Denote by $\mathbf{X}^{nT} = (x_{it})_{i=1 \dots n, t=1 \dots T}$ an $n \times T$ rectangular array of standardized observations for $\mathbf{x}_t$. Let $\hat{\Gamma}_{\mathbf{X}}^{nT}(k) = \frac{1}{T-k} \sum_{t=k+1}^T \mathbf{x}_t^{nT} \mathbf{x}_{t-k}^{nT'}$ be the k -lag sample covariance of $\mathbf{X}^{nT}$. Moreover, let $\hat{\Gamma}_{\mathbf{XY}}^{nT}(k) = \frac{1}{T-k} \sum_{t=k+1}^T \mathbf{y}_t^{nT} \mathbf{x}_{t-k}^{nT'}$, with $\mathbf{y}$ the non-standardized correspondent of $\mathbf{x}$. The common dynamic factors $\mathbf{f}_t$ or static factors $\mathbf{F}_t$ are latent and need to be estimated.

The static method

Let $\hat{\mathbf{S}}^{nT} = (\hat{\mathbf{S}}_1^{nT'} \dots \hat{\mathbf{S}}_r^{nT'})'$ be the $(r \times n)$ -matrix of static eigenvectors, which correspond to the r largest eigenvalues, of the sample contemporaneous correlation matrix $\hat{\Gamma}_{\mathbf{X}}^{nT}(0)$. The static principal components estimator reads in matrix notation as $\hat{\mathbf{F}}^{nT} = \hat{\mathbf{S}}^{nT} \mathbf{X}^{nT}$, where $\hat{\mathbf{F}}^{nT} = (\hat{\mathbf{F}}_1^{nT} \dots \hat{\mathbf{F}}_T^{nT})$ is the $(r \times T)$ -matrix of stacked estimated static factors. Moreover, orthogonal projection of the data on the factors yields the factor loading estimates $\hat{\boldsymbol{\Lambda}}_n = \hat{\mathbf{S}}^{nT}$ and the common component $\hat{\boldsymbol{\chi}}_t^{nT} = \hat{\mathbf{S}}^{nT'} \hat{\mathbf{S}}^{nT} \mathbf{x}_t^{nT}$. The static method requires only the specification of the number of static factors r . Bai and Ng's (2002) information criteria (BNIC) determine r as a trade-off between goodness-of-fit and an overfitting penalty that increases with sample size n and time series length T .

¹Moreover, the orthonormality assumption for the common dynamic factors $\mathbf{f}_t$ is effectively an identifying assumption. Consider a nonsingular $(q \times q)$ -lag polynomial $\mathbf{A}(L)$ of order $M \leq m+1$. Then, $\mathbf{B}_n(L) \mathbf{f}_t = \mathbf{B}_n(L) [\mathbf{A}(L)]^{-1} \mathbf{A}(L) \mathbf{f}_t = \tilde{\mathbf{B}}_n(L) \tilde{\mathbf{f}}_t$ shows that the factors and factor loadings can only be identified up to a rotation.

The dynamic method

The dynamic method takes explicitly the dynamic structure of the data into account as represented in (5.1) by the two-sided filter $\mathbf{B}_n(L)$. We can further decompose the common component χ_{it} into a cyclical medium- and long-run component ϕ_{it} and a non-cyclical seasonal and irregular part ψ_{it} , that is

$$x_{it} = \phi_{it} + \psi_{it} + \xi_{it}. \quad (5.2)$$

This decomposition is based on a two-sided, symmetric, square summable bandpass filter $\beta(L)$, which separates waves of periodicity larger than a given critical number of periods τ :

$$\phi_{ij,t} = \sum_{k=-\infty}^{\infty} \beta_k \chi_{ij,t-k}, \quad \beta_k = \begin{cases} \frac{1}{k\pi} \sin(2k\pi/\tau) & \text{for } k \neq 0, \\ 1/\tau & \text{for } k = 0. \end{cases} \quad (5.3)$$

The cyclical medium- and long-run component $\phi_{ij,t}$ is thereby filtered for short-run seasonal and erratic fluctuations and therefore signals more smoothly the underlying developments of the staffing employment growth².

In order to estimate the generalized dynamic factor model, we need to specify the number of dynamic factors q , the parameter m that determines the maximum lag of the auto-covariance matrix and the cyclicity parameter τ , see appendix 5.A.1 for details. The identifying factor model assumption requires that the q largest dynamic eigenvalues diverge, whereas the remaining $n - q$ eigenvalues remain bounded as the number of time series variables n increases. We follow Forni et al.'s (2000) approach and select $q = 3$ in the finite-sample as the marginal explained variance of the q^{th} dynamic eigenvalue is larger than 10% and the $(q + 1)^{\text{th}}$ one is smaller than 10%.

The corresponding q dynamic eigenvectors are the estimators for the common dynamic factors f_k , $k = 1, \dots, q$ and the dynamic factor loadings $b_{ik}(L)$ describe the dynamic impact of the k -th common factor f_k on the time series variable x_{it} . We use a data dependent rule to set the maximum lead and lag of m periods, that is $b_{ik,\pm n} L^{\pm n} f_{kt} = 0$ for $n > m$, at $m(T) = \text{ROUND}\left(2T^{(1/2)}\right)$. Finally, we set $\tau = 4$, so all seasonality, which by definition entails a duration shorter than 1 year, or 4 quarters, is filtered out.

The dynamic estimation method consists of two-sided filters and cannot be applied at the end of the sample, which is the most important for forecasting purposes. By truncating the time filters, the performance of the estimator deteriorates as t approaches T .

²CEPR's coincident indicator of the euro area (Eurocoin) reposes on a similarly composed measure that captures the cyclical signal underlying short-lived oscillations, see Altissimo et al. (2006).

In a second estimation step, the generalized principal components $\mathbf{F}^{nT}$ are therefore determined as the contemporaneous averages of $\mathbf{X}^{nT}$ that minimize the ratio of the variance of the idiosyncratic to common component. More precisely, the factor estimates are $\tilde{\mathbf{F}}^{nT} = \tilde{\mathbf{S}}^{nT} \mathbf{X}^{nT}$, where $\tilde{\mathbf{S}}^{nT} = \left(\tilde{S}_1^{nT'} \dots \tilde{S}_r^{nT'} \right)'$ is the $(r \times n)$ -matrix containing r generalized eigenvectors of the couple of matrices³ $\left(\tilde{\Gamma}_X^{nT}(0), \tilde{\Gamma}_\xi^{nT}(0) \right)$ with normalization such that $\tilde{S}_i^{nT'} \tilde{\Gamma}_\xi^{nT}(0) \tilde{S}_j^{nT'} = 1$ if $i = j$ and zero otherwise⁴. The contemporaneous correlation matrices of the common component $\tilde{\Gamma}_X^{nT}(0)$ and the idiosyncratic component $\tilde{\Gamma}_\xi^{nT}(0)$ are obtained in the first step of the estimation procedure. Like in the static method, BNIC determines the number of r generalized eigenvectors.

Comparing the dynamic and the static method

Evidently, the static factors $\hat{\mathbf{F}}^{nT}$ obtain as a special case of the dynamic factors $\tilde{\mathbf{F}}^{nT}$ in case $m = 0$ and assuming in the second step that $\text{DIAG}\left(\hat{\Gamma}_\xi^{nT}(0)\right) = I_N$, i.e. the identity matrix. Computing the generalized principal components of $\mathbf{x}_{nt}$ is equivalent to computing the standard principal components of $\mathbf{y}_{nt} = H\mathbf{x}_{nt}$ with $\det(H) \neq 0$ and H such that $H\xi_{nt}\xi_{nt}'H'$ is the $n \times n$ -identity matrix. When the idiosyncratic variance-covariance matrix is diagonal, the normalization amounts to dividing x by the standard deviation of its idiosyncratic component.

5.2.3 Factor forecasting

The object of interest is the h -step ahead forecast of the stationary time series variable $y_{i,T+h|T}$, whose standardized correspondent is x_i with mean μ_i and standard deviation σ_i . The factor forecasts read as

$$y_{i,T+h|T} = \mu_{i|T} + \sigma_{i|T} X_{i,T+h|T} = \mu_{i|T} + \sigma_{i|T} \beta_h \Lambda_{n,i} \mathbf{f}_T.$$

As the parameters are not observed and, hence, need to be estimated, the equivalence in population of the different forecast specifications breaks down in sample, so $\sigma_{i|T} \widehat{\beta}_h \widehat{\Lambda}_{n,i} \hat{\mathbf{f}}_T \neq \left(\sigma_{i|T} \widehat{\beta}_1 \widehat{\Lambda}_{n,i} \right)^h \hat{\mathbf{f}}_T \neq \hat{\sigma}_{i|T} \hat{\beta}_h \hat{\Lambda}_{n,i} \hat{\mathbf{f}}_T$. Using the same horizon for estimating and forecasting can modify the potential impact of the model specification error (cf. Clements and Hendry, 1998). Therefore, we disregard the parameter estimates that result from h times iterated one step

³Given the factor model assumption of orthogonality between χ and ξ , then $\tilde{\mathbf{S}}^{nT}$ are also the generalized eigenvectors of the couple $\left(\tilde{\Gamma}_X^{nT}(0) + \tilde{\Gamma}_\xi^{nT}(0), \tilde{\Gamma}_\xi^{nT}(0) \right) = \left(\tilde{\Gamma}_X^{nT}(0), \tilde{\Gamma}_\xi^{nT}(0) \right)$, with eigenvalues $\lambda_j^{nT} + 1$.

⁴In practice, only the diagonal elements of $\tilde{\Gamma}_\xi^{nT}(0)$ are employed in this step.

ahead forecasts $\left(\sigma_{i|T} \widehat{\beta_1 \Lambda_{n,i}}\right)^h$. Moreover, as the stochastic process driving the factors is generally not known, potential misspecification of $\widehat{\beta_h}$ can also be avoided by determining $\sigma_{i|T} \widehat{\beta_h \Lambda_{n,i}}$ as one parameter.

Therefore, the unrestricted χ_y , respectively restricted factor forecasts χ_χ are obtained by projecting the h -step ahead observations $y_{i,t+h}$, respectively the common component $\chi_{i,t+h}^{nT}$ on the t -dated factors $\mathbf{f}_t^{nT}$. The corresponding h -step ahead factor forecasts of the common component of the i -th variable given T observations of n time series variables reads as:

$$\begin{aligned} \chi_{y_{i,T+h|T}}^{nT} &= \left[\widehat{\Gamma}_{\mathbf{X}\mathbf{Y}}^{nT}(h) \right]_i \mathbf{S}^{nT'} \left(\mathbf{S}^{nT} \widehat{\Gamma}_{\mathbf{X}}^{nT}(0) \mathbf{S}^{nT'} \right)^{-1} \mathbf{S}^{nT} \mathbf{X}^{nT} \\ \chi_{\chi_{i,T+h|T}}^{nT} &= \widehat{\mu}_{i|T} + \widehat{\sigma}_{i|T} \left[\Gamma_{\mathbf{X}}^{nT}(h) \right]_i \mathbf{S}^{nT'} \left(\mathbf{S}^{nT} \widehat{\Gamma}_{\mathbf{X}}^{nT}(0) \mathbf{S}^{nT'} \right)^{-1} \mathbf{S}^{nT} \mathbf{X}^{nT} \end{aligned} \quad (5.4)$$

with sample mean $\widehat{\mu}_{i|T}$ and standard deviation $\widehat{\sigma}_{i|T}$. The static and dynamic factor forecasts can be obtained by employing $\widehat{\mathbf{S}}^{nT}$ and $\widehat{\Gamma}_{\mathbf{X}}^{nT}(h)$, respectively $\widetilde{\mathbf{S}}^{nT}$ and $\widetilde{\Gamma}_{\mathbf{X}}^{nT}(h)$. The restricted factor forecasts χ_χ adhere stronger to the factor structure as the matrix $[\Gamma_{\mathbf{X}}(h)]_i$ is involved instead of the data driven matrix $\widehat{\Gamma}_{\mathbf{X}\mathbf{Y}}^{nT}(h)$. This latter matrix also contains the first μ and second moment σ . So, the unrestricted (dynamic) factor forecasts⁵ χ_y are obtained as:

$$\sigma_{i|T} \widehat{\beta_h \Lambda_{n,i}} \mathbf{f}_T + \widehat{\mu}_{i|T} = \left[\widehat{\Gamma}_{\mathbf{X}\mathbf{Y}}^{nT}(h) \right]_i \widetilde{\mathbf{S}}^{nT'} \left(\widetilde{\mathbf{S}}^{nT} \widehat{\Gamma}_{\mathbf{X}}^{nT}(0) \widetilde{\mathbf{S}}^{nT'} \right)^{-1} \widetilde{\mathbf{S}}^{nT} \mathbf{X}^{nT}.$$

The sampling error of the factor estimates enters the forecasts and might even dominate the information gain in the factors. The forecast error is then affected both by the estimation of the factors and by the equation relating the estimated factors to the target variable. Rewriting the forecast equation $x_{i,T+h|T} = \chi_{i,T+h|T} + \xi_{i,T+h|T} = \beta_h \Lambda_{n,i} \mathbf{f}_T + \rho_i(L) \xi_{i,T}$

$= \left(1 - \rho_i(L) L^h\right) \beta_h \Lambda_{n,i} \mathbf{f}_T + \rho_i(L) y_{i,T}$ shows that forecasting the components separately is equivalent to forecasting the sum plus one of the two components separately. Boivin and Ng's (2005) factor augmented autoregressive forecasts (FAAR) simply augment lags of the estimated factors to an autoregressive forecast equation of the non-standardized target variable:

$$\widehat{y}_{i,T+h|T} = \widehat{\mu}_{i,h} + \widehat{\theta}_{i,h}(L) \widehat{\mathbf{f}}_T + \widehat{\gamma}_{i,h}(L) x_{iT} \quad (5.5)$$

⁵In case of the unrestricted factor forecasts, the dynamic factors $\widetilde{\mathbf{S}}^{nT}$ are estimated by the generalized eigenvectors of the couple $\left(\widehat{\Gamma}_{\mathbf{X}}^{nT}(0), \widehat{\Gamma}_{\xi}^{nT}(0)\right)$ instead of the couple $\left(\widetilde{\Gamma}_{\mathbf{X}}^{nT}(0), \widetilde{\Gamma}_{\xi}^{nT}(0)\right)$. The two estimators are identical in population, but in sample only to the extent that the factor model assumptions are satisfied.

The parameters $\hat{\mu}_{i,h}$, $\hat{\theta}_{i,h}(L)$ and $\hat{\gamma}_{i,h}(L)$ are obtained by regressing $y_{i,t+h}$ on b_f lags of $\hat{\mathbf{f}}_t$ and b_y lags of y_{it} . A likewise forecast $\tilde{y}_{i,T+h|T}$ is obtained using the generalized dynamic factors $\tilde{\mathbf{f}}_T$. The lags are chosen by the Bayesian Information Criterion (BIC) out of $1 \leq b_f \leq 4$ and $0 \leq b_y \leq 4$. So, the smallest candidate model that BIC can produce includes a constant, a single contemporaneous factor and no autoregressive lags. The parameter $\hat{\mu}_{i,h}$ is the estimated mean of the stationary variable. Since $\hat{\theta}_{i,h}(L)$ is not constrained to equal $(1 - \hat{\rho}_i(L)L^h)\sigma_i\widehat{\beta_h\Lambda_{n,i}}$, $\hat{\mu}_{i,h}$ is not restricted to equal the sample mean $\hat{\mu}_{i|T}$ and no restriction is imposed on the coefficients of the b_f lags of $\hat{\mathbf{f}}_t$ and b_y lags of y_{it} , the FAAR specification (5.5) is allowed more flexibility to adapt to the data.

5.3 Forecasting Dutch GDP

5.3.1 Real-time forecast simulation design

The aim is to generate forecasts of GDP-growth rates for the Netherlands, $y_{i,T+h|T}$ over a forecast horizon of $h = 1, \dots, 4$ quarters ahead. The forecasting exercise involves in-sample model selection, factor estimation, parameter estimation and, finally, generating factor forecasts (5.4) and FAAR forecasts (5.5). The in-sample selection of the factor models and the specification of the forecast equation are performed according to the various information criteria as explained in the previous sections. The precise specifications are based on data that cover the first half of the sample of observations, which runs from 1980Q2 until 1991Q1, and consists of 46 observations for each time series variable. Given the selected factor model and the factor forecast specification, the forecast $\hat{x}_{1991Q1+h}^h$ is generated in the first round. In the subsequent iteration, the factors and the parameters of the selected factor model and forecast specification are reestimated and utilized to generate the forecast $\hat{x}_{1991Q2+h}^h$. The iteration repeats 46 times and results in the final forecast $\hat{x}_{2002Q4+h}^h$. The factors and the parameters of the selected specifications are iteratively reestimated based on a rolling window scheme, which better takes into account the presence of structural breaks in the data set than a recursive window scheme. Eickmeier and Ziegler's (2008) meta-analysis of the empirical factor forecasting literature finds that correcting for structural breaks by applying a rolling window scheme overcompensates the loss of data for the beginning of the sample period.

5.3.2 Factor model diagnostics

The extracted factors represent the underlying specific data set, which then should capture correctly the main forces that drive the variable of interest, in our case GDP. Oversampling refers to the situation in which the data are more informative about some factors than the others. Including more variables in an oversampled data set could result in more precise factor estimates, which do however not improve the forecasting performance for the variables that depend on the less dominant factors. Let the commonality ratio $R_i^2 = \sum_{t=1}^T x_{it}^2 / \sum_{t=1}^T x_{it}^2$ indicate the relative importance of the common component of variable i and let the average commonality ratio of a specific data set be $\bar{R}^2 = \sum_{i=1}^n R_i^2$. A below average commonality ratio for the variable of interest, $R_{GDP}^2 < \bar{R}^2$ is then an indication of an oversampled data set.

In absence of oversampling, the features of the data that improve the precision of the factor estimates relate to the importance and dispersion of the common component. The estimation precision improves when the common component, as measured by $\bar{R}^2$, is important, but deteriorate with a larger dispersion of the importance of the common component. The cross-sectional dispersion R_q^2 is measured by the difference between the R_i^2 in the 90th and the 10th percentile: $R_q^2 = R_{9N}^2 - R_{1N}^2$.

So, adding data with large idiosyncratic errors or weak factor loadings deteriorate the factor diagnostics. Moreover, Boivin and Ng's (2006) simulation results suggest that the forecasts and factor estimates are adversely affected by the amount of cross-correlation among the idiosyncratic components. As shown in paragraph 5.2.2, the dynamic method downweights noisy data by the variance of its idiosyncratic component and reduces serial correlation by synchronizing the data. The dynamic method is therefore better capable to accommodate the properties of the data, which comes at a cost of a more elaborate estimation procedure.

5.3.3 Dutch data

The data set provides a balanced representation of the Dutch economy and of the forces it is exposed to. For this purpose, the data set of the macroeconomic model MORKMON of De Nederlandsche Bank for the Dutch economy (cf Fase et al., 1992) is supplemented with variables potentially possessing valuable information from a forecasting perspective. The data cover the Dutch national accounts on the expenditure components of GDP and describe the behaviour of the macro actors in the economy: households, firms, monetary financial sector, government and foreign sector. The data set is screened on

variables that are only available at a yearly frequency, especially related to the government sector, social security and the flow of funds like tax funds, insurance and pension premiums. The data set is supplemented with a more detailed extension of macro-wide variables and leading indicators.

The final data set consists of 124 time series variables, which can be divided into six different categories. The first category labeled 'GDP' consist of GDP, its expenditure components, labour market variables, real wages and the housing market. The second category labeled 'industrial production' consists of sectorally disaggregated time series on manufacturing turnover and capacity utilization. The third category labeled 'prices' consists of consumer, producer and commodity prices. The fourth category labeled 'financial' covers the financial developments captured by interest rates, exchange rates and the stock market. The fifth category labeled 'external' represents the external sector as recorded on the balance of payments in variables such as income transfers, direct and portfolio investment. The sixth and final category labeled 'surveys' consists of business expectations, assessments of stocks and order arrivals and confidence indicators.

The details of the data including the preprocessing are explained in appendix 5.A.2. The preprocessing includes outlier detection, removing seasonality and stationarity inducing transformations, rendering standardized time series variables of quarterly frequency. The entire data set was collected in the second quarter of 2004 and consists of the fully revised historical series available as of this date. The collected data set is the 2004Q2 snapshot of the variables and in this regard the forecasting results will be different from the results using real-time data.

5.4 Empirical results

In order to empirically determine the importance of the size and structure of the data set, we run the real-time forecast simulation design for different configurations of the data set. Apart from the complete data set consisting of 124 variables, we perform the forecast simulation on each of the six groups separately and on the complete data set excluding consecutively each one of the six groups. The forecasting performance is summarized by the Relative Mean Squared forecast Error (ReMSE), which is the mean squared forecast error of the particular forecast specification divided by the MSE of the AR(1)-process.

Table 5.1: Diagnostics for different factor model specifications

	h	$\hat{\chi}_\chi$	$\hat{\chi}_y$	$\tilde{\chi}_\chi$	$\tilde{\chi}_y$	$\tilde{\chi}_\chi^c$	$\tilde{\chi}_y^c$	$\bar{\chi}_\chi$	$\bar{\chi}_y$	$\bar{\chi}_\chi^c$	$\bar{\chi}_y^c$
R_{gdp}^2 (times 10)	1	1.80	1.80	0.86	1.30	0.74	1.27	2.41	2.41	1.33	1.77
	2	0.72	0.75	0.21	0.30	0.50	0.47	0.60	0.69	0.27	0.33
	3	0.20	0.25	0.26	0.22	0.26	0.29	0.21	0.33	0.24	0.25
	4	0.04	0.10	0.14	0.28	0.12	0.31	0.12	0.32	0.13	0.26
$\bar{R}^2$ (times 10)	1	1.00	1.00	0.76	1.00	0.66	0.99	2.44	2.44	1.88	2.34
	2	0.40	0.52	0.48	0.56	0.52	0.61	0.57	0.73	0.66	0.75
	3	0.11	0.25	0.23	0.30	0.27	0.36	0.30	0.46	0.35	0.46
	4	0.02	0.18	0.12	0.21	0.14	0.24	0.21	0.38	0.22	0.38
R_q^2	1	0.34	0.34	0.25	0.36	0.23	0.31	0.44	0.44	0.39	0.48
	2	0.14	0.10	0.15	0.17	0.19	0.21	0.13	0.14	0.15	0.17
	3	0.04	0.06	0.06	0.08	0.07	0.08	0.06	0.10	0.08	0.10
	4	0.01	0.05	0.03	0.05	0.03	0.06	0.04	0.08	0.05	0.08
ReMSE	1	0.84	0.85	0.87	0.87	0.89	0.89	0.90	0.90	0.92	0.93
	2	0.90	0.92	0.95	0.96	0.93	0.95	0.95	1.01	0.98	0.99
	3	0.96	0.93	0.99	1.01	1.00	1.03	0.96	1.02	1.02	1.05
	4	0.98	1.05	0.95	0.98	1.00	1.01	0.99	1.09	0.95	0.98

Notes.

$\hat{\chi}_y$ is the unrestricted static factor-forecast, $\hat{\chi}_\chi$ is the restricted static factor forecast, $\tilde{\chi}_y$ is the unrestricted dynamic factor forecast, $\tilde{\chi}_\chi$ is the restricted dynamic factor forecast, $\tilde{\chi}_y^c$ is the unrestricted cyclical dynamic factor-forecast, $\tilde{\chi}_\chi^c$ is the restricted cyclical dynamic factor forecast. The results of the forecast specifications are based on the complete data set of 124 variables. Moreover, the forecast specifications represented with a $\bar{\chi}$ present the average results over the 13 different configurations of the data. R_{gdp}^2 and $\bar{R}^2$ indicate the relative importance of the common component of GDP respectively the average relative importance of the common component of all the variables in the data set. R_q^2 measures the cross-section dispersion of the common component. Finally, ReMSE is the mean squared forecast error relative to the mean squared forecast error of the AR(1) process.

The different factor forecast specifications (5.4) consist of the unrestricted factor forecasts χ_y versus the restricted ones χ_x . Moreover, the factor forecasts employ static factors $\hat{\chi}$, dynamic factors $\tilde{\chi}$ or cyclical dynamic factors $\tilde{\chi}^c$. Table 5.1 reports the factor model diagnostics and the forecasting performance for the different forecast specifications. The results in the first part of the table are based on the complete data set, while the results of the forecast specifications represented with a bar, $\bar{\chi}$ are averages over the 13 different configurations of the data.

The results in the table clearly shows that $\bar{R}^2 < \overline{\bar{R}}^2$ and $\text{ReMSE} < \overline{\text{ReMSE}}$ for all model specifications and forecast horizons h . Smaller macroeconomic data sets exhibit stronger coherence, which is captured by the factors explaining a larger part of the correlation between the variables. However, the higher explanatory power of the factors does not lead to improved forecasting performance. One explanation is that the better factor fit does not relate to the variable of interest as shown by $\bar{R}_{GDP}^2 < \overline{\bar{R}}^2$. Contrarily for the case of the complete data set, it particularly holds that $R_{GDP}^2 > \bar{R}^2$ for the first forecast horizon $h = 1$. Here, oversampling, or rather missampling, refers to the misrepresentation of a small data set exhibiting a strong factor structure.

Comparing the diagnostics of the different factor forecast specifications, the table clearly shows that the unrestricted factor forecasts χ_y exhibit a better fit than its restricted equivalent χ_x for the different specifications, horizons h and data set sizes. Comparing the diagnostics of employing static factors $\hat{\chi}$ as compared to dynamic factors $\tilde{\chi}$ does not reveal structural differences, excepting that the static method seems to exhibit a better factor fit and forecasting performance at horizon $h = 1$. Comparing the diagnostics of employing dynamic factors $\tilde{\chi}$ as compared to cyclical dynamic factors $\tilde{\chi}^c$ only shows a worse fit of the cyclical factors in case of small data sets. Although the diagnostics reveal differences between the various factor specifications, the differences do at first sight not lead to substantial divergence in the forecasting performance.

5.4.1 Forecast accuracy of various model specifications

In order to systematically analyse the forecasting performance of the different factor forecasts (5.4) and the FAAR forecasts (5.5), we apply Harvey et al.'s (1997) small-sample modification of Diebold and Mariano's (1995) (DM) test of equal forecast accuracy. Under the null hypothesis, the squared difference between the forecast errors of two competing models is not statistically different from zero. We report the (symmetric) p -values of the DM-statistic that reject the null hypothesis.

As an additional summary statistic for the relative forecast accuracy over

time, we follow Schumacher (2007) and pairwise count the number of time periods for which model A has a smaller squared forecast error than model B . The counted number of time periods as a fraction of the total time span for which forecasts are generated provides a summary statistic, denoted as $I_{A<B}(h)$, for each forecast horizon h . So, if $I_{A<B}(h) > 0.5$ then in more than half of the forecast occasions, model A manages to outperform model B . Note that if, at the complementary occasions, model B outperforms model A with much smaller forecast errors, then it holds simultaneously that $MSE_A > MSE_B$ and $I_{A<B}(h) > 0.5$. So, $I_{A<B}(h)$ is a complementary statistic to MSE .

Table 5.2: Forecast accuracy of different model specifications; one quarter ahead forecast horizon

	$\hat{\chi}_\chi$	$\hat{\chi}_x$	$\hat{y}$	$\tilde{\chi}_\chi$	$\tilde{\chi}_x$	$\tilde{\chi}_\chi^c$	$\tilde{\chi}_x^c$	$\tilde{y}$
$\hat{\chi}_\chi$		0.57	0.04	0.17	0.19	0.25	0.29	0.03
$\hat{\chi}_x$	0.45		0.05	0.51	0.58	0.52	0.53	0.05
$\hat{y}$	0.34	0.34		0.09	0.08	0.08	0.12	0.84
$\tilde{\chi}_\chi$	0.45	0.51	0.57		0.50	0.85	0.96	0.07
$\tilde{\chi}_x$	0.47	0.53	0.60	0.47		0.97	0.88	0.06
$\tilde{\chi}_\chi^c$	0.45	0.49	0.60	0.51	0.55		0.82	0.10
$\tilde{\chi}_x^c$	0.51	0.53	0.55	0.51	0.53	0.47		0.15
$\tilde{y}$	0.40	0.47	0.40	0.40	0.43	0.40	0.40	

Notes.

See Table 5.1 for the explanation of the different factor forecast specifications. The results are based on the complete data set consisting of 124 variables. Moreover, $\hat{y}$ and $\tilde{y}$ are the unrestricted static respectively dynamic factor augmented autoregressive forecast. The p -values of the (symmetric) DM-test of pairwise equal forecast accuracy are presented in *italics* in the upper part of the table. The summary statistics of forecast accuracy over time $I_{A<B}(h)$ are presented in the lower part of the table for which model A is represented in the row and model B in the column.

The p -values of the DM-statistic and the forecast accuracy statistic over time $I_{A<B}(h)$ for the different specifications are reported in Table 5.2 for horizon $h = 1$ and in Table 5.3 as averages over four forecasting horizons. The results are based on the complete data set consisting of 124 time series variables. The table show that the FAAR specifications y are clearly being outperformed at the first horizon $h = 1$ according to the DM-statistic. The results across the factor forecast specifications are less pronounced. According to the statistic of forecast accuracy over time $I_{A<B}(h)$, the tables show that the restricted cyclical dynamic factor forecast $\tilde{\chi}_\chi^c$ outperforms almost all other specifications at all forecast horizons. Moreover, the next best performing specification is the

Table 5.3: Forecast accuracy of different model specifications; average over forecast horizons

	$\hat{\chi}_\chi$	$\hat{\chi}_x$	$\hat{y}$	$\tilde{\chi}_\chi$	$\tilde{\chi}_x$	$\tilde{\chi}_\chi^c$	$\tilde{\chi}_x^c$	$\tilde{y}$
$\hat{\chi}_\chi$		0.41	0.34	0.41	0.43	0.56	0.56	0.34
$\hat{\chi}_x$	0.47		0.35	0.38	0.40	0.46	0.47	0.35
$\hat{y}$	0.45	0.47		0.30	0.31	0.23	0.31	0.45
$\tilde{\chi}_\chi$	0.48	0.51	0.56		0.40	0.42	0.57	0.23
$\tilde{\chi}_x$	0.47	0.49	0.53	0.45		0.37	0.48	0.23
$\tilde{\chi}_\chi^c$	0.51	0.53	0.56	0.54	0.55		0.53	0.23
$\tilde{\chi}_x^c$	0.50	0.52	0.56	0.52	0.54	0.46		0.28
$\tilde{y}$	0.46	0.48	0.50	0.43	0.46	0.44	0.44	

Notes.

See Table 5.2

restricted static factor forecasts $\hat{\chi}_\chi$. These results indicate that imposing the factor structure, both in consideration of summarizing the dynamic cyclical structure of the data and by imposing the factor structure on the forecast equation, improves upon the forecasting performance despite weaker factor model diagnostics. Moreover, the FAAR specifications, which of all the possible specifications imposes the factor structure to the least extent and therefore allows the forecast equation most flexibility to adapt to the data, shows worst forecasting performance.

5.4.2 Forecast accuracy of various data configurations

The factor model diagnostics and forecasting performance of the best performing dynamic $\tilde{\chi}_\chi^c$ and static $\hat{\chi}_\chi$ factor forecast specifications and the FAAR specifications $\tilde{y}$, respectively $\hat{y}$, for different configurations of the data set are presented in Table 5.4 for forecast horizon $h = 1$ and in Table 5.5 as averages over four forecast horizons. The thirteen different configurations of the data set consists of each of the six groups separately, the complete data set consisting of 124 variables and the complete data set excluding consecutively each one of the six groups. The results in Table 5.4 and Table 5.5 show that size matters as the complete data set of 124 series outperforms for all horizons and almost all specifications the data configurations consisting of each of the six individual groups, which amounts to a size of around 20 variables.

Table 5.4: Forecasting performance of different data sets: one quarter forecast horizon

Group	Dynamic R_{gdp}^2	Method $\bar{R}^2$	ReMSE $\hat{\chi}_\chi^2$	$\bar{y}$	Static R_{gdp}^2	Method $\bar{R}^2$	ReMSE $\hat{\chi}_\chi^2$	$\bar{y}$	Ranking $\hat{\chi}_\chi^2$	$\hat{\chi}_\chi$	$\bar{y}$	Forecast $\hat{\chi}_\chi^2$	accuracy $\bar{y}$	$\hat{\chi}_\chi$	$\bar{y}$
1	0.33	0.11	0.32	0.93	0.57	0.12	0.95	1.08	2	4	3	6	0.45	0.47	0.45
2	0.23	0.38	0.83	0.87	0.26	0.38	0.77	0.94	1	2	2	4	0.51	0.43	0.36
3	0.07	0.53	0.74	1.02	0.19	0.63	0.51	1.12	3	6	1	2	0.45	0.49	0.34
4	0.10	0.69	0.77	0.95	0.13	0.69	0.77	1.04	6	5	5	5	0.45	0.47	0.53
5	0.12	0.14	0.33	0.94	0.13	0.15	0.33	1.07	4	3	4	3	0.47	0.57	0.49
6	0.56	0.47	0.68	0.96	0.57	0.48	0.56	1.11	5	1	6	1	0.40	0.51	0.53
Total	0.13	0.10	0.36	0.88	0.18	0.10	0.34	0.95	2	5	2	3	0.34	0.45	0.45
Total -/-1	0.11	0.11	0.39	0.89	0.15	0.11	0.38	0.96	2	3	6	2	0.45	0.47	0.49
Total -/-2	0.09	0.09	0.27	0.90	0.18	0.09	0.32	0.81	6	3	6	2	0.45	0.47	0.49
Total -/-3	0.13	0.11	0.38	0.88	0.17	0.12	0.34	0.87	1	6	1	4	0.36	0.30	0.51
Total -/-4	0.11	0.11	0.44	0.86	0.17	0.11	0.37	0.83	4	1	4	6	0.62	0.45	0.51
Total -/-5	0.12	0.11	0.40	0.88	0.17	0.11	0.35	0.84	3	4	3	5	0.55	0.43	0.49
Total -/-6	0.19	0.09	0.50	0.86	0.26	0.09	0.54	0.82	5	2	5	1	0.57	0.45	0.49

Notes.

The results are based on forecast horizon $h = 1$. R_{gdp}^2 and $\bar{R}^2$ indicate the relative importance of the common component of GDP respectively the average relative importance of the common component in the data group. R_q^2 measures the cross-section dispersion of the common component across the variables in the respective data groupings. ReMSE is the mean squared forecast error relative to the AR(1)-process. The section

"Ranking" represents the ranking of the different configurations of the data set according to the ReMSE of the various forecasting methods. The section

"Forecast accuracy" refers to the accuracy of the forecasts $I_{A < B}(h)$, where A refers to the corresponding data configuration and B to the complete data set. $\hat{\chi}_\chi$ represents the restricted static factor-forecast, $\hat{\chi}_\chi^c$ is the restricted cyclical dynamic factor-forecast, $\tilde{y}$ represents the dynamic factor augmented autoregressive forecast and $\bar{y}$ represents the static factor augmented autoregressive forecast.

Table 5.5: Forecasting performance of different data sets: average over forecast horizons

Group	Dynamic	Method		RMSE		Static		Method		RMSE		Ranking		Forecast		accuracy			
		R^2_{GDP}	R^2	$R^2_{\hat{y}}$	$\hat{y}$	$\hat{X}_Y$	$\hat{y}$	R^2_{GDP}	R^2	$R^2_{\hat{y}}$	$\hat{y}$	$\hat{X}_Y$	$\hat{y}$	$\hat{X}_Y$	$\hat{y}$	$\hat{X}_X$	$\hat{y}$		
1	0.15	0.03	0.09	1.02	0.25	0.02	0.02	0.05	0.99	1.09	2	1	2	4	0.50	0.58	0.47	0.48	
2	0.42	0.06	0.12	1.02	0.42	0.07	0.17	0.17	1.03	1.05	1	3	3	2	0.58	0.48	0.43	0.46	
3	0.45	0.10	0.17	1.04	0.60	0.12	0.12	0.19	1.13	1.18	6	5	5	5	0.52	0.46	0.44	0.40	
4	0.53	0.09	0.13	1.03	0.52	0.09	0.09	0.13	1.23	1.24	5	6	6	6	0.46	0.39	0.35	0.41	
5	0.18	0.02	0.06	0.99	0.19	0.02	0.02	0.04	0.97	1.07	4	4	1	3	0.50	0.51	0.47	0.50	
6	0.67	0.11	0.21	1.02	0.66	0.10	0.10	0.20	1.10	1.02	3	2	4	1	0.46	0.53	0.48	0.51	
Total	0.01	0.03	0.09	0.94	0.02	0.03	0.03	0.06	0.95	1.00	0.98								
Total +/- 1	0.01	0.03	0.08	0.96	0.01	0.03	0.03	0.07	0.95	1.00		2	2	3	3	0.40	0.41	0.46	0.45
Total +/- 2	0.00	0.02	0.06	0.98	0.03	0.02	0.02	0.04	0.96	1.00		1	1	2	1	0.45	0.48	0.45	0.47
Total +/- 3	0.01	0.03	0.09	0.95	0.01	0.02	0.03	0.08	0.94	0.99		4	4	4	4	0.47	0.40	0.55	0.49
Total +/- 4	0.01	0.04	0.12	0.97	0.03	0.03	0.03	0.10	0.97	0.95		6	6	1	6	0.59	0.53	0.49	0.53
Total +/- 5	0.00	0.03	0.09	0.94	0.01	0.03	0.03	0.07	0.94	0.99		5	3	5	5	0.53	0.44	0.48	0.52
Total +/- 6	0.01	0.03	0.07	0.95	0.03	0.02	0.05	0.05	0.93	1.00		3	5	6	2	0.53	0.48	0.53	0.49

Notes.
The results consists of averages over four forecast horizons. See Table 5.4 for further explanation.

The only exception consists of the cyclical dynamic method $\tilde{\chi}_X^c$ exclusively employing the second group containing the industrial production series, which shows a smaller forecast error, i.e. ReMSE, and a better performance over time, i.e. $I_{A<B}(h) > 0.5$. Also the full sized data equivalent result holds for this specification: the data configuration of the complete data set excluding the second group shows worse statistics on forecasting performance as compared to the complete data set case.

While the individual groups exhibit a strong factor structure (high $\bar{R}^2$), most are prone to missampling ($R_{GDP}^2 < \bar{R}^2$). In case of data group 6 consisting of surveys however, even a strong factor structure (high $\bar{R}^2$) that provide a strong signal for the target variable (high R_{GDP}^2) does not progress forecasting performance. Also the full sized data equivalent result holds at forecast horizon $h = 1$: excluding the survey leads to improved factor diagnostics, especially an enhanced R_{GDP}^2 and lower dispersion R_q^2 , that progresses the forecasting performance as shown by a lower ReMSE and a higher $I_{A<B}(1)$. Apparently, the survey variables expose the data compilation to oversampling, thereby hampering forecast improvement. The surveys convey an idiosyncratically confined signal, which is complementary to the common variation in the complete data set as captured by the common factors⁶.

Taking the big sized complete data set configuration as the benchmark, then Table 5.4 and Table 5.5 show that diagnostics matter in general as the marginal improvement of the factor diagnostics resulting from excluding a particular group progresses upon the forecasting performance.

5.5 Conclusions

This study compares the forecasts of GDP growth rates for the Netherlands over a forecast horizon up to 4 quarters ahead based on alternative factor model specifications and various data set configurations. Based on each possible combination of models and data, the aim is to determine a relationship between the factor diagnostics and the forecasting performance.

Regarding the model specification, the results indicate that imposing the factor structure, despite poorer factor model diagnostics, improves upon the forecasting performance. The factor forecasts outperform the FAAR specifications consistently for all specifications and all horizons, significantly so at the first horizon. Of all the possible specifications considered in this study,

⁶However, note that we abstract from the release timing of the various variables. As surveys typically become available relatively quickly, they possess a timely and valuable signal awaiting the first release of hard data.

the FAAR specification imposes the factor structure to the least extent and therefore allows the forecast equation most flexibility to adapt to the data. According to the statistic of forecast accuracy over time, the best performing specification is the restricted cyclical dynamic factor forecast. This specification rests upon the most comprehensive factor design, which encompasses both dynamics and cyclical, and moreover, imposes the factor structure on the forecast equation.

Despite the better forecasting performance, the diagnostics of the restricted factor forecasts do however not compare favourably to their unrestricted counterparts for different specifications, forecast horizons and data set configurations. Comparing the diagnostics of employing static factors, dynamic factors and cyclical dynamic factors do not reveal structural differences, excepting that the static method represents the target variable better at the first horizon.

Regarding the data set configurations, the results show that size matters as the complete data set of 124 series outperforms for all horizons and almost all specifications the data configurations consisting of each of the six individual groups separately. Even though smaller macroeconomic data sets exhibit stronger coherence, the factors being well fit do, however, generally not relate to the variable of interest. Here, oversampling, or rather missampling, refers to the misrepresentation of a small data set exhibiting a strong factor structure.

Taking the big sized complete data set configuration as the benchmark for the restricted factor forecast specifications, then the results show that diagnostics matter as the marginal improvement of the factor diagnostics resulting from excluding a particular group progresses upon the forecasting performance.

5.A Appendix

5.A.1 The estimator

In this appendix, we show in more detail how the common component χ can be estimated in a stepwise procedure. Moreover, we will highlight the parameter condition that makes the static factor model a special case of the dynamic one. Finally, we show in more detail the estimator for the factor model forecasts in case the forecast equation is restricted to admit the factor model structure.

The dynamic method

The dynamic method as outlined in Forni et al. (2000; 2001; 2001a; 2004; 2005) (FHLR) consists of the frequency-domain counterpart of the static method. The dynamic factors⁷ $\mathbf{u}_t = (u_{1t} \dots u_{qt})'$ are estimated by the dynamic principal components, which are the static principal components of the spectral density matrix as outlined by Brillinger (1981). Let $\mathbf{X}^{nT}$ be the observations and $\hat{\Gamma}_{\mathbf{X}}^{nT}(k)$ its k -lag sample correlation matrix. FHLR suggest the following stepwise procedure:

(i) estimate the spectral density matrix (cf. Brillinger, 1981) of $\mathbf{X}^{nT}$ as

$$\Sigma_{\mathbf{X}}^{nT}(\theta_h) = \sum_{k=-M}^M \hat{\Gamma}_{\mathbf{X}}^{nT}(k) \omega_k e^{-ik\theta_h}, \quad \theta_h = 2\pi h / (2M+1), \quad h = 0, \dots, 2M,$$

where $\omega_k = 1 - |k| / (M+1)$ is the Bartlett window of size M . Like Forni et al. (2000), we set $M(T) = \text{ROUND}(2T^{(1/2)})$ such that the convergence rate is $M(T)/T = O(T^{(1/2)})$;

(ii) calculate from the spectral density matrix $\Sigma_{\mathbf{X}}^{nT}(\theta_h)$ the q largest dynamic eigenvalues $\lambda_j^{nT}(\theta_h)$ and the corresponding dynamic eigenvectors $\mathbf{p}_j^{nT}(\theta_h)$, $j = 1, \dots, q$ for $h = 0, \dots, 2M$. We follow Forni et al.'s (2000) approach and select $q = 3$ in a finite-sample such that the marginal explained variance of the q^{th} dynamic eigenvalue is larger than 10% and the $(q+1)^{th}$ equivalent is smaller than 10%;

(iii) let $\mathbf{p}_q^{nT}(\theta_h) = (\mathbf{p}_1^{nT}(\theta_h) \dots \mathbf{p}_q^{nT}(\theta_h))'$ the $(q \times n)$ -matrix of dynamic eigenvectors and $\mathbf{\Lambda}_q^{nT}(\theta_h)$ a diagonal matrix with the q largest dynamic eigenvalues on the diagonal. Inverse Fourier transformation of

⁷The notation of the dynamic factors $\mathbf{u}_t$ differs with $\mathbf{f}_t$ in (5.1) as the latter notation is now employed for the generalized principal component estimator of the dynamic factors in step (v) below.

$\tilde{\Sigma}_X^{nT}(\theta_h) = \mathbf{p}_q^{nT*}(\theta_h) \mathbf{\Lambda}_q^{nT}(\theta_h) \mathbf{p}_q^{nT}(\theta_h)$ (* denotes complex conjugate) results in the correlation matrix of the common component

$$\tilde{\Gamma}_X^{nT}(k) = \frac{1}{(2M+1)} \sum_{k=-M}^M \tilde{\Sigma}_X^{nT}(\theta_h) \omega_k e^{ik\theta_h} \text{ for } h = 0, \dots, 2M. \text{ Moreover, the es-}$$

timated common dynamic factors are $\tilde{\mathbf{u}}_t^{nT} = \frac{1}{(2M+1)} \sum_{k=-M}^M \sum_{h=0}^{2M} \mathbf{p}_q^{nT}(\theta_h) e^{ik\theta_h} \mathbf{x}_{t-k}^{nT}$.

Projecting the data on the common dynamic factors gives the estimator of the cyclical medium- and long-run common component:

$$\tilde{\phi}_{nt}^{nT} = \frac{1}{(2M+1)} \sum_{k=-M}^M \sum_{h=0}^{2M} \beta_k \mathbf{p}_q^{nT*}(\theta_h) \mathbf{p}_q^{nT}(\theta_h) e^{ik\theta_h} \mathbf{x}_{t-k}^{nT}, \quad (5.6)$$

where the finite sample approximation of $\beta(L)$ consists of truncating the tails of the band-pass filter that involve unavailable data observations, i.e. $\beta_k = 0$ for $k > M$.

(iv) repeat step (iii) using the $(q+1)$ to n ordered eigenvalues to obtain $\tilde{\Gamma}_\xi^{nT}(k)$;

(v) let $\tilde{\mathbf{S}}^{nT} = \left(\tilde{S}_1^{nT'} \dots \tilde{S}_r^{nT'} \right)'$ the $(r \times n)$ -matrix containing the r generalized eigenvectors of the couple of matrices $\left(\tilde{\Gamma}_X^{nT}(0), \tilde{\Gamma}_\xi^{nT}(0) \right)$ with the normalization that $\tilde{S}_i^{nT'} \text{DIAG}\left(\tilde{\Gamma}_\xi^{nT}(0)\right) \tilde{S}_j^{nT'} = 1$ if $i = j$ and zero otherwise. We use Bai and Ng's (2002) information criteria (BNIC) to determine the r generalized static factors as a trade-off between the goodness-of-fit and overfitting. The factors can then be estimated by the generalized principal components, i.e. $\tilde{\mathbf{F}}^{nT} = \tilde{\mathbf{S}}^{nT} \mathbf{X}^{nT}$, with $\tilde{\mathbf{F}}^{nT} = \left(\tilde{\mathbf{f}}_1^{nT} \dots \tilde{\mathbf{f}}_T^{nT} \right)$ a $(r \times T)$ -matrix of the stacked estimated factors;

(vi) let $\tilde{\chi}_{i,T+h|T}^{nT}$ be the h -step ahead factor forecasts of the common component of the i -th variable given T observations of n time series variables. The forecasts for the dynamic common component can be obtained by projecting the $(t+h)$ -dated unobserved common component $\mathbf{x}_{t+h}^{nT}$ on the t -dated factors $\tilde{\mathbf{f}}_t^{nT}$, which for variable i results in:

$$\tilde{\chi}_{i,T+h|T}^{nT} = \left[\tilde{\Gamma}_X^{nT}(h) \right]_i \tilde{\mathbf{S}}^{nT'} \left(\tilde{\mathbf{S}}^{nT} \tilde{\Gamma}_X^{nT}(0) \tilde{\mathbf{S}}^{nT'} \right)^{-1} \tilde{\mathbf{S}}^{nT} \mathbf{X}^{nT} \quad (5.7)$$

Evidently, the in-sample estimator for the common component can be obtained by setting $h = 0$.

Step (i) until step (iii) allow to estimate the dynamic factor model. The estimated cyclical common component $\tilde{\phi}_{nt}$ is calculated by applying time filters to the x 's before averaging along the cross-section. The dynamic estimation

method consists of two-sided filters and cannot be applied at the end of the sample, which is the most important part for forecasting. By truncating the time filters, the performance of the estimator deteriorates as t approaches T . Therefore in step (v), FHRLR construct generalized principal components $\mathbf{F}^{nT}$, which are contemporaneous averages of $\mathbf{X}^{nT}$ that minimize the ratio of the variance of the idiosyncratic to common component.

5.A.2 Dutch data set

The appendix describes the data set for the Dutch economy. The aim is to construct an exhaustive collection covering different economic spheres, which gives a balanced representation of the economy and of the forces influencing it. For this purpose, the data set for the macroeconometric model MORKMON of De Nederlandsche Bank is screened and supplemented with a set of macro variables of forward looking nature. The data set consists of stock variables of five sectors, namely households, business, monetary financial institutions, government and external world, and the variables describing the flows between these sectors. The data set is screened on variables that are available only at a yearly frequency, especially within the sphere of public finance and social security, taxation and capital formation. This data set is supplemented with sectorally disaggregated production series, surveys and leading indicators, external economic developments and international financial developments as transmitted by equity prices, a broader set of interest rates, exchange rates and commodity prices. The data is preferably collected on a seasonally (and calendar effects) adjusted basis at a quarterly sampling frequency. Some of the series available on a quarterly frequency are only disposable in raw format and are seasonally adjusted by applying the census-X12 method. Other series like interest rates, exchange rates and equity prices are kept in raw format.

Table A.5.6 lists all the series and the columns report respectively the description of the variable, unit of measurement, transformation code to render the variable stationary and the original data source. The automated procedure of TRAMO (cf Gómez and Maravall, 1996) is applied to correct the data for outliers and missing observations. Subsequently, the time series are rendered stationary by following one of the codes: 1 = no transformation for capacity utilization rates, unemployment rates, ratios and interest rate spreads, 2 = first difference for interest rates, surveys, sentiment indicators and, in general, (nonstationary) series possessing negative values like balance of payments statistics, 3 = first difference of logarithms producing quarterly growth rates for the vast majority of the series and 4 = second difference of logarithms for

nonstationary series like wages, consumer prices, producer prices, commodity prices and monetary aggregates.

As is required for factor estimation, the variables were standardized by subtracting their mean and then dividing by their standard deviation. This standardization is necessary to avoid overweighting of large variance series in the factor estimation. The full data set consists of 124 series that can be divided equally into six different categories labeled GDP, industrial production, prices, financial, external and surveys. The sample period runs from 1980Q1 until 2002Q4. Moreover, the data are collected in the second quarter of 2004 and represents therefore the fully revised historical series, or equivalently, the 2004Q2 snapshot of the data.

Table A.5.6: Description of data set

#	Description	Unit	Transformation code ^a	Original source ^b
Group 1: GDP, gross value added and real activity				
1	Gross domestic product by expenditure, constant prices	mil. euro 95	3	CBS
2	Private final consumption expenditure incl. NPI-h, constant prices	mil. euro 95	3	CBS
3	Government final consumption expenditure, constant prices	mil. euro 95	3	CBS
4	Gross fixed capital formation, constant prices	mil. euro 95	3	CBS
5	Gross fixed capital formation of dwellings, constant prices	mil. euro 95	3	CBS
6	Gross fixed capital formation of machinery and equipment, constant prices	mil. euro 95	3	CBS
7	Gross domestic product by expenditure, OECD (25), constant prices	index 1995=100	3	OECD, QNA
8	Compensation of employees	mil. euro	4	CBS
9	Unemployment	persons*1000	4	CBS
10	Number of jobs employees	jobs*1000	4	CBS
11	Negotiated wage (monthly base)	index 1995=100	4	CBS
12	Collective final consumption expenditure of general government	mil. euro	3	CBS
13	Capital formation excluding changes in inventories (sector), constant prices	mil. euro 95	3	CBS
14	Negotiated wage (monthly base)	index 1995=100	4	CBS
15	Residence permits granted	number	3	CBS
16	Houseprices	euro * 1000	4	Kadaster
17	Negotiated wage (all sectors: monthly base)	index 1995=100	4	CBS
18	hourly wages, industry	index 2000=100	4	CBS
19	issued vehicle registration certificates	number (end of period)	3	CBS
20	Composite Leading Indicator (trend restored)	indicator	2	DS
21	WO business cycle indicator NL	indicator	1	DNB - division WO
Group 2: Industrial Production and capacity utilization				
22	Productive hours worked per employee in construction	index 1995=100	4	CBS
23	Capacity utilization in manufacturing industry	%	3	CBS
24	World capacity utilization in manufacturing industry	index 1995=100	3	OECD
25	production of consumptiongoods (average daily production)	index 2000=100	3	CBS
26	production of investmentgoods (average daily production)	index 2000=100	3	CBS
27	average daily production - production industries	index 2000=100	3	CBS
28	average daily production - energycompanies and waterworks	index 2000=100	3	CBS
29	average daily production - mineral extraction	index 2000=100	3	CBS
30	average daily production - industry	index 2000=100	3	CBS
31	labour productivity, production per employed person	index 2000=100	3	CBS
32	earnings per employee, private businesses, general government and other sectors	index 1995=100	3	DNB/CBS
33	production per employee, private businesses, general government and other sectors	index 1995=100	3	DNB/CBS
34	labour costs per unit, private businesses, general government and other sectors	index 1995=100	3	DNB/CBS
35	capacity utilization manufacturing industry	percentage	2	DS
36	capacity utilization intermediate and final goods	percentage	2	CBS
37	capacity utilization consumer goods	percentage	2	CBS
38	capacity utilization investment goods	percentage	2	CBS
39	capacity utilization intermediate products	percentage	2	CBS
40	labour costs per unit product, processing industry	index 1995=100	3	EC
41	industrial turnover, foreign market, manufacturing	index 2000=100	3	CBS
42	industrial turnover, domestic market, manufacturing	index 2000=100	3	CBS

Notes.

^a: 1 = no transformation; 2 = first differences; 3 = first difference of logarithms producing quarterly growth rates; and 4 = second difference of logarithms for nonstationary series.

^b: BIS Bank of International Settlements; CBS Central Bureau of Statistics; DNB-FM De Nederlandsche Bank, divisie Financiële Markten; DS Datastream; EC European Commission; ECB European Central Bank; HWWA Institut für Wirtschaftsforschung; OECD, QNA Organisation for Economic Co-operation and Development, Quarterly National Accounts.

Description of data set (2)

#	Description	Unit	Transformation code ^a	Original source ^b
Group 3: Prices				
43	Large scale price of natural gas	eurocent p/m3	4	CBS
44	Foreign consumer price	index 1995=100	4	SIR
45	HICP component housing	index 1995=100	4	CBS
46	gas price index, small-scale, excl.vat	index 1990=100	4	CBS
47	spot crude oil UK Brent	US-dollar per barrel	4	OPEC
48	world market commodity prices, overall (euro area)	index 2000=100	4	HWWA
49	world market commodity prices, overall excl. energy (euro area)	index 2000=100	4	HWWA
50	world market commodity prices, food and luxury foods (euro area)	index 2000=100	4	HWWA
51	world market commodity prices, industrial materials (euro area)	index 2000=100	4	HWWA
52	world market commodity prices, agricultural-industrial materials (euro area)	index 2000=100	4	HWWA
53	world market commodity prices, metals (euro area)	index 2000=100	4	HWWA
54	world market commodity prices, energy-components (euro area)	index 2000=100	4	HWWA
55	Producer prices, sale, industry, dom.+for.market, total interm.+final products	Index 2000=100	4	CBS
56	Producer prices, large-scale gas consumption, dom.+for.market (index)	Index 1990=100	4	CBS
57	Consumerprice index NL, total CPI, all households	Index 2000=100	4	CBS
58	Consumerprice index NL, underlying inflation	Index 2000=100	4	CBS
59	Consumerprice index NL, energy	Index 2000=100	4	CBS
60	Consumerprice index NL, vegetables and fruit	Index 2000=100	4	CBS
Group 4: Financial				
61	Short term interest rate	%-point	2	DNB-FM
62	Long term interest rate	%-point	2	DS
63	Exchange rate	dollar per euro	3	ECB
64	Domestic stock market prices	1983-IV=100 (end)	4	CBS
65	British pound per euro	number	3	ECB
66	Japanese yen per euro	number	3	ECB
67	Effective return om government bonds	percent	2	CBS
68	Effective return on national loan (3-5 year)	percent	2	CBS
69	Effective return on national loan (5-8 year)	percent	2	CBS
70	Effective return on bank-bonds	percent	2	CBS
71	Effective return on mortgage bonds	percent	2	CBS
72	M1	mil. euro	4	ECB
73	M3 (money in circulation inclusive)	mil. euro	4	ECB
74	spread (67 - 61)		1	calculation
75	spread (68 - 61)		1	calculation
76	spread (69 - 61)		1	calculation
77	spread (r1 - 61)		1	calculation
78	Amsterdam Midkap-index	index 830103=45,4(EUR)	4	Euronext Amsterdam
79	stock prices, CBS General	index 1983=100	4	CBS
80	stock prices, internationals	index 1983=100	4	CBS
81	stock prices, inland (domestic)	index 1983=100	4	CBS
82	stock prices, financial institutions	index 1983=100	4	CBS
83	stock prices, non-financial institutions	index 1983=100	4	CBS
84	stock prices, general reinvestment index	index 1983=100	4	CBS

Notes. See Table A.5.6.

Description of data set (3)

#	Description	Unit	Transformation code ^a	Original source ^b
Group 5: External Sector				
85	gdp germany	mil. euro 95	3	ECB
86	gdp belgium	mil. euro 95	3	ECB
87	gdp united kingdom	mil. Br. pound 95	3	ECB
88	gdp united states	bil. US-dollar 2000	3	BIS
89	gdp japan	bil. Jap. yen 95	3	BIS
90	gdp france	mil. euro 95	3	ECB
91	gdp italy	mil. euro 95	3	ECB
92	exports of goods	mil. euro	4	DNB, balance of payments section
93	imports of goods	mil. euro	4	DNB, balance of payments section
94	balance on goods	mil. euro	1	DNB, balance of payments section
95	exports of services	mil. euro	3	DNB, balance of payments section
96	imports of services	mil. euro	3	DNB, balance of payments section
97	balance on services	mil. euro	2	DNB, balance of payments section
98	receipts (income account)	mil. euro	3	DNB, balance of payments section
99	expenditures (income account)	mil. euro	3	DNB, balance of payments section
100	balance on income	mil. euro	2	DNB, balance of payments section
101	receipts (current transfers account)	mil. euro	3	DNB, balance of payments section
102	expenditures (current transfers account)	mil. euro	3	DNB, balance of payments section
103	net current transfers	mil. euro	2	DNB, balance of payments section
104	balance on current account	mil. euro	2	DNB, balance of payments section
105	inland volume of trade NCM	mil. euro	3	Nederlandse Crediet Maatschappij
106	IFO-indicator	index 2000=100	1	IFO-Instituut
Group 6: Surveys				
107	consumer confidence	percentage	2	DS
108	producer confidence	percentage	2	DS
109	nl business tendency survey: mfg. - export orders inflow	percentage	2	DS - OECD
110	nl business tendency survey: mfg. - finished goods stocks	percentage	2	DS - OECD
111	nl business tendency survey: mfg. - future production	percentage	2	DS - OECD
112	nl business tendency survey: manufacturing - order books	percentage	2	DS - OECD
113	nl business tendency survey: manufacturing - ordersinflow	percentage	2	DS - OECD
114	nl business tendency survey: manufacturing - production	percentage	2	DS - OECD
115	nl construction survey: order book position	index - diffusion	2	DS - EC
116	nl consumer opinion survey: confidence indicator	percentage	2	DS - OECD
117	nl industry survey: capacity utilisation	index - diffusion	2	DS - EC
118	nl industry survey: current production capacity	index - diffusion	2	DS - EC
119	nl industry survey: export expectation for mo. ahead	index - diffusion	2	DS - EC
120	nl industry survey: mth. prod. assured by order book	index - diffusion	2	DS - EC
121	nl industry survey: new order pstn. in recent months	index - diffusion	2	DS - EC
122	nl industry survey: order book position	index - diffusion	2	DS - EC
123	nl industry survey: prod.expectation for mth. ahead	index - diffusion	2	DS - EC
124	nl industry survey: stocks of finished goods	index - diffusion	2	DS - EC

Notes. See Table A.5.6. Bank of International Settlements; CBS Central Bureau of Statistics; DNB-FM De Nederlandsche Bank, divisie Financiële Markten; DS Datastream; EC European Commission; ECB European Central Bank; HWWA Institut für Wirtschaftsforschung; OECD, QNA Organisation for Economic Co-operation and Development, Quarterly National Accounts.

Chapter 6

Regional and Sectoral Dynamics of the Dutch Staffing Labour Cycle.

Statistical pointillism: finding patterns and relationships where none seem to exist.

Jason Benderly

Abstract

This study analyses the dynamic characteristics of staffing employment across different business sectors and across different geographical regions in the Netherlands. We analyse a micro data set of the market leader of the Dutch staffing employment market, i.e. Randstad. We apply the dynamic factor model to extract common information out of a large data set and to isolate business cycle frequencies with the aim of forecasting staffing employment. We identify regions and sectors whose cyclical developments lead the staffing labour cycle at the country level. The second question is then which model specification can best exploit the identified leading indicators at the disaggregate level to forecast the country aggregate? The dynamic factor model turns out to outperform univariate benchmark forecasting models by exploiting the substantial temporal variation of the staffing labour market at the disaggregate level.

6.1 Introduction

Flexible staffing agency work is characterized by a triangular relationship between the user firm, the employee and the private labour market intermediary (cf Gottfried, 1992). The staffing agency is a private matchmaker that acts as an intermediary between temporary labour supply and demand. Staffing agencies derive their income from fees charged to user firms for the temporary employment of workers registered with the agency. Staffing agencies perform

I would like to thank Randstad Nederland for providing the confidential data. Moreover, I would like to thank Lex Hoogduin, Massimiliano Marcellino, Franz Palm, Guido Schotten and seminar participants at the 5th *studiedag conjunctuur* at Nijenrode University and an internal seminar for comments and suggestions.

the recruitment of personnel and provide a ready source of labour for their business clients. The flexible staffing industry effectively creates a spot market for labour, so user firms can replace absent employees or adjust the labour force to short-term changes and fluctuations in market demand without incurring the usual hiring and firing costs (cf. Katz and Krueger, 1999). From the perspective of the client firm, flexible staffing labour constitutes a mere variable factor of production. Peck and Theodore (2007) and Theodore and Peck (2002) show that the American flexible staffing industry is not just a purveyor of flexibility at the micro level of meeting the needs of individual enterprises, but also at the macro level of mediating macroeconomic pressures and socioeconomic risks across the labour market as a whole. During the last 30 years, temporary employment expanded rapidly prior to macroeconomic upturns, while sharp declines in temporary employment preceded recessions (cf. Segal and Sullivan, 1997; Theodore and Peck, 2002). Hence, fluctuations in staffing employment are timely indicators of broader business cycle motions. Berkhout and Van Leeuwen's (2004) international comparison shows a mature Dutch flexible staffing industry that serves a relatively large part of total employment. Goldschmeding (2003), Franses and de Groot (2005b) and den Reijer (2006) analyse the Dutch staffing labour market developments to monitor and forecast macroeconomic business cycles. The primary objective of this paper is to document the cyclical developments of staffing employment in the Netherlands at the disaggregate level and to identify the regions and sectors that show leading properties, (cf. Forni et al., 2001). Like Kvasnicka's (2003) German data set, the observations are directly obtained from the administrative source of a market participant instead of using survey based data. The second question is then how the disaggregate information, particularly the identified leading indicators at the sectoral and geographical level, can be exploited to forecast the country aggregate of staffing employment. The paper is structured as follows. Section 6.2 describes the staffing labour market and in the available data set. Section 6.3 introduces the factor model that is employed to extract the staffing labour cycle from the data. Section 6.4 classifies the staffing labour cycle at the disaggregate level and identifies the leading and lagging regions and sectors. Finally, section 6.5 compares different model specifications that exploit the information at the disaggregate level to forecast the staffing labour developments at the country level.

6.2 Staffing agency work

The private employment agency, which is often referred to as staffing services organization¹, transforms labour from a quasi-fixed into a variable factor of production and therefore effectively create an efficient spot market for labour. The structure of the labour market and the importance of temporary and agency work differ between countries because of the legislative framework, see Berkhout and Van Leeuwen (2004) for an international comparison and Dunnewijk (2001) for a brief history in the Netherlands. The comparatively mature Dutch staffing services market grew from its inception as a percentage of the labour force from 0% in 1960 to 5% in 2004. Randstad Netherlands (Randstad hereafter) is the market leader and covers a stable market share of 40% over this entire period (cf. Franses and de Groot, 2005a). Randstad is the country branch of Randstad Holding², which is one of the largest temporary and contract staffing organizations in the world.

6.2.1 Data

The data set is directly obtained from the administrative source of Randstad and are nearly real-time available. The available data set consists of 1.276.393 observations on the number of contracted staffing hours. The data run from 1998 until 2005 and each year is divided into 13 subsequent administrative periods of a four week duration. Every observation consists of four dimensions; the number of staffing hours for each time period is sectorally and geographically disaggregated. The sectoral classification occurs along the four digit SBI-code and the geographical classification along the four digit system of postal codes, see appendix 6.A for details.

At this level of disaggregation, each observation almost always corresponds to a single user firm. By nature of staffing employment, a single user firm does not make use of staffing services continuously during the entire sample period. In order to create a balanced data set, we aggregate the individual observations to the level of 15 regions and 58 sectors. The regions consist of the 12 provinces from which the agglomerations of the three largest cities are separated out. The sectors correspond to the two digit SBI-code. Now, $X_{ij,t}$ represents the total number of hours of staffing employment in region $i = 1, \dots, 15$ and sector $j = 1, \dots, 58$ during period t , running from the

¹The terminology of "agency work", "agency worker" and "employment agency" is practiced by the *International Confederation of Temporary Work Businesses* (CIETT). The alternative terminology of "staffing work", "staffing employee" and "staffing company" is used by the *American Staffing Association* (ASA).

²see <http://www.randstad.com>

first period in 1998 until the second period in 2005 consisting of 92 four-weeks periods. Moreover, we create a balanced data set by deleting the combination of region i and sector j if the time series shows missing observations, that is delete X_{ij} if $\exists t$ such that $X_{ij,t} = 0$. Out of the $15 \times 58 = 870$ possible combinations, the balanced data set consists of $N=536$ different time series. Some combinations are not feasible as the type of economic activity is hardly performed in the particular region, e.g. the activity Fishing in the province Drenthe, or not present at all in the Netherlands, e.g. the activity Mining of uranium and thorium ores. The resulting balanced data set covers 97.3% of the total data set in terms of the number of observations and 98.2% in terms of the number of staffing hours. Seven sectors disappear for the balanced data set and, on the other hand, 22 sectors do not lose observations at all as a result of balancing the data set. The overall loss of roughly 2% of observations is not concentrated within a specific remaining sector, region or time period.

We calculate the aggregates of each sector i , each region j and the country total as $X_{i*,t} = \sum_j X_{ij,t}$, $X_{*,j,t} = \sum_i X_{ij,t}$ and $X_t = \sum_i \sum_j X_{ij,t}$ respectively. In order to apply the dynamic factor model, all series are transformed to remove non-stationarity and corrected for outliers. The stationarity inducing transformation amounts to calculating the period-on-period growth rates³, so we analyse $x_{ij,t} = (1 - L) \ln(X_{ij,t})$, where L is the lag operator. The time series of growth rates are corrected for outliers by replacing those observed growth rates that are more than three sample standard deviations away from the sample mean with the average of the remaining observed growth rates. In order to apply the factor method as outlined below, we construct standardized growth rates $x_{ij,t}^s$ by subtracting the sample average from the outlier corrected growth rates and dividing by the sample standard deviation⁴.

6.3 Dynamic factor model

In order to extract the cyclical developments of staffing employment in the Netherlands at the disaggregate level and to identify the regions and sectors that show leading properties for the staffing cycle at the aggregate level, we fit a dynamic factor model to the balanced stationary data set. We apply the

³Considering the country total as a time series variable sampled at a quarterly frequency over the sample period 1967.1-2004.4, Franses and de Groot (2005a) find no evidence for a seasonal unit root performing the HEGY test statistic.

⁴Preprocessing the data by stationarity inducing transformation, outlier correction and standardisation is common practice in the literature, see Breitung and Eickmeier (2006) for an overview of factor models and their applications to economic indicators, forecasting and business cycle analysis.

methodology of Forni et al. (2000; 2001; 2001a; 2004; 2005) that was developed to extract coincident and leading indicators for the euro area from a large panel of economic variables of member countries. Factor models are a tool to cope with many variables without running into problems of too little degrees of freedom often faced in regression based analysis.

Firstly, factor models summarize large data sets in few underlying forces. The extracted low-dimensional common information is then used to discern the "common signal" χ from the "idiosyncratic noise" ξ for each of the underlying variables, so

$$x_{ij,t}^s = \chi_{ij,t} + \xi_{ij,t} \quad (6.1)$$

The idiosyncratic motion of a variable includes the effects of local shocks that are typically sector or region specific, while the common signal affects all sectors and regions. The common component $\chi_{ij,t}$ is driven by the impact of $k = 1, \dots, q$ unobserved "dynamic factors" u_{kt} that are common to all the variables in the data set.

Secondly, the dynamic factor model allows for factor loadings $\alpha_{ijk}(L)$, $k = 1, \dots, q$, which describe the dynamic impact of the common dynamic factors u_{kt} on the common component:

$$\chi_{ij,t} = \alpha_{ij1}(L) u_{1t} + \dots + \alpha_{ijq}(L) u_{qt}. \quad (6.2)$$

The common driving forces u_k can affect the individual variables with different leads and lags, which enables to classify the variables, regions and sectors as leading, coincident and lagging.

Thirdly, we further decompose the common component χ_{ijt} into a cyclical medium- and long-run component ϕ_{ijt} and a non-cyclical seasonal and irregular part ψ_{ijt} , that is

$$x_{ijt}^s = \phi_{ijt} + \psi_{ijt} + \xi_{ijt}. \quad (6.3)$$

This decomposition is based on a two-sided, symmetric, square summable bandpass filter $\beta(L)$, which separates waves of periodicity larger than a given critical number of periods τ :

$$\phi_{ij,t} = \sum_{k=-\infty}^{\infty} \beta_k \chi_{ij,t-k}, \quad \beta_k = \begin{cases} \frac{1}{k\pi} \sin(2k\pi/\tau) & \text{for } k \neq 0 \\ 1/\tau & \text{for } k = 0 \end{cases}. \quad (6.4)$$

The cyclical medium- and long-run component $\phi_{ij,t}$ is thereby filtered for short-run seasonal and erratic fluctuations and therefore signals more smoothly

the underlying development of the staffing employment growth⁵.

In order to estimate the generalized dynamic factor model, we need to specify the number of dynamic factors q , the parameter M that determines the maximum lag of auto-covariance matrix and the cyclical parameter τ , see appendix 6.A.3 for details. The identifying factor model assumption requires that the q largest dynamic eigenvalues diverge, whereas the remaining $N - q$ eigenvalues remain bounded as the number of time series variables N increases. We follow Forni et al.'s (2000) approach and select $q = 3$ in the finite-sample, because the marginal explained variance of the q^{th} dynamic eigenvalue is larger than 10% and the $(q + 1)^{th}$ one is smaller than 10%. The corresponding q dynamic eigenvectors are the estimators for the common dynamic factors u_k , $k = 1, \dots, q$ and the dynamic factor loadings $\alpha_{ijk}(L)$ describe the dynamic impact of the k -th common factor u_k on the time series variable x_{ijt} . We use a data dependent rule to set the maximum lead and lag of M periods, that is $\alpha_{ijk, \pm n} L^{\pm n} u_{kt} = 0$ for $n > M$, at $M(T) = \text{round}\left(2T^{(1/2)}\right) = 19$ for our data set of $T = 92$ observations in the time dimension. Finally, we set $\tau = 13$, so all seasonality, which by definition entails a duration shorter than 1 year, or 13 periods, is filtered out. The medium- and long-run component then describes the cyclical of duration longer than one year and, given the length of the sample of observations of $T = 92$ periods, shorter than seven years.

Figure 6.1 plots the year-on-year growth rates of the total employment in the Netherlands as reported by Statistics Netherlands, the year-on-year growth rates of the country aggregate of the turnover of Randstad⁶ together with the aggregate common signal $\hat{\phi}_t$.

6.3.1 Aggregate and aggregated staffing employment

The signal at the country level ϕ_t , at the sectoral level $\phi_{i*,t}$, at the regional level $\phi_{*,j,t}$ and at the disaggregate level $\phi_{ij,t}$ can be determined by projecting the corresponding aggregates x_t^s , $x_{i*,t}^s$, $x_{*,j,t}^s$ and $x_{ij,t}^s$ respectively on the dynamic factors u_{kt} , which can be estimated by dynamic principal components, see appendix 6.A.3. The linear projection of the data on the dynamic principal components provides the parameter estimates of (6.2), that is the factor

⁵CEPR's coincident indicator of the euro area (Eurocoin) reposes on a similarly composed measure that captures the cyclical signal underlying short-lived oscillations, see Altissimo et al. (2006).

⁶The history of the country aggregate of turnover data of Randstad is obtained from Franses and de Groot (2005a). Moreover, the staffing data from 2005 onwards originates from the Dutch association of temporary work agencies, which covers 60 per cent of the market, see <http://www.abu.nl>

Figure 6.1: Year-on-year growth rate of total employment, staffing employment and turnover

loadings $\hat{\alpha}_k(L)$, $\hat{\alpha}_{i*,k}(L)$, $\hat{\alpha}_{*j,k}(L)$ and $\hat{\alpha}_{ijk}(L)$, $k = 1, \dots, q$ respectively.

Alternatively, the aggregated signal of the country, each sector i and each region j can be constructed as the weighted aggregate of the individual signals $\phi_{ij,t}$, that is $\bar{\phi}_{i*,t} = \sum_j a_{*j,t-1} \phi_{ij,t}$, $\bar{\phi}_{*j,t} = \sum_i a_{i*,t-1} \phi_{ij,t}$ and $\bar{\phi}_t = \sum_i \sum_j a_{ij,t-1} \phi_{ij,t}$ respectively. The time-varying weights $a_{ij,t}$ are the shares of the individual variables in the aggregate multiplied by the ratio of the standard deviations of x_t and $x_{ij,t}$: $a_{ij,t} = \frac{\sigma_{x_{ij}}}{\sigma_x} b_{ij,t}$ with $b_{ij,t} = \frac{X_{ij,t}}{\sum_i \sum_j X_{ij,t}}$ the share of variable $X_{ij,t}$, which belongs to sector i and region j , in the total staffing turnover X_t at time t . Appendix 6.A.2 shows that calculating the standardized growth rates x_t^s from the aggregate X_t is mathematically equivalent to aggregating the standardized growth rates of the disaggregates $x_{ij,t}^s$ using the delayed weights $a_{ij,t-1}$. However, projecting the aggregate x_t^s on the dynamic principal components is only mathematically equivalent to aggregating the projected disaggregates $x_{ij,t}^s$ if the weights are constant $\alpha_{ij,t} = \alpha_{ij}$. So, the mathematical equivalence between the aggregate signal and the aggregated signal, $\phi_t = \bar{\phi}_t$, does not hold exactly because of time varying weights.

Figure 6.2 shows the aggregate growth rate x_t , its aggregate common component $\hat{\chi}_t$, the time-varying aggregated signal $\hat{\bar{\phi}}_t$ and the aggregate signal

Figure 6.2: Invoiced staffing hours and its model model decomposition

$\hat{\phi}_t$. The figure suggests that the aggregated signal is empirically equivalent to the aggregate signal even though time varying aggregation weights are employed.

The correlation structure of the panel of observations across both the regional and sectoral dimensions characterizes the staffing labour cycle in the Netherlands at a disaggregate level. The correlation is calculated using a sample period of seven years of available observations and is therefore only based on at most one complete business cycle. The correlation structure is summarized by the cross-correlation ρ of each individual variable's cyclical medium- and long-run component $\phi_{ij,t}$ with the aggregate cycle ϕ_t . The optimal lead l^* is determined as $l_{ij}^* = \arg \max_l |\rho_{ij}(\phi_t, \phi_{ij,t-l})|$ and its corresponding correlation $\rho_{ij}^* = \rho_{ij}(\phi_t, \phi_{ij,t-l_{ij}^*})$. The optimal aggregate correlation and lead measures (ρ_{i*}^*, l_{i*}^*) and (ρ_{*j}^*, l_{*j}^*) are likewise obtained by employing the regional aggregate $\phi_{i*,t}$ and the sectoral aggregate $\phi_{*j,t}$ respectively.

Optimal aggregated correlation and lead measures $(\bar{\rho}_{i*}^*, \bar{l}_{i*}^*)$ and $(\bar{\rho}_{*j}^*, \bar{l}_{*j}^*)$ at the regional and sectoral level, respectively, are alternatively obtained as the weighted average of the optimal disaggregate measures, i.e. $\bar{l}_{i*}^* = \sum_j b_{*j|T} l_{ij}^*$

and $\bar{l}_{*j}^* = \sum_i b_{i*|T} l_{ij}^*$, respectively, where the weights $b_{ij|T}$ represents the variable's share in the total staffing turnover that is constant over time: $b_{ij|T} = \frac{1}{T} \sum_{t=1}^T b_{ij,t}$. The empirical measures $\bar{\rho}^*$ are likewise obtained.

The following stylized example illustrates the difference between the optimal aggregate measures (ρ^*, l^*) and the aggregated optimal measures $(\bar{\rho}^*, \bar{l}^*)$. Consider a data set of three series y_{ij} , $i = 1, 2, 3$ with weights $b_{ij} = \frac{1}{6}, \frac{1}{6}, \frac{2}{3}$ respectively. Let the correlation coefficient be $\rho_{ij}(y_t, y_{ij,t-l_{ij}}) = \frac{1}{2}$ and zero otherwise with the corresponding leads $l_{ij} = -1, 0, 1$ respectively. The weighted average correlation is $\bar{\rho}_{*j}^* = \frac{1}{2}$ and $\bar{l}_{*j}^* = -\frac{1}{6} + \frac{2}{3} = \frac{1}{2}$. The aggregated series reads as $y_{*j} = \frac{1}{6}y_{1j} + \frac{1}{6}y_{2j} + \frac{2}{3}y_{3j}$ and its optimal measures are $\rho_{*j}^* = \frac{1}{2}$ and $l_{*j}^* = 1$. The stylized example shows that the optimal aggregate measures capture the characteristics of the underlying time series variable that is most dominant in terms of weight and cross-correlation. Given equal weights and positive correlation coefficients, the underlying time series variable that shows the highest cross-correlation will be selected: $\rho^* \geq \bar{\rho}^* (\geq 0)$.

6.4 The empirics of staffing employment

We summarize the correlation structure across both the regional and sectoral dimensions of the staffing labour cycle as represented by the medium- and long-run signal $\hat{\phi}$ and the year-on-year growth rate of X_{ijt} , i.e. $x_{ijt}^{13} = (1 - L^{13}) \ln(X_{ijt})$. Christiano and Fitzgerald's (2003) band-pass filter is applied to the latter growth rates to smooth away the irregularities and remaining seasonalities with periodicity smaller than one year. The year-on-year growth rates relate to the period-on-period growth rates as follows⁷:

$x_{it}^{13} = (1 + L + \dots + L^{12})(\phi_{it} + \psi_{it} + \xi_{it})$. The cross-correlation of x_{it}^{13} with another variable, say x_{kt}^{13} , involves, apart from the cross-correlation $\rho(\phi_{it}, \phi_{kt})$, at least the auto-correlations $\rho(\phi_{it}, \phi_{kt \pm j})$, $j = 1, \dots, 12$ and the correlations between the seasonal components $\rho(\psi_{it}, \psi_{kt \pm j})$ and the idiosyncratic components $\rho(\xi_{it}, \xi_{kt \pm j})$, $j = 0, \dots, 12$.

⁷Note that $(1 - L^{13}) = (1 - L)(1 + L + \dots + L^{12})$. Then, $x_{it}^{13} = (1 + L + \dots + L^{12})x_{it} = (1 + L + \dots + L^{12})(\phi_{it} + \psi_{it} + \xi_{it})$.

Table 6.1: The empirical results of the staffing labour cycle at the regional level.

	$\overline{var}(\bar{x})$	$\overline{var}(\bar{y})$	$\bar{\rho}_{\bar{x}\bar{y}}(\phi)$	$\bar{\rho}_{\bar{x}\bar{y}}(x^{13})$	$\bar{\rho}_{\bar{x}\bar{y}}^{\perp}(\phi)$	$\bar{\rho}_{\bar{x}\bar{y}}^{\perp}(x^{13})$	$\bar{\tau}_{\bar{x}\bar{y}}(\phi)$	$\bar{\tau}_{\bar{x}\bar{y}}(x^{13})$	$\bar{\tau}_{\bar{x}\bar{y}}^{\perp}(\phi)$	$\bar{\tau}_{\bar{x}\bar{y}}^{\perp}(x^{13})$
Drenthe	0.38	0.80	0.11	0.16	0.60	0.65	-0.31	-2.62	0.00	0.00
Noord-Brabant	0.49	0.84	0.05	0.16	0.89	0.76	-3.43	-0.71	6.00	3.00
Noord-Holland	0.46	0.79	0.26	0.07	0.96	0.78	-0.75	0.85	-2.00	-4.00
Gelderland	0.54	0.96	0.51	0.38	0.77	0.85	0.04	1.11	4.00	1.00
Zuid-Holland	0.49	0.90	0.36	0.10	0.68	0.68	-2.54	2.49	0.00	-7.00
Zeeland	0.29	0.53	0.03	-0.14	-0.73	-0.45	0.12	3.03	17.00	7.00
Friesland	0.38	0.82	0.16	0.10	0.86	0.37	-5.90	-0.73	0.00	0.00
Overijssel	0.47	0.80	0.10	0.16	0.70	0.85	2.00	1.89	2.00	0.00
Flevoland	0.31	0.60	0.34	0.25	0.81	0.63	-0.35	2.42	0.00	0.00
Limburg	0.40	0.74	0.16	0.11	0.98	0.90	5.31	2.02	0.00	-2.00
Utrecht	0.51	0.77	0.11	0.31	0.95	0.74	0.66	1.59	0.00	0.00
Groningen	0.36	0.81	0.06	0.02	0.52	0.59	3.22	-1.06	0.00	-2.00
Groot-Amsterdam	0.39	0.69	0.44	0.43	0.84	0.78	2.81	1.01	1.00	-1.00
Groot-Rijnmond	0.40	0.89	0.44	0.15	0.81	0.68	0.59	1.49	5.00	-1.00
Agg.-s-Gravenhage	0.24	0.74	-0.34	-0.01	-0.86	0.53	-0.12	3.24	11.00	3.00

Notes:

The variance ($\overline{var}$) of the common component reports the fraction of the total variance that is explained by the static factor model. The correlation ρ and the time shift l are calculated for the cycle of the corresponding region with respect to the aggregate cycle at the country level. The table reports the result for both the signal ϕ , which is extracted with the factor model, and the growth rates x^{13} . The aggregate measures without a bar summarize the results for the aggregate variables at the regional level. The aggregated measures with a bar summarize the results for the disaggregate variables. The disaggregate results are aggregated to the regional level using as aggregation weights $b_{*|l7}$.

Table 6.1: The empirical results of the staffing labour cycle at the sectoral level.

	$\overline{var}_{s,j}(\bar{x})$	$\overline{var}_{s,j}(\bar{x})$	$\overline{p}_{s,j}^*(\hat{\phi})$	$\overline{p}_{s,j}^*(x^{13})$	$\rho_{s,j}^*(\hat{\phi})$	$\rho_{s,j}^*(x^{13})$	$\overline{p}_{s,j}^*(\hat{\phi})$	$\overline{p}_{s,j}^*(x^{13})$	$\overline{p}_{s,j}^*(\hat{\phi})$	$\overline{p}_{s,j}^*(x^{13})$
Agriculture, hunting and related service activities	0.29	0.70	0.19	0.10	0.87	0.88	3.22	-1.75	21.00	16.00
Manufacture of food products and beverages	0.56	0.81	-0.02	-0.05	0.71	-0.05	6.36	0.39	0.00	1.00
Manufacture of tobacco products	0.18	0.26	-0.69	-0.60	-0.74	-0.64	6.04	0.00	7.00	1.00
Manufacture of textiles	0.31	0.48	0.39	0.52	0.87	0.84	4.47	-0.54	3.00	-2.00
Tanning and dressing of leather	0.45	0.31	0.29	0.57	0.51	0.79	4.48	4.54	8.00	7.00
Manufacture of wearing apparel, dressing and dyeing of fur	0.23	0.31	0.29	0.57	0.51	0.79	4.48	4.54	8.00	7.00
Manufacture of wood and of products of wood and cork, except furniture	0.45	0.48	-0.72	-0.72	0.40	0.40	14.00	0.00	14.00	0.00
Manufacture of paper and paper products	0.32	0.69	0.45	0.31	0.74	0.83	4.63	3.21	8.00	7.00
Publishing, printing and reproduction of recorded media	0.35	0.73	-0.17	0.18	0.67	0.37	0.78	-0.80	5.00	3.00
Manufacture of coke, refined petroleum products and nuclear fuel	0.25	0.54	0.07	0.31	0.64	0.59	1.38	5.40	17.00	-1.00
Manufacture of chemicals and chemical products	0.25	0.54	0.07	0.31	0.64	0.59	1.38	5.40	17.00	-1.00
Manufacture of rubber and plastics products	0.44	0.84	0.28	0.09	0.68	0.54	2.11	-0.80	11.00	6.00
Manufacture of other non-metallic mineral products	0.44	0.79	0.51	0.47	0.68	0.61	4.77	0.04	-1.00	-8.00
Manufacture of basic metals	0.43	0.86	0.45	0.41	0.95	0.78	-3.13	1.75	4.00	2.00
Manufacture of fabricated metal products, except machinery and equipment	0.39	0.72	0.21	0.46	0.51	0.77	7.47	-1.06	7.00	3.00
Manufacture of machinery and equipment n.e.c.	0.50	0.88	0.13	0.47	0.72	0.67	-5.31	2.77	7.00	4.00
Manufacture of office, accounting and computing machinery	0.42	0.84	0.29	0.49	0.45	0.73	2.67	3.36	6.00	2.00
Manufacture of electrical machinery and apparatus n.e.c.	0.10	0.39	0.05	0.03	-0.17	-0.04	2.94	1.03	0.00	0.00
Manufacture of radio, television and communication eq.	0.34	0.66	0.50	0.36	0.67	0.66	3.65	3.10	9.00	2.00
Manufacture of medical, precision and optical instruments	0.29	0.47	0.60	0.47	0.74	0.71	7.54	2.90	9.00	4.00
Manufacture of motor vehicles, trailers and semi-trailers	0.30	0.74	0.33	0.32	0.87	0.82	-2.94	-1.93	0.00	-6.00
Manufacture of other transport equipment	0.39	0.74	0.63	0.65	0.83	0.91	0.56	-7.15	0.00	-4.00
Manufacture of furniture; manufacturing n.e.c.	0.27	0.64	0.03	0.37	0.84	0.87	-1.21	3.03	8.00	2.00
Recycling	0.29	0.62	-0.22	-0.13	0.66	-0.62	5.53	2.84	-15.00	6.00
Electricity, gas, steam and hot water supply	0.02	0.02	0.51	0.09	0.38	0.49	-2.73	-1.48	0.00	-6.00
Collection, purification and distribution of water	0.22	0.74	0.25	-0.07	-0.81	0.55	-10.61	3.70	12.00	-17.00
	0.12	0.19	0.02	0.21	-0.23	0.20	7.51	10.51	0.00	4.00

Notes:

The variance (*var*) of the common component reports the fraction of the total variance that is explained by the static factor model. The correlation ρ and the time shift l are calculated for the cycle of the corresponding sector with respect to the aggregate cycle at the country level. The table reports the result for both the signal ϕ , which is extracted with the factor model, and the growth rates x^{13} . The aggregate measures without a bar summarize the results for the aggregate variables at the sectoral level. The aggregated measures with a bar summarize the results for the disaggregate variables. The disaggregate results are aggregated to the sectoral level using as aggregation weights $b_{i*|T}$.

Table continued.

	$\overline{var}_j(X)$	$\overline{var}_j(X)$	$\overline{P}_j(\phi)$	$\overline{P}_j(x^{13})$	$\rho_{x^j}^*(\phi)$	$\rho_{x^j}^*(x^{13})$	$\overline{I}_{x^j}(\phi)$	$\overline{I}_{x^j}(x^{13})$	$I_{x^j}^*(\phi)$	$I_{x^j}^*(x^{13})$
Agriculture, hunting and related service activities	0.29	0.20	0.19	0.10	0.87	0.88	3.22	-1.75	21.00	16.00
Construction	0.63	0.90	0.14	0.35	0.34	0.54	3.27	-0.52	1.00	-2.00
Sale, maintenance and repair of motor vehicles and motorcycles	0.19	0.26	0.21	0.34	0.57	0.77	6.54	5.64	9.00	11.00
Wholesale trade and commission trade	0.59	0.90	0.35	0.20	0.85	0.68	-1.86	1.53	5.00	0.00
Retail trade, except of motor vehicles and motorcycles	0.33	0.61	0.38	0.56	0.92	0.87	1.86	0.88	6.00	5.00
Hotels and restaurants	0.43	0.75	0.59	0.55	0.88	0.82	5.45	-1.12	2.00	3.00
Land transport; transport via pipelines	0.38	0.61	0.56	0.56	0.95	0.80	0.03	-0.81	-2.00	-12.00
Water transport	0.16	0.17	0.73	0.54	0.56	0.56	-8.00	-11.00	-8.00	-11.00
Air transport	0.18	0.15	0.78	0.48	0.79	0.49	18.90	0.11	19.00	0.00
Supporting and auxiliary transport activities	0.27	0.20	0.58	0.51	0.86	0.91	5.77	2.66	9.00	9.00
Post and telecommunications	0.25	0.33	0.55	0.29	0.87	0.55	3.63	-0.16	14.00	-2.00
Financial intermediation, except insurance and pension funding	0.61	0.88	-0.14	-0.11	-0.68	-0.26	1.73	3.11	9.00	6.00
Insurance and pension funding, except compulsory social security	0.29	0.47	0.32	0.23	0.68	0.52	-8.00	-1.44	-12.00	0.00
Activities auxiliary to financial intermediation	0.10	0.20	-0.09	0.06	-0.92	0.67	-2.29	-4.31	-25.00	2.00
Real estate activities	0.26	0.74	0.43	0.51	0.75	0.70	1.87	-1.73	8.00	1.00
Renting of machinery and equipment without operator	0.23	0.54	0.26	0.43	0.69	0.69	-0.54	0.27	4.00	0.00
Computer and related activities	0.21	0.45	0.21	0.63	0.93	0.94	-0.43	-1.35	1.00	-3.00
Research and development	0.26	0.69	0.20	0.32	0.28	0.51	3.52	1.53	3.00	0.00
Other business activities	0.39	0.67	0.00	0.16	0.78	0.56	3.76	3.26	-1.00	0.00
Public administration and defence; compulsory social security	0.55	0.80	0.31	0.24	0.93	0.69	-3.71	-2.49	-20.00	-9.00
Education	0.45	0.68	0.08	0.18	0.89	0.48	2.45	0.29	-26.00	-1.00
Health and social work	0.59	0.85	0.00	-0.52	-0.92	-0.91	-0.55	6.21	13.00	10.00
Activities of membership organizations n.e.c.	0.45	0.78	0.53	0.30	0.70	0.37	4.39	-0.84	6.00	-4.00
Recreational, cultural and sporting activities	0.16	0.46	0.09	0.11	0.16	0.24	1.03	0.92	0.00	-4.00
Other service activities	0.19	0.42	0.47	0.35	0.75	0.47	2.40	-0.68	8.00	-3.00
	0.37	0.66	0.25	0.15	0.62	0.73	5.85	-0.64	1.00	14.00

Notes:

The variance ($\overline{var}$) of the common component reports the fraction of the total variance that is explained by the static factor model. The correlation ρ and the time shift l are calculated for the cycle of the corresponding sector with respect to the aggregate cycle at the country level. The table reports the result for both the signal ϕ , which is extracted with the factor model, and the growth rates x^{13} . The aggregate measures without a bar summarize the results for the aggregate variables at the sectoral level. The aggregated measures with a bar summarize the results for the disaggregate variables. The disaggregate results are aggregated to the sectoral level using as aggregation weights $b_{i|j}^{17}$.

For $y = \{\hat{\phi}, x^{13}\}$ the reference cycle is the aggregate series y_t and each of the series $y_{ij,t}$, $y_{i*,t}$ and $y_{*j,t}$ can be classified as pro- or counter-cyclical according to the phase angle with the reference cycle at the zero frequency⁸. Tables 6.1 and 6.1 show at the regional respectively sectoral level both the optimal aggregate results $\{l_{i*}^*, \rho_{i*}^*\}$ and the aggregated optimal results $\{\bar{l}_{i*}^*, \bar{\rho}_{i*}^*\}$ for both $\{\hat{\phi}, x^{13}\}$. The first column of both tables reports the number of time series variables present in the corresponding region respectively sector. The subsequent two columns of both tables show respectively the weighted average of the variance of the common components $\overline{var}(\hat{\chi})$ and the variance of the aggregate common component $var(\hat{\chi})$. Both measures report the fraction of the variance explained by the static factor model. The difference between the two measures shows that the static factor model explains the covariation at the aggregate level much better than at the disaggregate level. The idiosyncratic motions of the variables die out in the aggregation as they are only weakly cross-correlated. The common factors explain on average 75% respectively 60% of the variation of the aggregate at the regional and sectoral level.

Table 6.1 reports the empirical results of the staffing labour cycle at the regional level. The four different measures for the lead almost always indicate the number of lead periods of $-6 < l^* < 6$. The regions that show a robust, but modest lead across the four different measures are Gelderland and Overijssel. The leading characteristics correspond with the relatively dominant presence of leading sectors like Wholesale and Manufacture of motor vehicles.

Table 6.1 reports the empirical results of the staffing labour cycle at the sectoral level. The differences across the sectors are more pronounced than across the regions according to the different statistical measures. The five sectors for which the variation is best explained by the common dynamics are: Manufacture of food products and beverages, Construction, Financial intermediation, Health and social work, and Wholesale trade and commission trade. Some sectors are more driven by idiosyncratic dynamics instead of by the common dynamics of the staffing labour cycle. Due to the hub function of the Netherlands for international freight flows in Europe, sectors like Air transport and Water transport are likely more reactive to world trade developments than to the national business cycle. The sectors Manufacture of tobacco products and Financial Intermediation show anti-cyclicity with respect to the aggregate

⁸We recall that the cross-spectral density between two variables h and j can be expressed, in its 'polar form' as $S_{hj}(\theta) = A_{hj}(\theta) e^{-i\phi_{hj}(\theta)}$ where $\phi_{hj}(\theta)$ is the 'phase'. The phase measures the angular shift between the cosine waves of h and j at frequency θ , $-\pi < \theta \leq \pi$. At frequency zero, the phase may be either 0 or π depending on whether the long-run correlation is positive or negative, respectively.

staffing labour cycle. The four different measures of the lead period almost always indicate a number of lead periods of $-26 < l^* < 17$. The variation in leading and lagging patterns is more pronounced across sectors than across regions. The two most leading sectors in addition to some Manufacturing sectors according to the four different measures are Supporting and auxiliary transport activities, and Sale, maintenance and repair of motor vehicles and motorcycles. Retail trade shows a modest lead of less than half a year. The latter two mentioned sectors' leading properties are confirmed by business cycle analysts. The variable of issued motor vehicle permits is incorporated in Nardo et al.'s (2008) composite leading indexes for many countries. Moreover, the retail sales are part of the total sales series, which is a key indicator in The Conference Board Index, (cf. McGuckin, 2001). The five manufacturing sectors of Electrical machinery and apparatus, Machinery and equipment and Wearing apparel, Wood and products of wood and Radio, television and communication also show a modest lead of less than half a year. The three sectors that show lagging characteristics across the four different measures are Public administration, defence and compulsory social security, Insurance and pension funding and Water transport.

6.5 Forecasting aggregate staffing employment

The natural question is how the identified leading indicators at the disaggregate level can be exploited to forecast the country aggregate of the staffing employment. This variable is closely monitored to discern patterns of the Dutch macroeconomic business cycle, cf. Goldschmeding (2003), Franses and de Groot (2005b) and den Reijer (2006). As the data set is closed in the sense that the country aggregate is by definition the sum of the disaggregates, we can moreover analyse the forecasting performance from a different perspective: does forecasting the aggregate improve upon aggregating the forecasts?

We refer to the h -period ahead forecast of the growth rate of the country aggregate based on observations until time t as $x_{t+h|t}$ and to the aggregated forecast as $\bar{x}_{t+h|t} = \sum_{ij} b_{ij|t} x_{ij,t+h|t}$. Moreover, the different forecasts are labelled as $x_{ij,t+h|t}^{m,fe}$, where the different model specifications $m = \{SF, DF, DFC, AR, \mu\}$ employ static factors (SF), one-sided dynamic factors (DF), one-sided cyclical dynamic factors (DFC), a second order autoregressive model (AR) and the first moment of the time series (μ), respectively. The latter two model specifications are purely univariate and act as a benchmark for multivariate factor specifications.

The one-sided dynamic factors f_t^{DF} are the contemporaneous weighted cross-sectional averages of the times series variables for which the weights depend on the common-to-idiosyncratic variance ratios as determined by Forni et al.'s (2005) two-step procedure. The two-sided filters in (6.2) render the dynamic method less suitable for forecasting purposes, since only a poor signal can be extracted at the end of the sample. However, the two-sided filters extract the temporal variation in the data to determine the in-sample common-to-idiosyncratic variance decomposition that can be exploited in a second step to determine the one-sided dynamic factors f_t^{DF} as the contemporaneously weighted cross-sectional averages of the times series variables, see appendix 6.A.3 for details. We employ Bai and Ng's (2002) information criteria (BNIC) to determine the number of $r = 2$ one-sided dynamic factors f_t^{DF} as a trade-off between the goodness-of-fit and overfitting. Finally, note that the second order autoregressive model (AR) is capable of capturing oscillatory motion. Before employing the univariate specifications $\{AR, \mu\}$, the observed time series x_{ijt} gets seasonally adjusted in an automated procedure if at least one of the seasonal dummies shows a significant coefficient at the 2.5% level.

The different forecasts $x_{ij,t+h|t}^{m,fe}$ differ not only with respect to the model specification m , but also regarding the forecast equation fe that relates the factors to the target variable. If the forecast equation $fe = \{u, fm\}$ admits the factor model structure $fe = fm$, the h -step ahead common components $\hat{X}_{ij,t+h}^m$ are projected on the t -dated estimated factors $\hat{f}_t^m$ and so, the autocorrelation structure of the common factor structure is exploited. The factor forecasts of the common components $\hat{X}_{ij,T+h|T}^m$ are transformed to forecasts of the target variable by inverse standardization, i.e. $\hat{x}_{ij,T+h|T}^m = \hat{\sigma}_{ij,T} \hat{X}_{ij,T+h|T}^m + \hat{\mu}_{ij,T}$, where $\hat{\mu}_{ij,T}$ and $\hat{\sigma}_{ij,T}$ are the sample mean and, respectively, the standard deviation of the univariate time series variable x_{ij} . If the forecast equation $fe = \{u, fm\}$ is unrestricted $fe = u$, the h -step ahead data $x_{ij,t+h}$ are projected on the t -dated estimated factors $\hat{f}_t^m$ and a constant. Boivin and Ng (2005) show that both approaches $fe = \{u, fm\}$ deliver identical forecasts only if the data adheres to the factor model assumptions and the model parameters are known, see also appendix 6.A.3.

6.5.1 Results

The out-of-sample forecasting exercise starts in 2002.9 and produces 32 forecasts⁹ for each horizon $h = 1, \dots, 13$. We perform the forecasting exercise using

⁹ All forecasts are incorporated in the evaluation since the country aggregate data are available until 2006.11

both a recursive estimation window and a rolling window of a size of 61 periods¹⁰. The forecasting performance of the different specifications is measured by the forecast error, i.e. the difference between the forecasts and the realizations. The summary statistics are the mean squared forecast error (*mse*) and the variance of the forecast error (*var*). The difference between these two statistics is the squared mean forecast error, i.e. the bias.

Along the lines of Eickmeier and Ziegler's (2008) meta-analytic analysis of the recent macroeconomic factor forecasting literature, we report the forecasting performance along four dimensions: 1) multivariate factor models $m = \{SF, DF, DFC\}$ vs. univariate models $m = \{AR, \mu\}$; 2) free vs. factor restricted forecast equation $fe = \{u, fm\}$; 3) aggregate forecasts $x_{t+h|t}$ vs. aggregating disaggregate forecasts $\bar{x}_{t+h|t}$; 4) recursive vs. rolling window estimation. The detailed forecasting performance statistics *mse* and *var* are reported along the four dimensions for each forecast horizon h in case of the recursive window in Table A.6.5 vs. the rolling window in Table A.6.6 in Appendix 6.A.4.

Table 6.3: The relative forecasting performance of different model specifications

Forecast specification	Ratio of average <i>mse</i>
$\frac{\text{multivariate } m=\{SF,DF,DFC\}}{\text{univariate } m=\{AR,\mu\}} \text{ model}$	0.993
$\frac{\text{free } fe=u}{\text{factor restricted } fe=fm} \text{ equation}$	0.997
$\frac{\text{aggregate } x_{t+h t}}{\text{aggregated } \bar{x}_{t+h t}} \text{ forecasts}$	0.964
$\frac{\text{rolling}}{\text{recursive}} \text{ window}$	0.976

Notes:

The table reports the ratio of mean squared errors (*mse*) that are averaged over the forecast horizon $h = 1, \dots, 13$ and the dimensions of the forecasting exercise. So, the *mse* in the first row is averaged over forecast horizon h and, moreover, over both specifications of the forecast equation *fe*, both aggregate and aggregated forecasts and both rolling and recursive estimation window. The results are based on the out-of-sample forecasting exercise, which starts in 2002.9 and produces 32 forecasts for each horizon.

¹⁰So, the first forecasting round is based on the sample that runs from 1998.1 until 2002.9 and consists of 61 periods. The recursive and the rolling window forecasts are by construction identical in the first round.

Table 6.3 summarizes these detailed statistics on the forecasting performance by averaging the detailed *mse* statistics over forecast horizon and over the relevant dimensions of the forecasting exercise. For example, the first row of Table 6.3 reports the ratio of the *mse* that is the average of the multivariate specification $m = \{SF, DF, DFC\}$ against its univariate $m = \{AR, \mu\}$ equivalent. Moreover, both multivariate and univariate *mse* are averaged over forecast horizon h and over the three remaining dimensions of the forecasting exercise: the two specifications of the forecast equation $fe = \{u, fm\}$, aggregate $x_{t+h|t}$ and aggregated $\bar{x}_{t+h|t}$ forecasts and both rolling and recursive estimation window.

The ratio of *mse* of the multivariate specification versus its univariate equivalent¹¹ shows some value added of forecasting based on a large cross-section of data. As we will show, the slight outperformance masks substantial differences among the multivariate specifications. The ratio of *mse* between the two specifications of $fe = \{u, fm\}$ confirms Boivin and Ng's (2005) notion that the unrestricted forecast equation performs better. The slight outperformance again masks the observation that the outperformance is more substantial for the *DF* specification. The ratio of *mse* between the aggregate x and aggregated $\bar{x}$ forecasts shows a preference for forecasting the aggregate directly. For the recursive window method, forecasting the aggregate performs on average better than aggregating the forecasts, while the rolling window method shows no distinctive differences. One explanation for the result is that the time-varying and for outlier-corrected aggregation weights $a_{ij,t}$ might be biased, see appendix 6.A.2. Note that forecasting the aggregate employing the factor models $m = \{SF, DF, DFC\}$ incorporates the disaggregate information as contained by the factors. So, these results confirm Hendry and Hubrich's (2006) theoretical result on predictability that forecasting the aggregate using disaggregate information outperforms aggregating the forecasted disaggregates. Finally, the ratio of *mse* between the rolling vs. the recursive window shows a slight preference for the rolling window estimation. This result is consistent with Eickmeier and Ziegler's (2008) analysis of the empirical macroeconomic factor forecasting literature. Apparently, the data shows substantial temporal variation, which is accounted for by employing a rolling window scheme that allows the model specification to adapt more flexibly to the data and so overcompensates the gains from using long time series.

¹¹Note that we reported the 2nd order autoregressive process *AR* as being the best performing one. An example of an alternative specification is an $AR(p)$ process for which the lag length p is determined by the Akaike information criteria in an automated procedure. Other examples are autoregressive specifications of x_{ijt}^{13} with or without including a lagged term at $(t-13)$, i.e. $\alpha_{13}x_{ijt-13}^{13}$.

Table 6.4: The relative forecasting performance of the dynamic factor specification

h	p-values				mse/mse*			
	SF	DFC	AR	μ	SF	DFC	AR	μ
1	0.01	0.05	0.60	0.72	1.31	1.12	0.97	0.98
2	0.03	0.68	0.64	0.86	1.09	0.98	0.97	0.99
3	0.31	0.00	0.04	0.34	1.09	1.06	1.06	1.03
4	0.00	0.00	0.07	0.14	1.26	1.18	1.06	1.05
5	0.73	0.58	0.04	0.78	1.02	1.03	0.89	1.02
6	0.01	0.07	0.01	0.01	1.85	1.24	1.32	1.28
7	0.00	0.00	0.02	0.01	1.52	1.49	1.45	1.47
8	0.00	0.13	0.11	0.05	1.20	1.10	1.08	1.10
9	0.00	0.02	0.00	0.08	1.16	1.07	1.13	1.08
10	0.06	0.00	0.47	0.42	1.08	1.15	1.03	1.04
11	0.83	0.08	0.47	0.97	1.01	1.12	1.05	1.00
12	0.08	0.56	0.13	0.17	1.35	0.95	1.20	1.16
13	0.09	0.00	0.00	0.00	1.08	1.55	1.90	2.01

Notes:

All specifications in this table correspond to the unrestricted forecast equation $fe = u$, aggregate forecasts $x_{t+h|t}$ and rolling window estimation. The reported p-values for each forecast horizon $h = 1, \dots, 13$ correspond to the Harvey et al.'s (1997) small-sample modification of Diebold and Mariano's (1995) (DM) test statistic of equal forecast accuracy. Moreover, the table reports the ratio of mean squared errors (*mse*) of the model specification presented in the second row of the table vs. the *mse** equivalent of the dynamic factor model (DF) specification. The results are based on the out-of-sample forecasting exercise, which starts in 2002.9 and produces 32 forecasts for each horizon.

Following upon the results of table 6.3, the remaining analysis is based on aggregate forecasts x based on a rolling window scheme and an unrestricted forecast equation $fe = u$. Table 6.4 reports for every forecast horizon h , the relative *mse* and the *p*-values of the respective reported model specifications against the *DF* specification. The *p*-values correspond to Harvey et al.'s (1997) small-sample modification of Diebold and Mariano's (1995) (DM) test of equal forecast accuracy. Table 6.4 shows that the *DF* specification outperforms its competitor specifications significantly so at most forecast horizons. As shown in appendix 5.A.1, the *DF* specification nests the *SF* specification, which does not exploit the dynamic structure of the data. So, the outperformance by the *DF* model implies that the underlying data exhibit substantial

dynamics. Moreover, the *DFC* specification nests the *DF* specification, which can be obtained by setting $\tau = 1$ in (5.3) and so, does not filter out any short-lived cyclical. The outperformance of the *DF* over the *DFC* specification shows that exploiting only the cyclical common dynamics results in a loss of information as the observed staffing employment series exhibits a substantial seasonal pattern.

6.6 Conclusion

This paper analyses the developments of the staffing labour cycle in the Netherlands at a disaggregate level using the data set from Randstad. We create a balanced data set that describes the number of hours of staffing employment for 15 different regions and 58 different sectors. We apply factor models to extract low-dimensional common information from the data set and show that the extracted signal resembles the year-on-year growth rate of aggregate staffing employment. The common signal, which excludes the effects of sector or region specific shocks, is also extracted at the disaggregate level. We analyse the correlation structure and classify the disaggregate cycles as leading and lagging according to eight empirical measures. Almost all regions lead or lag the staffing labour cycle by less than half a year. The regions, whose modest leading characteristics are robust across four different empirical measures, are Gelderland and Overijssel.

Almost all sectors show a lead that lies between -2 years and +1.5 years. The differences across the sectors are more pronounced than across the regions. Three leading sectors, whose leading characteristics are robust across four different empirical measures, are Supporting and auxiliary transport activities, Sale, maintenance and repair of motor vehicles and motorcycles, and Retail trade. The turnover in the latter two sectors are known to be stylized business cycle leading indicators.

We then explored how the identified leading indicators at the disaggregate level can be exploited to forecast the country aggregate of the staffing employment. The final section compares different model specifications that employ static factors, dynamic factors, cyclical dynamic factors, a second order autoregressive model and the first moment of the time series. The performance is measured by the forecast bias and the mean squared forecast error. Detailed information does not necessarily improve the forecasting performance as the static factor model does not outperform the benchmark autoregressive model. Due to substantial temporal and seasonal variation in the staffing labour market, the dynamic factor model manages to outperform the univariate benchmark forecasting models.

6.A Appendix

6.A.1 Data

The data set consists of 1.276.393 observations on the number of contracted staffing hours employed by Randstad. An observation consists of the number of paid hours of contracted staffing agency work. The number of paid hours is larger than the number of invoiced hours since for instance holiday hours and sickness leave are included in addition to worked hours. Every observation consists of four dimensions: the number of paid hours of contracted staffing agency work, the time dimension denoted by the year and the period, the sector to which the user company belongs and the geographical area where the user firm is located. So, the sector and location refers to the company at which the staffing agency worker is employed. The user firm that hires a staffing employee is located in a geographical area that is classified according to the 4-digit postal code. The data consists of 2882 different postal codes that can be aggregated to 466 different municipalities and 15 different regions. The time dimension of the data is divided into years and periods. Every year consists of 13 subsequent administrative periods of 4 weeks. Observations are available from the first period of 1998 until and including the first period of 2005.

Statistics Netherlands (*Centraal Bureau voor de Statistiek*, CBS) employs a systematic hierarchical classification system for economic activities called the *Standaard Bedrijfsindeling* (SBI). The currently operational SBI '93 was settled in 1993 and consists of five levels. On the highest 2-digits level of compartments, the classification consists of 58 different sectors, which are not all present in the Netherlands¹². The SBI-code equals for the highest 4-digits level Eurostat's *Nomenclature statistique des activités économiques dans la Communauté Européenne* (NACE, in particular Rev. 1). On the highest 2-digits level, the SBI'93 and NACE Rev.1 are compatible with the International Standard Industrial Classification of All Economic Activities (ISIC, in particular Rev 3.1), which is the recommended classification of economic activities as established in March 2002 by the statistics commission of the United Nations.

Statistics Netherlands employs a systematic hierarchical classification system for regional units called the *Nomenclatuur van Territoriale eenheden voor de Statistiek* (NUTS). The system provides a non-overlapping countrywide division of the Netherlands in 40 regional units (NUTS3), which can be aggregated to 12 provinces (NUTS2). In this study, we employ 15 regions that correspond

¹²Compartments describing activities that are not present in the Netherlands are for instance the "mining of uranium and thorium ores" and the "mining of metal ores".

to the 12 provinces from which the agglomerations of the three largest cities are separated out. The three NUTS3 regional units *Groot-Amsterdam*, *Groot-Rijmond* and *Agglomeratie 's Gravenhage* correspond with the agglomerations of the three biggest cities Amsterdam, Rotterdam and the Hague, respectively.

6.A.2 On aggregation

The balanced data set consists of the series $X_{ij,t}$, which are the total number of hours of staffing employment in region $i = 1, \dots, 15$ and sector $j = 1, \dots, 58$ during period t , running from 1998.1 until 2005.1 consisting of 92 four-weeks periods. The time series of aggregate staffing employment at the regional level consists of $X_{i*,t} = \sum_j X_{ij,t}$, the sectoral level equivalent $X_{*j,t} = \sum_i X_{ij,t}$ and the time series of total staffing employment in the Netherlands equals $X_t = \sum_i \sum_j X_{ij,t}$. Let the time-varying shares of the individual variables in the aggregate be defined as: $\alpha_{ij,t} = \frac{X_{ij,t}}{\sum_i \sum_j X_{ij,t}}$. Lemma (1) shows that calculating the growth rates x_t of the aggregate X_t is equivalent to aggregating the growth rates of the disaggregates $x_{ij,t}$ using the delayed weights $\alpha_{ij,t-1}$.

1. LEMMA. Let $X_t = X_{1,t} + X_{2,t}$ be a time series variable. Let x_t be the growth rate of X_t and let $b_{i,t} = \frac{X_{i,t}}{X_t}$. Then $x_t = b_{1,t-1}x_{1,t} + b_{2,t-1}x_{2,t}$

Proof. $\Delta \ln(X_t) \approx x_t = \frac{\Delta X_t}{X_{t-1}} = \frac{\Delta X_{1,t} + \Delta X_{2,t}}{X_{1,t-1} + X_{2,t-1}} = \frac{\Delta X_{1,t}}{X_{1,t-1}} \frac{X_{1,t-1}}{X_{1,t-1} + X_{2,t-1}} + \frac{\Delta X_{2,t}}{X_{2,t-1}} \frac{X_{2,t-1}}{X_{1,t-1} + X_{2,t-1}} = x_{1,t} \left(\frac{X_{1,t-1}}{X_{1,t-1} + X_{2,t-1}} \right) + x_{2,t} \left(\frac{X_{2,t-1}}{X_{1,t-1} + X_{2,t-1}} \right) = b_{1,t-1}x_{1,t} + b_{2,t-1}x_{2,t}$. ■

The dynamic factor model decomposition is only defined for data sets that consist of a panel of stationary time series.

2. LEMMA. Let $X_t = X_{1,t} + X_{2,t}$ be a first-order integrated, $I(1)$, time series variable. For $i = 1, 2$ let x_t and $x_{i,t}$ be the corresponding growth rates, x_t^s and $x_{i,t}^s$ the standardized growth rates with means μ_x and μ_{x_i} and standard deviations σ_x and σ_{x_i} respectively. Let the time-varying weights be $b_{i,t} = \frac{X_{i,t}}{X_t}$. Then $x_t^s = a_{1,t-1}x_{1,t}^s + a_{2,t-1}x_{2,t}^s$ with $a_{i,t} = b_{i,t} \frac{\sigma_{x_i}}{\sigma_x}$.

Proof. Lemma (1) shows that $\sigma_x x_t^s + \mu_x = b_{1,t-1}(\sigma_{x_1} x_{1,t}^s + \mu_{x_1}) + b_{2,t-1}(\sigma_{x_2} x_{2,t}^s + \mu_{x_2})$. Rewriting this equation leads to $x_t^s = b_{1,t-1} \frac{\sigma_{x_1}}{\sigma_x} x_{1,t}^s + b_{2,t-1} \frac{\sigma_{x_2}}{\sigma_x} x_{2,t}^s$ and $\mu_x = b_{1,t-1}\mu_{x_1} + b_{2,t-1}\mu_{x_2}$. ■

Projecting the aggregate x_t^s on the estimated factors is only mathematically equivalent to aggregating the projected disaggregates $x_{ij,t}^s$ if the weights

are constant $a_{ij,t} = a_{ij}$. Let x_{it}^s denote a vector of T observations for $i = 1, 2$ and u a $(T \times q)$ -matrix of q common components u_{it} . Say, you have $x_t^s = a_1 x_{1t}^s + a_2 x_{2t}^s$, then $\chi_t' = (u_t' u_t)^{-1} u_t' x_t^s u_t' = (u_t' u_t)^{-1} u_t' (a_1 x_{1t}^s + a_2 x_{2t}^s) u_t' = a_1 (u_t' u_t)^{-1} u_t' x_{1t}^s u_t' + a_2 (u_t' u_t)^{-1} u_t' x_{2t}^s u_t' = a_1 \chi_{1t}' + a_2 \chi_{2t}'$.

The assumptions needed for this mathematical equivalence are not necessarily met in practice. Firstly, the weights a_i are time varying. Secondly, lemma (1) does not hold precisely since the disaggregate growth rates $x_{i,t}$ and the aggregate growth rate x_t are independently corrected for outliers. Thirdly, lemma (2) only holds in population as the estimated variances may vary over sample, i.e. $\hat{\sigma}_{x|1,\dots,t} \neq \hat{\sigma}_{x|1,\dots,t-1}$.

6.A.3 The Estimator

In this appendix, we show in more detail how the common component χ can be estimated in a stepwise procedure. Moreover, we will highlight the parameter condition that makes the static factor model a special case of the dynamic one. Finally, we show in more detail the estimator for the factor model forecasts in case the forecast equation is restricted to admit the factor model structure, i.e. $fe = fm$.

The dynamic method

The dynamic method as outlined in Forni et al. (2000; 2001; 2001a; 2004; 2005) (FHLR) consists of the frequency-domain counterpart of the static method. The dynamic factors $\mathbf{u}_t = (u_{1t} \dots u_{qt})'$ are estimated by the dynamic principal components, which are the static principal components of the spectral density matrix as outlined by Brillinger (1981). Denote by $\mathbf{X}^{nT} = (x_{it})_{i=1\dots n, t=1\dots T}$ an $n \times T$ rectangular of observations¹³, which are realisations of real-valued stationary stochastic processes $\{\mathbf{x}_t = (x_{1t} \dots x_{nt})'\}$.

Let $\hat{\Gamma}_{\mathbf{X}}^{nT}(k) = \frac{1}{T-k} \sum_{t=k+1}^T \mathbf{x}_t^{nT} \mathbf{x}_{t-k}^{nT'}$ be the k -lag sample covariance of $\mathbf{x}_t^{nT}$. FHLR suggest the following stepwise procedure: (i) estimate the spectral density matrix (cf. Brillinger, 1981) of $\mathbf{X}^{nT}$ as $\Sigma_{\mathbf{X}}^{nT}(\theta_h) = \sum_{k=-M}^M \hat{\Gamma}_{\mathbf{X}}^{nT}(k) \omega_k e^{-ik\theta_h}$, $\theta_h = 2\pi h / (2M+1)$, $h = 0, \dots, 2M$, where $\omega_k = 1 - |k| / (M+1)$ is the Bartlett window of size M . Like Forni et al. (2000), we set $M(T) = \text{round}(2T^{(1/2)})$ such that the convergence rate is $M(T)/T = O(T^{(1/2)})$;

¹³The notation is simplified in that the time series variables are indexed by the single subscript i instead of by the sector-region coordinates ij as in the main text. Moreover, the additional superscript nT denotes the sample size.

(ii) calculate from the spectral density matrix $\Sigma_X^{nT}(\theta_h)$ the q largest dynamic eigenvalues $\lambda_j^{nT}(\theta_h)$ and the corresponding dynamic eigenvectors $\mathbf{p}_j^{nT}(\theta_h)$, $j = 1, \dots, q$. for $h = 0, \dots, 2M$. We follow Forni et al.'s (2000) approach and select $q = 3$ in a finite-sample such that the marginal explained variance of the q^{th} dynamic eigenvalue is larger than 10% and the $(q + 1)^{th}$ equivalent is smaller than 10%;

(iii) let $\mathbf{p}_q^{nT}(\theta_h) = (\mathbf{p}_1^{nT'}(\theta_h) \dots \mathbf{p}_q^{nT'}(\theta_h))'$ the $(q \times n)$ -matrix of dynamic eigenvectors and $\mathbf{\Lambda}_q^{nT}(\theta_h)$ a diagonal matrix with the q largest dynamic eigenvalues on the diagonal. Inverse Fourier transformation of

$\hat{\Sigma}_X^{nT}(\theta_h) = \mathbf{p}_q^{nT*}(\theta_h) \mathbf{\Lambda}_q^{nT}(\theta_h) \mathbf{p}_q^{nT}(\theta_h)$ (* denotes complex conjugate) results in the correlation matrix of the common component

$$\hat{\Gamma}_X^{nT}(k) = \frac{1}{(2M+1)} \sum_{k=-M}^M \hat{\Sigma}_X^{nT}(\theta_h) \omega_k e^{ik\theta_h} \text{ for } h = 0, \dots, 2M. \text{ Moreover, the es-}$$

timated common dynamic factors are $\hat{\mathbf{u}}_t^{nT} = \frac{1}{(2M+1)} \sum_{k=-M}^M \sum_{h=0}^{2M} \mathbf{p}_q^{nT}(\theta_h) e^{ik\theta_h} \mathbf{x}_{t-k}^{nT}$.

Projecting the data on the common dynamic factors gives the estimator of the cyclical medium- and long-run common component::

$$\hat{\Phi}_{nt}^{nT} = \frac{1}{(2M+1)} \sum_{k=-M}^M \sum_{h=0}^{2M} \beta_k \mathbf{p}_q^{nT*}(\theta_h) \mathbf{p}_q^{nT}(\theta_h) e^{ik\theta_h} \mathbf{x}_{t-k}^{nT}, \quad (6.5)$$

where the band-pass filter coefficients $\beta(L)$ are defined in (6.4). The finite sample approximation consists of truncating the tails of the filter that involve unavailable data observations, so $\beta_k = 0$ for $k > M$.

(iv) repeat step (iii) using the $(q + 1)$ to n ordered eigenvalues to obtain $\hat{\Gamma}_\xi^{nT}(k)$;

(v) let $\hat{\mathbf{S}}^{nT} = (\hat{\mathbf{S}}_1^{nT'} \dots \hat{\mathbf{S}}_r^{nT'})'$ the $(r \times n)$ -matrix containing the r generalized eigenvectors of the couple of matrices $(\hat{\Gamma}_X^{nT}(0), \hat{\Gamma}_\xi^{nT}(0))$ with the normalization that $\hat{\mathbf{S}}_i^{nT'} \text{diag}(\hat{\Gamma}_\xi^{nT}(0)) \hat{\mathbf{S}}_j^{nT'} = 1$ if $i = j$ and zero otherwise. We use Bai and Ng's (2002) information criteria (BNIC) to determine the r generalized static factors as a trade-off between the goodness-of-fit and overfitting. The factors can then be estimated by the generalized principal components, i.e. $\hat{\mathbf{F}}^{nT} = \hat{\mathbf{S}}^{nT} \mathbf{X}^{nT}$, with $\hat{\mathbf{F}}^{nT} = (\hat{\mathbf{f}}_1^{nT} \dots \hat{\mathbf{f}}_T^{nT})$ a $(r \times T)$ -matrix of the stacked estimated factors;

(vi) let $\tilde{\chi}_{i,T+h|T}^{nT}$ be the h -step ahead factor forecasts of the common component of the i -th variable given T observations of n time series variables. The forecasts for the dynamic common component can be obtained by project-

ing the $(t + h)$ –dated unobserved common component $\hat{\chi}_{t+h}^{nT}$ on the t -dated factors $\hat{\mathbf{x}}_t^{nT}$, which for variable i results in:

$$\hat{\chi}_{i,T+h|T}^{nT} = \left[\hat{\Gamma}_{\mathbf{x}}^{nT}(h) \right]_i \hat{\mathbf{S}}^{nT'} \left(\hat{\mathbf{S}}^{nT} \hat{\Gamma}_{\mathbf{x}}^{nT}(0) \hat{\mathbf{S}}^{nT'} \right)^{-1} \hat{\mathbf{S}}^{nT} \mathbf{x}^{nT} \quad (6.6)$$

Evidently, the in-sample estimator for the common component can be obtained by setting $h = 0$.

Step (i) until step (iii) allow to estimate the dynamic factor model. The estimated cyclical common component $\hat{\phi}_{nt}$ is calculated by applying time filters to the x 's before averaging along the cross-section. The dynamic estimation method consists of two-sided filters and cannot be applied at the end of the sample, which is the most important part for forecasting. By truncating the time filters, the performance of the estimator deteriorates as t approaches T . Therefore in step (v), FHLR construct generalized principal components $\mathbf{F}^{nT}$, which are contemporaneous averages of $\mathbf{X}^{nT}$ that minimize the ratio of the variance of the idiosyncratic to common component.

The static method

Evidently, Stock and Watson's (2002a) static factors obtain as a special case of $\hat{\mathbf{F}}^{nT}$ with $M = 0$ and assuming in step (v) that $\text{diag} \left(\hat{\Gamma}_{\mathbf{x}}^{nT}(0) \right) = I_N$, i.e. the identity matrix. Computing the generalized principal components of $\mathbf{x}_{nt}$ is equivalent to computing the standard principal components of $\mathbf{y}_{nt} = H\mathbf{x}_{nt}$ with $\det(H) \neq 0$ and H such that $H\mathbf{x}_{nt}\mathbf{x}_{nt}'H'$ is the $n \times n$ -identity matrix. When the idiosyncratic variance-covariance matrix is diagonal, the normalization amounts to dividing each of the x 's by the standard deviation of its idiosyncratic component.

Factor based forecasting

If the forecast equation $fe = \{u, fm\}$ admits the factor model structure $fe = fm$, the h -step ahead forecasts for the common component $\hat{\chi}_{i,T+h|T}^{nT}$ is obtained by (6.6). The forecasts for the target variable are then obtained by inverse standardization, i.e. $\hat{y}_{i,T+h|T} = \hat{\sigma}_{iT} \hat{\chi}_{ij,T+h|T}^m + \hat{\mu}_{iT}$, where $\hat{\mu}_{iT}$ and $\hat{\sigma}_{iT}$ are the sample mean and, respectively, the standard deviation of the i th variable y_i ¹⁴.

If the forecast equation $fe = \{u, fm\}$ is unrestricted $fe = u$, the h -step ahead forecasts are obtained from an unrestricted forecast equation: $\hat{y}_{i,T+h|T} = \hat{\mu}_{iT,h} + \hat{\theta}_{iT,h} \hat{\mathbf{f}}_T$. The parameters $\hat{\mu}_{iT,h}$, $\hat{\theta}_{iT,h}$ are obtained by a linear regression

¹⁴So, we refine the notation and introduce the time series variable y_i whose standardized equivalent is x_i .

of $x_{i,t+h}$ on the estimated factors $\hat{\mathbf{f}}_t$ and a constant. The orthogonal projection of the $(t+h)$ -dated variable $y_{i,t+h}$ on the t -dated factors results in $\theta_{iT,h} = \sigma_{iT} [\Gamma_{\mathbf{X}}^{nT}(h)] \mathbf{S}^{nT'} \left(\mathbf{S}^{nT'} \Gamma_{\mathbf{X}}^{nT}(0) \mathbf{S}^{nT'} \right)^{-1}$. This expression for $\theta_{iT,h}$ differs in two respects from its equivalent for the restricted factor model (6.6). First, the data are directly projected on the estimated factors, i.e. $\Gamma_{\mathbf{X}}^{nT}(h)$ instead of imposing the factor decomposition (6.1) and therefore only project the common component $\hat{\Gamma}_{\mathbf{X}}^{nT}(h)$. Second, projecting $y_{i,t+h}$ instead of its standardized equivalent $x_{i,t+h}$ makes the corresponding standard deviation σ_{iT} appear in the expression.

So as shown by Boivin and Ng (2005), the two different forecast equations $fe = \{u, fm\}$ would provide identical forecasts in population if the data admits the factor decomposition (6.1). Moreover, identical forecasts in sample would require that the estimated constant of the linear regression equals the in-sample standard deviation: $\hat{\mu}_{iT,h} = \hat{\mu}_{iT}$ and that the estimate for the parameter of the linear regression model $\theta_{iT,h}$ equals:

$$\hat{\theta}_{iT,h} = \hat{\sigma}_{iT} \left[\hat{\Gamma}_{\mathbf{X}}^{nT}(h) \right]_i \hat{\mathbf{S}}^{nT'} \left(\hat{\mathbf{S}}^{nT} \hat{\Gamma}_{\mathbf{X}}^{nT}(0) \hat{\mathbf{S}}^{nT'} \right)^{-1}.$$

6.A.4 Empirical results

The detailed results of the forecasting exercise are reported along four dimensions: multivariate factor models $m = \{SF, DF, DFC\}$ vs. univariate models $m = \{AR, \mu\}$, free vs. factor restricted forecast equation $fe = \{u, fm\}$, aggregate forecasts $x_{t+h|t}$ vs. aggregated disaggregate forecasts $\bar{x}_{t+h|t}$ and recursive vs. rolling window estimation. The forecasting performance along the first three dimensions are reported in Table A.6.5 and Table A.6.6 employing recursive respectively rolling window estimation. The first column reports for each forecast horizon $h = 1, \dots, 13$ the forecasting performance of the AR-model. The upper part of the column reports the *mse* and the lower part the *var* of the forecast errors. The other columns report the forecasting performance of the respective models as a ratio to the performance of the benchmark AR-model.

Table A.6.5: Forecasting performance of different models using recursive windows

h	$AR_{t+h t}^m$ mse	$\mu_{t+h t}^m$ ratio	$SE_{t+h t}^m$ ratio	$SE_{t+h t}^{f,m}$ ratio	$DF_{t+h t}^m$ ratio	$DF_{t+h t}^{f,m}$ ratio	$DFC_{t+h t}^m$ ratio	$DFC_{t+h t}^{f,m}$ ratio	$\overline{AR}_{t+h t}^m$ ratio	$\overline{\mu}_{t+h t}^m$ ratio	$\overline{SE}_{t+h t}^m$ ratio	$\overline{SE}_{t+h t}^{f,m}$ ratio	$\overline{DF}_{t+h t}^m$ ratio	$\overline{DF}_{t+h t}^{f,m}$ ratio	$\overline{DFC}_{t+h t}^m$ ratio	$\overline{DFC}_{t+h t}^{f,m}$ ratio
1	0.75	1.04	1.39	1.36	1.09	1.08	1.13	1.07	1.04	1.04	1.36	1.37	1.05	1.04	1.09	1.06
2	0.75	1.03	1.09	1.08	0.92	0.92	1.05	1.05	1.01	1.03	1.08	1.08	0.91	0.91	1.04	1.04
3	0.75	1.02	1.07	1.06	0.98	0.98	1.04	1.04	1.01	1.02	1.06	1.06	0.98	0.98	1.04	1.04
4	0.75	1.01	1.20	1.20	0.97	0.97	1.07	1.02	0.98	1.01	1.19	1.20	0.95	0.95	1.08	1.02
5	0.80	0.98	1.00	0.98	0.89	0.89	0.97	0.98	0.97	0.99	0.99	0.99	0.90	0.90	0.98	0.99
6	0.80	0.97	1.35	1.30	0.77	0.79	0.95	0.98	0.88	0.98	0.99	1.32	0.77	0.79	0.96	0.99
7	0.75	1.00	0.98	1.03	0.63	0.68	0.80	1.02	0.93	1.00	1.07	1.04	0.64	0.69	0.82	1.03
8	0.75	0.99	1.06	1.05	0.50	0.92	1.01	1.00	1.01	1.00	1.06	1.07	0.90	0.92	1.02	1.01
9	0.77	0.99	1.01	1.02	0.86	0.99	0.99	0.99	0.98	0.99	1.04	1.04	0.85	0.89	0.99	1.00
10	0.74	1.03	1.05	1.06	0.98	0.98	1.05	1.04	1.05	1.04	1.06	1.07	0.98	0.99	1.06	1.05
11	0.78	0.97	1.00	1.00	0.97	0.96	1.10	0.98	0.96	0.98	1.01	1.01	0.98	0.97	1.09	0.99
12	0.74	0.96	1.07	1.10	0.82	0.83	0.90	0.97	0.65	0.97	1.09	1.09	0.79	0.84	0.86	0.98
13	0.79	1.00	0.55	0.56	0.52	0.74	0.60	0.99	0.61	1.00	0.54	0.54	0.52	0.75	0.60	0.99
h	mse	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio
1	0.75	1.04	1.39	1.36	1.09	1.08	1.13	1.06	1.04	1.04	1.36	1.37	1.05	1.04	1.09	1.06
2	0.75	1.03	1.09	1.08	0.92	0.92	1.05	1.04	1.01	1.03	1.08	1.08	0.91	0.90	1.04	1.04
3	0.75	1.02	1.07	1.06	0.98	0.98	1.04	1.04	1.01	1.02	1.06	1.06	0.98	0.97	1.04	1.04
4	0.75	1.01	1.21	1.20	0.97	0.97	1.07	1.02	0.98	1.01	1.19	1.20	0.95	0.95	1.08	1.02
5	0.79	0.99	1.00	0.98	0.89	0.89	0.98	0.99	0.96	0.99	0.99	0.99	0.91	0.90	0.98	0.99
6	0.78	0.97	1.37	1.32	0.79	0.80	0.97	0.97	0.85	0.97	1.33	1.32	0.78	0.79	0.97	1.03
7	0.73	1.01	1.00	1.04	0.63	0.69	0.82	1.03	0.91	1.01	1.08	1.05	0.64	0.69	0.83	1.03
8	0.73	1.01	1.08	1.06	0.92	0.93	1.01	1.02	0.99	1.01	1.06	1.07	0.92	0.93	1.02	1.02
9	0.76	1.00	1.02	1.03	0.87	0.90	1.00	1.00	0.97	1.00	1.04	1.04	0.86	0.90	1.00	1.00
10	0.73	1.04	1.06	1.07	0.99	0.99	1.07	1.04	1.04	1.04	1.07	1.07	0.99	0.99	1.07	1.04
11	0.76	0.99	1.01	1.01	0.98	0.97	1.11	0.99	0.95	0.99	1.02	1.02	0.98	0.97	1.10	0.99
12	0.73	0.97	1.08	1.11	0.81	0.84	0.88	0.98	0.64	0.97	1.10	1.10	0.78	0.83	0.84	0.98
13	0.79	1.00	0.54	0.55	0.51	0.74	0.56	0.99	0.60	1.00	0.54	0.54	0.51	0.75	0.56	0.99

Notes:

$x_{t+h|t}^{m,f,e}$, represents the h -step ahead forecast with the different model specifications $m = \{SE, DF, DFC, AR, \mu\}$. The specifications employ respectively static factors, dynamic factors, cyclical dynamic factors, 2nd order autoregressive model and the first moment of the time series. Moreover, the forecast equation $f^e = \{u, f^m\}$ is unrestricted such that the time series variables are forecasted directly or, respectively, admits the factor model structure such that the common component is forecast. The first column reports for each forecast horizon h the forecast performance of the AR-model. The upper part of the column reports the mean squared error (*mse*) and the lower part the variance (*var*) of the forecast errors. The other columns report the forecasting performance of the respective models as a ratio to the performance of the benchmark AR-model. The forecasting exercise starts in 2002.9 and produces 32 forecasts for each horizon. All forecasts are evaluated since the aggregate data at the country level are available until 2006.4. The forecasts are generated using a recursive estimation window.

Table A.6.6: Forecasting performance of different models using rolling windows

h	$x_{t+h t}^{AR,u}$	$x_{t+h t}^{\mu,u}$	$x_{t+h t}^{SF,u}$	$x_{t+h t}^{SF,m}$	$x_{t+h t}^{DF,u}$	$x_{t+h t}^{DF,m}$	$x_{t+h t}^{AR,u}$	$x_{t+h t}^{\mu,u}$	$x_{t+h t}^{SF,u}$	$x_{t+h t}^{SF,m}$	$x_{t+h t}^{DF,u}$	$x_{t+h t}^{DF,m}$	$x_{t+h t}^{AR,u}$	$x_{t+h t}^{\mu,u}$	$x_{t+h t}^{SF,u}$	$x_{t+h t}^{SF,m}$	$x_{t+h t}^{DF,u}$	$x_{t+h t}^{DF,m}$
	mse	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio
1	0.76	1.01	1.35	1.03	1.03	1.16	1.07	1.01	1.02	1.32	1.01	1.01	1.01	1.02	1.34	1.00	1.13	1.06
2	0.75	1.02	1.12	1.03	1.02	1.01	1.04	1.02	1.02	1.10	1.09	1.01	1.00	1.00	1.09	1.01	1.00	1.04
3	0.79	0.97	1.03	0.94	0.94	0.99	0.98	0.96	0.96	1.02	1.02	0.93	0.99	0.98	0.99	0.99	0.99	0.98
4	0.77	0.99	1.19	0.94	0.95	1.11	1.02	0.98	1.00	1.16	1.17	0.94	1.11	1.02	1.17	0.94	1.11	1.02
5	0.70	1.14	1.14	1.12	1.10	1.16	1.16	1.11	1.16	1.15	1.15	1.13	1.11	1.16	1.15	1.13	1.11	1.16
6	0.81	0.98	1.41	0.76	0.77	0.94	0.98	0.94	0.97	1.35	1.33	0.77	0.78	0.96	1.33	0.77	0.96	0.98
7	0.73	1.02	1.05	0.69	0.73	1.03	1.04	1.00	1.02	1.09	1.06	0.70	1.00	1.04	1.09	0.70	1.00	1.04
8	0.74	1.02	1.11	0.93	0.95	1.02	1.04	1.03	1.02	1.10	1.11	0.93	0.96	1.01	1.10	0.93	1.01	1.04
9	0.80	0.96	1.04	0.89	0.90	0.95	0.97	0.97	0.97	1.03	1.03	0.88	0.90	0.94	1.03	0.88	0.90	0.94
10	0.76	1.01	1.05	0.97	0.97	1.11	1.01	1.01	1.02	1.04	1.05	0.96	1.02	1.02	1.05	0.96	1.10	1.02
11	0.82	0.95	0.96	0.95	0.92	1.07	0.94	0.97	0.93	0.96	0.96	0.92	1.02	0.94	0.95	0.92	1.02	0.94
12	0.72	0.97	1.12	0.83	0.87	0.79	1.00	0.82	0.98	1.10	1.10	0.83	0.87	1.00	1.10	0.83	0.78	1.00
13	0.76	1.06	0.57	0.53	0.76	0.81	1.05	0.90	1.06	0.56	0.55	0.52	0.79	1.05	0.52	0.76	0.79	1.05
13	mse	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio	ratio
1	0.76	1.02	1.36	1.04	1.03	1.16	1.07	1.01	1.03	1.33	1.34	1.01	1.01	1.03	1.34	1.01	1.13	1.07
2	0.75	1.02	1.12	1.02	1.01	1.01	1.04	1.02	1.02	1.10	1.09	1.00	1.00	1.00	1.09	1.00	1.00	1.04
3	0.79	0.97	1.03	0.94	0.93	0.99	0.98	0.96	0.96	1.02	1.02	0.92	0.93	0.99	1.02	0.92	0.99	0.98
4	0.77	0.99	1.18	0.94	0.95	1.11	1.02	0.98	1.00	1.16	1.16	0.93	0.94	1.10	1.16	0.93	1.10	1.01
5	0.70	1.13	1.12	1.11	1.09	1.13	1.14	1.10	1.14	1.12	1.12	1.12	1.10	1.12	1.12	1.12	1.10	1.14
6	0.80	0.97	1.38	0.74	0.75	0.92	0.96	0.93	0.96	1.32	1.30	0.75	0.75	0.93	1.30	0.75	0.93	0.95
7	0.70	1.04	1.08	0.69	0.74	1.05	1.05	1.02	1.04	1.11	1.08	0.71	0.75	1.02	1.08	0.71	1.02	1.05
8	0.72	1.02	1.10	0.92	0.95	1.01	1.03	1.02	1.01	1.09	1.10	0.93	0.95	1.00	1.10	0.93	1.00	1.02
9	0.79	0.95	1.01	0.88	0.89	0.92	0.94	0.96	0.95	1.01	1.01	0.86	0.88	0.91	1.01	0.86	0.91	0.94
10	0.75	1.01	1.04	0.96	0.96	1.11	1.00	1.01	1.02	1.03	1.04	0.95	0.96	1.09	1.04	0.95	1.09	1.00
11	0.80	0.96	0.97	0.94	0.91	1.07	0.93	0.97	0.93	0.96	0.96	0.93	0.91	1.03	0.96	0.93	1.03	0.93
12	0.71	0.96	1.11	0.79	0.84	0.75	0.99	0.79	0.98	1.08	1.08	0.79	0.85	0.75	1.08	0.79	0.85	0.98
13	0.76	1.06	0.56	0.52	0.75	0.79	1.04	0.89	1.05	0.54	0.54	0.50	0.76	1.04	0.54	0.50	0.76	1.04

Notes:

$x_{t+h|t}^{m,fe}$, represents the h -step ahead forecast with the different model specifications $m = \{SF, DF, DFC, AR, \mu\}$. The specifications employ respectively static factors, dynamic factors, cyclical dynamic factors, 2nd order autoregressive model and the first moment of the time series. Moreover, the forecast equation $fe = \{u, fm\}$ is unrestricted such that the time series variables are forecasted directly, or, respectively, admits the factor model structure such that the common component is forecast. The first column reports for each forecast horizon h the forecast performance of the AR-model. The upper part of the column reports the mean squared error (mse) and the lower part the variance (var) of the forecast errors. The other columns report the forecasting performance of the respective models as a ratio to the performance of the benchmark AR-model. The forecasting exercise starts in 2002.9 and produces 32 forecasts for each horizon. All forecasts are evaluated since the aggregate data at the country level are available until 2006.4. The forecasts are generated using a rolling estimation window of 61 periods.

This chapter is based on joint work with Jan P.A.M. Jacobs and Pieter W. Otter from the University of Groningen and is reprinted with their kind permission.

Chapter 7

Information, data dimension and factor structure

Only entropy comes easy.

Anton Chekhov

Abstract

This paper employs concepts from information theory to choosing the dimension of a data set. We propose two relative information measures, one based on eigenvalues, the other connected to Kullback-Leibler numbers. By ordering the series of the data set according to these measures, we are able to obtain a subset of the data set that is most informative to model a variable of interest. The method can be used as a first step in the construction of a dynamic factor model or a leading index, as illustrated with the U.S. macroeconomic data set of Stock and Watson (2005).

7.1 Introduction

With the proliferation of huge data sets a natural question to ask is how much information there is in a data set. Is there an ‘optimal’ size of the data set in relation to some variable(s) of interest, in other words can we confine attention to a subset of the series instead of having to monitor all series in a data set? The question seems especially relevant for factor models, which exploit the idea that movements in a large number of series are driven by a limited number of common ‘factors’. For a recent overview see Bai and Ng (2008).

The paper has benefited from comments received following presentations of previous versions at the Far Eastern Meeting of the Econometric Society, Beijing, China, July 2006, the 13th International Conference on Computing in Economics and Finance, Montréal, Canada, June 2007, the Research Forum: New Developments in Dynamic Factor Modelling, Centre for Central Banking Studies of the Bank of England, London, October 2007, the Far Eastern Meeting of the Econometric Society, Singapore, July 2008, the Conference on Factor Structures for Panel and Multivariate Time Series Data, Maastricht, September 2008, the First Macroeconomic Forecasting Conference, Rome, March 2009 and seminars at De Nederlandsche Bank and the Sveriges Riksbank.

Factor models are a tool to cope with many variables without running into degree of freedom problems often faced in regression based analysis. The extracted factors represent the underlying specific data set, which then should capture correctly the main forces that drive the target variable of interest. The factor diagnostics that improve the precision of the factor estimates relate to the importance of the common component, i.e. the common-to-idiosyncratic ratio and its dispersion across the variables in the data set. Boivin and Ng (2006) conclude that adding variables exhibiting a large idiosyncratic component that is highly correlated only increases the sampling error.

Although convergence of factor estimates requires large cross-sections and large time dimensions, see e.g. Forni and Lippi (2001) and Bai (2003), the data set need however not be very large to obtain reasonably precise factor estimates. Boivin and Ng (2006) and Inklaar, Jacobs, and Romp (2005) find that a selection of some 40 variables are sufficient using Monte Carlo simulations and a comparison to conventional NBER-type business cycle indicators, respectively. Bai and Ng (2002) also conclude that the number of series need not be very large to get precise factor estimates.

Oversampling refers to the situation in which the data are more informative about some factors than the other ones. Including more variables in an oversampled data set could result in more precise factor estimates, which do however not improve the forecasting performance for the target variables that depend on the less dominant factors. The question whether we can confine attention to a subset of the variables is also relevant for the construction of leading indexes, which aims at selecting indicators with predictive power for a target variable, out of a large number of candidates.

Building upon Otter and Jacobs (2006), this paper exploits concepts from information theory, in particular Kullback-Leibler numbers, to analyse information in the data.¹ We propose two relative information measures, one based on eigenvalues, the other based on gaussian distributed data with a clear link to Kullback-Leibler numbers. The second measure is discussed in more detail assuming an approximate factor structure in the data; a test procedure is given whether an additional variable adds information. Ordering the series of the data set according to these measures enables us to identify a subset of the data set that is most informative to modelling a variable of interest. The method can be used as a first step in the construction of a dynamic factor model or a leading index.

We illustrate the concepts with the macroeconomic data set of Stock and Watson (2005), which consists of 132 monthly U.S. variables and runs from

¹Jacobs and Otter (2008) apply similar information concepts to derive a formal test for the number of common factors and the lag order in a dynamic factor model.

1959–2003. We find that relative information is maximized for 40–50 series if we are interested in modelling industrial production and CPI inflation. Allowing for pure leads only or both leads and lags does not affect the size of the subset to a great extent.

The paper is structured as follows. Section 7.2 discusses our two relative information measures, how one of the measures works out assuming an approximate factor structure in the data, and a test procedure. Section 7.3 applies our methods to the U.S. data set. Section 7.4 concludes.

7.2 Information in data

7.2.1 Information measure based on eigenvalues

Consider a N -dimensional stationary normalized random vector $\mathbf{x}_t$ with zero mean and covariance matrix $\mathbf{\Gamma}$. Note that $\mathbf{x}_t$ need not necessarily be normally distributed. In order to establish a relation between the proposed information measure based on the eigenvalues and correlation (factor structure) between the variables consider the Euclidian (Schur) norm of a matrix.

Let $\|\mathbf{A}\|_E = \left(\sum_{i,j} |a_{i,j}|^2 \right)^{1/2} = \text{tr}(\mathbf{A}'\mathbf{A})^{1/2}$ be the Euclidean (Schur) norm of the matrix $\mathbf{A}$. So, $\|\mathbf{\Gamma}\|_E^2 = \sum_{i,j} |\gamma_{i,j}|^2$ measures the magnitude of correlation between $\mathbf{x}_{i,t}$ and $\mathbf{x}_{j,t}$ with $\gamma_{i,i} = 1$. From the spectral decomposition $\mathbf{\Gamma} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}'$ we have

$$\|\mathbf{\Gamma}\|_E^2 = \text{tr}(\mathbf{C}\mathbf{\Lambda}^2\mathbf{C}') = \text{tr}(\mathbf{\Lambda}^2) = \sum_{i=1}^N \lambda_i^2.$$

Let $\boldsymbol{\lambda} = (\lambda_1 \dots \lambda_N)'$, then $\|\mathbf{\Gamma}\|_E^2 = \boldsymbol{\lambda}'\boldsymbol{\lambda}$ which under the restriction $\boldsymbol{\lambda}'\mathbf{1} = \sum \lambda_i = N$ attains its maximum if $\lambda_1 = N$ and $\lambda_j = 0, j = 2, \dots, N$, i.e. $\gamma_{i,j} = 1$ for all i and j and its minimum if $\lambda_j = 1$ for all j , i.e. $\mathbf{\Gamma} = \mathbf{I}_N$ (no correlation at all).

The foregoing shows that the distribution of the eigenvalues is important, which can be measured by the entropy of the eigenvalues, denoted by H , considering the relative eigenvalues as “probabilities”. Because $\text{tr}(\mathbf{\Lambda}) = N$ we have $\bar{\lambda}_j = \lambda_j/N$ with $0 \leq \bar{\lambda}_j \leq 1$ and

$$H_{\bar{\lambda}} = - \sum_j \bar{\lambda}_j \log \bar{\lambda}_j, \quad (7.1)$$

with $H_{\bar{\lambda}}^{\max} = \log(N)$ for $\bar{\lambda}_j = 1/N$ for all j which corresponds to the case of no correlation and $H_{\bar{\lambda}}^{\min} = 0$ in case $\bar{\lambda}_1 = 1$ and $\bar{\lambda}_j = 0$ for $j = 2, \dots, N$ (with the usual convention that $\bar{\lambda}_j \log \bar{\lambda}_j = 0$ for $\bar{\lambda}_j = 0$). The information contained

in the eigenvalues is $\text{Inf}_{\bar{\lambda}} = \log(N) - H_{\bar{\lambda}}$. So the *relative eigenvalues information* can be defined as

$$\text{Inf}_{\bar{\lambda}}^R = 1 - \frac{H_{\bar{\lambda}}}{\log(N)}, \quad (7.2)$$

with $0 \leq \text{Inf}_{\bar{\lambda}}^R \leq 1$. $\text{Inf}_{\bar{\lambda}}^R = 0$ in case $\lambda_j = 1/N$ for $j = 1, \dots, N$ and $\text{Inf}_{\bar{\lambda}}^R = 1$ if $H_{\bar{\lambda}} = 0$. Note that the sensitivity of $\text{Inf}_{\bar{\lambda}}^R$ with respect to the j -th ordered eigenvalue can be measured as

$$\frac{\partial \text{Inf}_{\bar{\lambda}}^R}{\partial \bar{\lambda}_j} = \frac{\log(\bar{\lambda}_j) + 1}{\log(N)},$$

which is positive for $\bar{\lambda}_j > e^{-1} \approx .37$, but decreasing with increasing N .

7.2.2 Kullback-Leibler numbers and information

Let $f_1(\tilde{x}) : \tilde{x} \sim \mathcal{N}_N(\mathbf{0}, \Gamma = \mathbf{C}\mathbf{A}\mathbf{C}')$ be the density function of an N -dimensional data vector x (time index suppressed), then $f_1(x) : x \sim \mathcal{N}_N(\mathbf{0}, \mathbf{A})$ where $x = \mathbf{C}'\tilde{x}$. Let $f_2(\tilde{x}) : \tilde{x} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N)$. Then $f_2(x) : x \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N)$ with $x = \mathbf{C}'\tilde{x}$. The so-called *Kullback-Leibler* numbers are defined as

$$G_1 = E_{f_1} \left(\log \left(\frac{f_1(x)}{f_2(x)} \right) \right) \text{ and } G_2 = E_{f_2} \left(\log \left(\frac{f_2(x)}{f_1(x)} \right) \right), \quad (7.3)$$

and $G = G_1 + G_2$ is the measure of information for discriminating between the two density functions with $G = 0$ in case $f_1(x) = f_2(x)$ and $G = \infty$ in case of perfect discrimination, see Young and Calvert (1974, p245). For a general background see Burnham and Anderson (2002).

For $\text{tr}(\Gamma) = \text{tr}(\mathbf{A}) = N$ we have $G_1 = -\log \det(\mathbf{A})$ and $G_2 = \log \det(\mathbf{A}) + \frac{1}{2} (\text{tr}(\mathbf{A}^{-1}) - N)$. Therefore

$$2G = \text{tr}(\mathbf{A}^{-1}) - N = \text{tr}(\mathbf{A}^{-1}) - \text{tr}(\mathbf{A}) = \sum_{j=1}^N \frac{(1 - \lambda_j^2)}{\lambda_j}, \quad (7.4)$$

from which it can be seen that G is small (not discriminating) if the eigenvalues λ_j are close to 1, but becomes large (discriminating) for “small” eigenvalues.

For the Gaussian case we can use alternative measures of entropy and information. Let x_t again be an N -dimensional vector of observed data at time t , $t = 1, \dots, T$. The data is demeaned and normalized, and normally distributed with mean zero and variance $E(x_t x_t') = \Gamma$, i.e. $x_t \sim \mathcal{N}(\mathbf{0}, \Gamma)$, where $\text{diag}(\Gamma) = (1, 1, \dots, 1)$ and $\text{tr}(\Gamma) = N$. Here we make the additional assumption that all eigenvalues are positive. Note that this assumption can be relaxed when considering the information measure based on eigenvalues. The

entropy as measure of disorder for a stationary, normally distributed vector is given by

$$2H_x = cN + \log \det(\Gamma),$$

where $c \equiv \log(2\pi) + 1 \approx 2.84$, with $2H_{x,max} = cN$ in case $\Gamma = I_N$, see e.g. Goodwin and Payne (1977). The information or negentropy is defined as

$$2\text{Inf}_x \equiv 2(H_{x,max} - H_x) = -\log \det(\Gamma) \geq 0, \quad (7.5)$$

which is zero in case $\Gamma = I_N$. We define the *relative information* as

$$\text{Inf}_N^R = \frac{2H_{max} - 2H_{x(N)}}{2H_{max}} = \frac{2\text{Inf}_N}{2H_{max}} = \frac{2\text{Inf}_N}{cN}. \quad (7.6)$$

If $H_{x(N)}$ is equal to H_{max} then $\text{Inf}_N^R = 0$; if $H_{x(N)} = 0$ then $\text{Inf}_N^R = 1$.

7.2.3 Relative information measure Inf_N^R in the approximate factor model

In this section we consider the relative information measure Inf_N^R in more detail assuming an approximate factor structure in the data. Let x_t be driven by k factors

$$x_t = B_N F_t + \varepsilon_t, \quad x_t \in \mathbb{R}^N, \quad F_t \sim \mathcal{N}_k(\mathbf{0}, I_k), \quad \varepsilon_t \sim \mathcal{N}_N(\mathbf{0}, \Psi_{11}), \quad (7.7)$$

where $B_N \in \mathbb{R}^{N \times k}$ is the matrix of factor loadings. Note that this approximate factor model is sufficiently general to cover the static and the dynamic case. The variance between the first N elements of x_t is equal to $\Gamma(N) = B_N B_N' + \Psi_{11}$.

Adding a variable $x_{N+1,t}$ we have

$$\begin{pmatrix} x_t \\ x_{N+1,t} \end{pmatrix} = \begin{pmatrix} B_N \\ b_{N+1} \end{pmatrix} F_t + \begin{pmatrix} \varepsilon_t \\ \varepsilon_{N+1,t} \end{pmatrix}, \quad (7.8)$$

with covariance $\Gamma(N+1) = \begin{pmatrix} \Gamma(N) & \Gamma_{12} \\ \Gamma_{21} & 1 \end{pmatrix}$, where $\Gamma_{12} = B_N b_{N+1}' + \Psi_{12}$

with $\Psi_{12} = E(\varepsilon_t \varepsilon_{N+1,t})$. Because of the normalisation we have $b_{N+1} b_{N+1}' + \sigma_{N+1}^2 = 1$, where $\sigma_{N+1}^2 = E(\varepsilon_{N+1,t}^2)$. Using the rule of determinants for partitioned matrices we get

$$\det(\Gamma(N+1)) = \det(\Gamma(N))(1 - a_{N+1}),$$

with $a_{N+1} \equiv (\mathbf{b}_{N+1}\mathbf{B}'_N + \boldsymbol{\Psi}'_{12})\boldsymbol{\Gamma}^{-1}(N)(\mathbf{B}_N\mathbf{b}'_{N+1} + \boldsymbol{\Psi}_{12})$ and $0 \leq (1 - a_{N+1}) \leq 1$. After some calculations the following relation between the relative information measures Inf_{N+1}^R and Inf_N^R can be established:

$$\text{Inf}_{N+1}^R = \text{Inf}_N^R - \frac{1}{N+1} \left(\frac{\log(1 - a_{N+1})}{c} + \text{Inf}_N^R \right). \quad (7.9)$$

Therefore a variable $x_{N+1,t}$ adds relative information, i.e. $\text{Inf}_{N+1}^R > \text{Inf}_N^R$, if $-\log(1 - a_{N+1}) > c\text{Inf}_N^R$, provided $E(x_{N+1,t}\mathbf{x}'_t) = (\mathbf{b}_{N+1}\mathbf{B}'_N + \boldsymbol{\Psi}'_{12}) = \boldsymbol{\Gamma}'_{12} \neq 0$. The latter condition can be tested by means of the procedure described in Section 7.2.4 below. Rewriting a_{N+1} as $a_{N+1} = E(x_{N+1,t}\mathbf{x}'_t) \mathbf{C}\boldsymbol{\Lambda}^{-1}\mathbf{C}' E(x_{N+1,t}\mathbf{x}'_t)'$ it is seen that the covariance between $x_{N+1,t}$ and $\mathbf{x}_t$ is amplified by the inverse of the eigenvalues consistent with the criterion based on the KL-numbers, see Equation (7.4).

The second term on the right-hand side of Equation (7.9) serves as a threshold which can be seen as follows. From the equation we have $\text{Inf}_{N+1}^R = \text{Inf}_N^R$ if $a_{N+1} = 1 - \exp(-c\text{Inf}_N^R)$. Whenever Inf_N^R is close to zero, Inf_{N+1}^R increases for relative small values of a_{N+1} whereas if Inf_N^R is close to one, a_{N+1} should be close to one to allow $x_{N+1,t}$ to add information. Note that the threshold in Equation (7.9) is determined by the covariance between $x_{N+1,t}$ and $\mathbf{x}_t$, $\boldsymbol{\Lambda}^{-1}$ which measures the degree of correlation between the components of $\mathbf{x}_t$ consistent with the KL-measure G , the magnitude of Inf_N^R , and the dimension of the data vector.

Equation (7.9) can be simplified by the following procedure. Let $\boldsymbol{\Gamma}(N) = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'$ with $\boldsymbol{\Gamma}(N)$ regular and consider the linear transforms $\tilde{\mathbf{x}}_t = \mathbf{U}'\boldsymbol{\Lambda}^{-1/2}\mathbf{C}'\mathbf{x}_t$ and $\tilde{x}_{N+1,t} = v^{-1}x_{N+1,t}$ with $\mathbf{U}$ orthogonal, i.e. $\mathbf{U}'\mathbf{U} = \mathbf{U}\mathbf{U}' = \mathbf{I}_N$, and $v^2 = 1$ obtained by the singular value decomposition (SVD) $\boldsymbol{\Lambda}^{-1/2}\mathbf{C}'\boldsymbol{\Gamma}_{12} = \mathbf{U}\boldsymbol{\Sigma}v$ with $\boldsymbol{\Sigma} = (\phi_1 \ 0 \ \dots \ 0)'$. From this SVD it can be seen that $\boldsymbol{\Gamma}_{12} = 0$ implies $\boldsymbol{\Sigma} = 0$. Then

$$\begin{pmatrix} \tilde{\mathbf{x}}_t \\ \tilde{x}_{N+1,t} \end{pmatrix} \sim \mathcal{N}_{N+1} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tilde{\boldsymbol{\Gamma}}(N+1) \equiv \begin{pmatrix} \mathbf{I}_N & \boldsymbol{\Sigma} \\ \boldsymbol{\Sigma}' & 1 \end{pmatrix} \right),$$

and hence $\det(\tilde{\boldsymbol{\Gamma}}(N+1)) = \det(\mathbf{I}_N)(1 - \phi_1^2)$. So the information of the transformed vector $(\tilde{\mathbf{x}}'_t \ \tilde{x}_{N+1,t})'$ becomes $\tilde{\text{Inf}}_{N+1} = -\log(1 - \phi_1^2)/2$, where $\phi_1 \in [0, 1]$ is the canonical correlation coefficient. Because $\tilde{\text{Inf}}_N = 0$, Equation (7.9) becomes $\tilde{\text{Inf}}_{N+1}^R = -\log(1 - \phi_1^2)/c(N+1)$.

7.2.4 A test procedure

Replacing $\tilde{\boldsymbol{\Gamma}}(N+1)$ by a consistent estimate $\hat{\tilde{\boldsymbol{\Gamma}}}(N+1)$ and applying the same procedure yields $\hat{\tilde{\text{Inf}}}_{N+1} = -\log(1 - \hat{\phi}_1^2)/2$. Under $H_0 : \phi_1 = 0$, the Bartlett

test statistic

$$-[T - 1/2(N + 2)] \log(1 - \hat{\phi}_1^2) = [T - 1/2(N + 2)] 2\hat{\text{Inf}}_{N+1}$$

follows asymptotically a χ^2 -distribution with N degrees of freedom, see e.g. Muirhead (1982). Testing the hypothesis $\phi_1 = 0$ is basically testing whether the transformed vector $(\tilde{x}'_t \tilde{x}_{N+1,t})'$ has maximum entropy, i.e. no correlation at all. If the null hypothesis is rejected, the estimated relative information of the transformed variables equals

$$\hat{\text{Inf}}_{N+1}^R = -\log(1 - \hat{\phi}_1^2)/c(N + 1).$$

Under the null hypothesis the expected value of $2\hat{\text{Inf}}_{N+1}^R$ is $N/[T - 1/2(N + 2)]$ and hence the expected value of the relative information of the transformed vector $(\tilde{x}'_t \tilde{x}_{N+1,t})'$ is

$$E \left\{ \hat{\text{Inf}}_{N+1}^R \right\} \approx 1/c[T - 1/2(N + 2)] \approx 1/[cT - H_{\max}]$$

where $H_{\max} = c(N + 1)/2$ is the maximum entropy of $(\tilde{x}'_t \tilde{x}_{N+1,t})'$.

7.2.5 MSE-prediction

From the foregoing we have $\tilde{x}_t = Ax_t$ with $A = U'\Lambda^{-1/2}C'$. Given a realization $\tilde{x}_{N+1,t} = v^{-1}x_{N+1,t}$ the conditional mean (predictor) of $\tilde{x}_t$ is $\tilde{x}_t^P = \Sigma \tilde{x}_{N+1,t}$ with conditional variance $\text{var}\{\tilde{x}_t^P\} = I - \Sigma\Sigma' = \text{diag}((1 - \phi_1^2), 1, \dots, 1)$ and information $-\log(1 - \phi_1^2)/2$. Hence if $\phi_1 = 0$ implying $\Sigma = 0$ the vector $\tilde{x}_t^P$ has maximum entropy and no information.

The conditional MSE-predictor of x_t itself is

$$x_t^P = A^{-1}\tilde{x}_t^P = \phi_1 C\Lambda^{1/2}u_1 \tilde{x}_{N+1,t},$$

where u_1 is the first column of the orthonormal matrix U . The conditional variance of x_t^P is

$$\text{var}\{x_t^P\} = A^{-1} \left(A^{-1} \right)' - A^{-1} \Sigma \Sigma' \left(A^{-1} \right)^{-1} = \Gamma(N) - A^{-1} \Sigma \Sigma' \left(A^{-1} \right)^{-1},$$

from which it can be seen that $\Gamma(N)$ exceeds $\text{var}\{x_t^P\}$ by a positive definite matrix if $\phi_1 \neq 0$.

7.3 Application

In the application below, we use the relative information measures introduced above to order a macroeconomic data set. Plots of the relative information measures against the number of variables indicate which subset is most informative for modelling a variable of interest.

7.3.1 The data set

In this section we evaluate the performance of the suggested approach on the Stock and Watson (2005) U.S. macroeconomic data set, which consists of monthly observations on 132 macroeconomic time series from 1959M1 up to and including 2003M12. The series cover 14 categories: real output and income; employment and hours; real retail, manufacturing and trade sales; consumption; housing starts and sales; real inventories; orders; stock prices; exchange rates; interest rates and spreads; money and credit quantity aggregates; price indexes; average hourly earnings; and miscellaneous. The series are transformed by taking logarithms and/or differencing when necessary to assure approximate stationarity. In general, first differences of logarithms (growth rates) are used for real quantity variables, first differences are used for nominal interest rates, and second differences of logarithms for price series (changes in inflation). Moreover, the series are adjusted for outliers by replacing the observations of the transformed variables with absolute median deviations larger than 6 times the interquartile range with the median value of the preceding 5 observations. The specific transformations and the list of series are given in Appendix A of Stock and Watson (2005).

7.3.2 Information in the data set

We order the data set according to the relative information measures with respect to two target variables: the first difference of the log of total industrial production (IP hereafter) and the second difference of the log of the consumer price index (CPI hereafter) using the following procedure: (i) the initial variable of the ordered data set is the target variable; (ii) the variable that maximizes the respective relative information from the remaining data is added to the ordered data set, and so on. Let $\mathcal{D}(n)$ be the ordered data set that consists of n variables. Then variable x_i , $i = 1, \dots, N - n$ of the remaining data set is chosen for which holds that $i = \arg\max_{\lambda(\mathcal{D}(n+1))} I_{\lambda(\mathcal{D}(n+1))}^R$, or $i = \arg\max_{N(\mathcal{D}(n+1))} I_{N(\mathcal{D}(n+1))}^R$ with $\mathcal{D}(n+1) = \{\mathcal{D}(n), x_i\}$. The full data set consists of $N = 132$ time series variables, with $T = 540$ observations covering the sample 1959M1–2003M12. Since the number of observations T is much larger than the number of series N , all eigenvalues of the covariance matrix of x_t differ from zero and both relative information measures are computationally stable. Below we focus on relative information (Inf_N^R) outcomes.

Table 7.1: Ranking of series according to relative information criteria

order	IP		CPI	
	rel. info series #	rel. eigenvalues info series #	rel. info series #	rel. eigenvalues info series #
1	6	6	115	115
2	16	16	124	124
3	20	20	123	123
4	7	7	119	119
5	8	8	125	125
6	13	13	127	127
7	14	14	122	122
8	9	9	117	117
9	12	12	128	128
10	11	19	121	121
11	19	62	39	110
12	62	61	37	56
13	61	50	38	51
14	50	37	34	54
15	64	38	33	59
16	37	34	40	60
17	38	33	41	55
18	34	41	43	58
19	33	40	50	53
20	40	43	61	57
21	41	42	19	52
22	43	64	62	61
23	42	63	64	50
24	63	39	42	19
25	114	114	63	62
26	39	36	114	64
27	102	52	102	37
28	101	57	101	38
29	100	44	100	34
30	99	68	99	33
31	97	2	98	41
32	96	1	97	40
33	98	102	96	43
34	95	101	95	42
35	59	100	59	63
36	54	99	54	39
37	56	98	56	114
38	51	97	51	36
39	60	96	60	44
40	55	95	55	12
41	58	51	58	16
42	53	56	53	20
43	57	54	57	6
44	52	59	52	7
45	49	60	49	8
46	47	55	47	13
47	44	58	44	14
48	36	53	36	9
49	74	49	74	68
50	68	47	68	102

Notes. See the table in the appendix for the description of the variables.

Figure 7.1: Relative information (dashed line) and relative eigenvalues information (solid line) of ordered data set
Target variable: IP

Target variable: CPI

Table 7.1 presents the orders of the first 50 variables according to the two relative information criteria for both target variables. The table allows the following observations. The first ten series that are included in the subset for IP belong to the group of Industrial Production; the first ten series for CPI are price indices. Second, the overlap between the ordered data sets according to the different relative information criteria in the first 50 series is very high: 96% for IP and 90% for CPI. Third, price indices are generally speaking not informative for IP (the exception is series # 114: NAPM commodity price index), while production series do not appear in the first fifty variables of the ordered data subset for CPI (with one exception series # 19: NAPM production). Finally, variables enter the ordered data sets in clusters. For IP, the relative information measure first selects a group of industrial production variables, followed by employment series, interest rates and spreads, and housing starts and sales. With CPI as target variable, the relative information measure starts with picking price indices, followed by employment, orders, interest rates and spreads, housing starts and sales, and employment. Adoption of the relative eigenvalue information results in slightly different sequences for both target variables.

Figure 7.1 shows the evolution in relative information if we order the data set according to the target variables IP (top panel) and CPI (bottom panel). The figure reveals that sometimes relative information decreases with the addition of a single series, but increases if a batch of variables is added. This pattern is most pronounced if we order the data set according to CPI using the relative eigenvalue information measure. For both target variables relative information attains a global maximum if we take between 40 and 50 series in line with the findings of Boivin and Ng (2006) and Inklaar et al. (2005).

Figure 7.2 shows outcomes (p-values) of the test described in Section 7.2.4 whether an additional variable adds information. The null hypothesis is that an additional variable is not correlated with the variables already included in the set. Hence, low p-values indicate that an additional variable adds information. We note that the outcomes of the test are not sensitive to the initial condition, i.e. the choice of the target variable. The figure suggest that some 120 series are informative. This finding does not contradict our conclusion that relative information, measured by the ratio of information, Inf_N , and maximum entropy cN , is maximized for 40–50 series. More than this number of series add information to the ordered data set, i.e. $\text{Inf}_{N+1} > \text{Inf}_N$ for $40 < N < 120$, but apparently the additional information does not exceed the increase in entropy in these series, $\text{Inf}_{N+1} - \text{Inf}_N < c(N+1) - cN = c$, and therefore $\text{Inf}_{n+1}^R < \text{Inf}_n^R$.

Figure 7.2: Does an additional variable add information?

7.3.3 Allowing for pure leads and leads and lags

Our methods can be quite useful to reduce the size of a data set as a first step in the construction of dynamic factor models or leading indexes. To that purpose we calculate the relative information within the data set allowing for both leads and lags and pure leads only. If leads and lags are allowed, individual variables can be selected with a ‘lead’ j between $-k$ and k periods, and a pure lead of $j = 0, \dots, k$ periods. In either case, variable $x_{i,t-j}$, $i = 0, \dots, N - n$ and $j = 0, \dots, k$ of the remaining data set is chosen with a ‘lead’ of j periods for which holds that $\{i, j\} = \arg \max \text{Inf}_{\lambda(D(n+1,j))}^R$ or $\{i, j\} = \arg \max \text{Inf}_{N(D(n+1,j))}^R$ with $D(n+1, j) = \{D(n), x_{i,t-j}\}$. Here we take a three-year horizon and set $k = 36$. This implies that the data is shortened to around 500 observations for the pure lead case, and to around 460 observations if we allow leads and lags.

Table 7.2 shows the order of the first 50 variables on the basis of relative information I_N^R when we allow for pure leads only, and both leads and lags. The overlap between the two cases for the first 50 series is close to 100% for both target variables. If we allow only pure leads, the first 30 series enter the

ordered data set for IP without a lead. Note that this sequence may differ from the static case, because the reduction in the number of observations affects the eigenvalues of the covariance matrix of the data matrix x_t . Housing starts and sales variables and hours enter the ordered data set with a lead of more than two years. If series may enter with leads and lags, the majority of the first fifty variables is selected with no lead/lag or a small lag. Strikingly, most housing starts and sales variables now get a lag of five months, whereas hours enter the ordered data set with a lag of over one year.

If leads and lags are allowed with CPI as target variable, the first ten variables — all price indices — enter the ordered data set without a lead or a lag. All other variables enter the ordered data set with considerable lags, with the exception of two housing variables which get a twenty months lead. The maximum lag is two and a half years for interest rates and spreads. Allowing pure leads only yields a maximum lead of 31 months for hours worked, and a lead of 14 months for housing starts and sales variables. The majority of the series however enters the ordered data set with no lead or a small lead.

Figure 7.3 compares relative information in the Stock and Watson data set with respect to IP (top panel) and CPI (bottom panel) in the static case, pure leads only, and both leads and lags. The relative information patterns for the three cases are similar. However relative information is higher if the series are allowed to enter the ordered data set with pure leads only or both leads and lags than in the static case. Hence, there is scope in the data for constructing a dynamic factor model and a leading index for GDP. The maximum is attained at around 40 series for both target variables.

7.4 Conclusion

This paper fruitfully applied concepts from information theory in the analysis of large data sets. We defined two relative information measures, one based on eigenvalues, the other linked to Kullback-Leibler numbers. The application of the measures enabled us to order a data set and to identify a subset of the data that is most informative to modelling a variable of interest.

We illustrated our methods with the Stock and Watson (2005) U.S. macroeconomic data set consisting of 132 times series variables with 540 observations. With around 40–50 series relative information is maximized for industrial production and inflation. Approximately the same number of series enter the ordered data set if leads and lags or pure leads are allowed. We conclude that our method can indeed produce a considerable reduction in the dimension of a data set.

Table 7.2: Ranking of series according to relative information criterion: pure leads, and leads and lags

order	IP				CPI			
	pure leads		leads and lags		pure leads		leads and lags	
	series #	lead	series #	lead(+)/lag(-)	series #	lead	series #	lead(+)/lag(-)
	IP				CPI			
	Relative information				Relative information			
	pure leads		leads + lags		pure leads		leads + lags	
1	6		6		115		115	
2	16	0	16	0	124	0	124	0
3	20	0	20	0	122	0	125	0
4	7	0	7	0	123	0	127	0
5	8	0	8	0	119	0	119	0
6	13	0	13	0	125	0	123	0
7	14	0	14	0	127	0	122	0
8	9	0	9	0	117	0	117	0
9	12	0	12	0	128	0	128	0
10	11	0	11	0	121	0	121	0
11	38	0	38	0	56	14	44	-20
12	37	0	37	0	59	14	40	-20
13	34	0	34	0	54	14	33	-20
14	33	0	33	0	51	14	34	-20
15	40	0	50	-1	60	14	37	-20
16	41	0	61	-1	55	14	38	-20
17	43	0	19	-1	57	14	41	-20
18	42	0	62	-1	52	14	43	-20
19	50	0	63	-2	58	14	50	-21
20	61	0	64	-2	53	14	61	-21
21	19	0	40	0	47	31	19	-21
22	62	0	41	0	49	31	62	-21
23	64	0	43	0	64	5	63	-22
24	63	0	42	0	50	6	64	-22
25	114	0	114	-3	61	6	42	-20
26	102	0	49	-26	19	6	114	-22
27	101	0	47	-26	62	6	102	-30
28	100	0	53	9	63	6	101	-30
29	99	0	58	9	102	0	100	-30
30	97	0	100	0	101	0	99	-30
31	96	0	101	0	100	0	49	-25
32	59	26	102	0	99	0	47	-25
33	56	26	99	0	97	0	52	20
34	54	26	59	-5	96	0	57	20
35	51	26	56	-5	98	0	59	-20
36	60	26	54	-5	38	6	56	-20
37	55	26	51	-5	37	6	54	-20
38	58	26	60	-5	34	6	51	-20
39	53	26	55	-5	33	6	60	-20
40	57	26	57	-5	40	6	55	-20
41	52	26	52	-5	41	6	58	-20
42	98	0	97	0	43	6	53	-20
43	49	27	96	0	42	6	97	-30
44	47	27	98	0	114	3	96	-30
45	95	0	95	23	95	0	98	-30
46	39	0	44	-34	39	6	95	-30
47	44	0	39	0	36	6	36	-20
48	36	0	74	0	44	19	39	-20
49	68	3	36	0	68	8	74	-15
50	74	28	45	0	74	19	68	-20

Notes. See the table in the appendix for the description of the variables.

Our relative information measures are based on the eigenvalues of the covariance matrix of the data, which is however only defined if the number of observations T exceeds the number of series N . Future research will deal with the mirror situation of $N > T$.

Figure 7.3: Comparison of relative information
Target variable: IP

Target variable: CPI

7.A Appendix: The U.S. macroeconomic data set

Table A.7.3 lists the 132 series of the Stock and Watson (2005) U.S. data set, with number, mnemonic, and description of the variable. For details like the transformation applied to the series and sources see Stock and Watson (2005) Appendix A. As is required for factor estimation, the variables are standardized by subtracting their mean and then dividing by their standard deviation. This standardization is necessary to avoid overweighting of large variance series in the factor estimation.

Table A.7.3: Description of the Stock and Watson data set

#	Short name	Mnemonic	Description
1	PI	A0M052	Personal income (AR, bil. chain 2000 \$)
2	PI less transfers	A0M051	Personal income less transfer payments (AR, bil. chain 2000 \$)
3	Consumption	A0M224_R	Real Consumption (AC) A0m224/gmdc
4	M&T sales	A0M057	Manufacturing and trade sales (mil. Chain 1996 \$)
5	Retail sales	A0M059	Sales of retail stores (mil. Chain 2000 \$)
6	IP: total	IPS10	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX
7	IP: products	IPS11	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL
8	IP: final prod	IPS299	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS
9	IP: cons gds	IPS12	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS
10	IP: cons dble	IPS13	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS
11	iIP:cons nondble	IPS18	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS
12	IP:bus eqpt	IPS25	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT
13	IP: matls	IPS32	INDUSTRIAL PRODUCTION INDEX - MATERIALS
14	IP: dble mats	IPS34	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS
15	IP:nondble mats	IPS38	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS
16	IP: mfg	IPS43	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)
17	IP: res util	IPS307	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES
18	IP: fuels	IPS306	INDUSTRIAL PRODUCTION INDEX - FUELS
19	NAPM prodn	PMP	NAPM PRODUCTION INDEX (PERCENT)
20	Cap util	A0M082	Capacity Utilization (Mfg)
21	Help wanted indx	LHEL	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)
22	Help wanted/emp	LHELX	EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF
23	Emp CPS total	LHEM	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)
24	Emp CPS nonag	LHNAG	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)
25	U: all	LHUR	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%;SA)
26	U: mean duration	LHU680	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)
27	U < 5 wks	LHU5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
28	U 5-14 wks	LHU14	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)
29	U 15+ wks	LHU15	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)
30	U 15-26 wks	LHU26	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)
31	U 27+ wks	LHU27	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS.,SA)
32	UI claims	A0M005	Average weekly initial claims, unemploy. insurance (thous.)
33	Emp: total	CES002	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE
34	Emp: gds prod	CES003	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING
35	Emp: mining	CES006	EMPLOYEES ON NONFARM PAYROLLS - MINING
36	Emp: const	CES011	EMPLOYEES ON NONFARM PAYROLLS - CONSTRUCTION
37	Emp: mfg	CES015	EMPLOYEES ON NONFARM PAYROLLS - MANUFACTURING
38	Emp: dble gds	CES017	EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS
39	Emp: nondbles	CES033	EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS
40	Emp: services	CES046	EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING
41	Emp: TTU	CES048	EMPLOYEES ON NONFARM PAYROLLS - TRADE, TRANSPORTATION, AND UTILITIES
42	Emp: wholesale	CES049	EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE
43	Emp: retail	CES053	EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE
44	Emp: FIRE	CES088	EMPLOYEES ON NONFARM PAYROLLS - FINANCIAL ACTIVITIES
45	Emp: Govt	CES140	EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT
46	Emp-hrs nonag	A0M048	Employee hours in nonag. establishments (AR, bil. hours)
47	Avg hrs	CES151	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFARM PAYROLLS - GOODS-PRODUCING
48	Overtime: mfg	CES155	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFARM PAYROLLS - MFG OVERTIME HOURS
49	Avg hrs: mfg	AOM001	Average weekly hours, mfg. (hours)
50	NAPM empl	PMEMP	NAPM EMPLOYMENT INDEX (PERCENT)
51	HStarts: Total	HSFR	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SAAR)
52	HStarts: NE	HSNE	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.
53	HStarts: MW	HSMW	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.
54	HStarts: South	HSSOU	HOUSING STARTS:SOUTH (THOUS.U.)S.A.
55	HStarts: West	HSWST	HOUSING STARTS:WEST (THOUS.U.)S.A.
56	BP: total	HSBR	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)
57	BP: NE	HSBNE	HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A
58	BP: MW	HSBMW	HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.
59	BP: South	HSBSOU	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.
60	BP: West	HSBWST	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.
61	PMI	PMI	PURCHASING MANAGERS' INDEX (SA)
62	NAPM new ordrs	PMNO	NAPM NEW ORDERS INDEX (PERCENT)
63	NAPM vendor del	PMDEL	NAPM VENDOR DELIVERIES INDEX (PERCENT)
64	NAPM Invent	PMNV	NAPM INVENTORIES INDEX (PERCENT)

#	Short name	Mnemonic	Description
65	Orders: cons gds	A0M008	Mfrs' new orders, consumer goods and materials (bil. chain 1982 \$)
66	Orders: dble gds	A0M007	Mfrs' new orders, durable goods industries (bil. chain 2000 \$)
67	Orders: cap gds	A0M027	Mfrs' new orders, nondefense capital goods (mil. chain 1982 \$)
68	Unf orders: dble	A1M092	Mfrs' unfilled orders, durable goods indus. (bil. chain 2000 \$)
69	M&T invent	A0M070	Manufacturing and trade inventories (bil. chain 2000 \$)
70	M&T invent/sales	A0M077	Ratio, mfg. and trade inventories to sales (based on chain 2000 \$)
71	M1	FM1	MONEY STOCK: M1(CURR,TRAV,CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL\$,SA)
72	M2	FM2	MONEY STOCK:M2(M1+O'NITE RPS,EURO\$,G/P&B/D MMMFS&SAV&SM TIME DEP)(BIL\$,SA)
73	M3	FM3	MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL\$,SA)
74	M2 (real)	FM2DQ	MONEY SUPPLY - M2 IN 1996 DOLLARS (BCI)
75	MB	FMFBA	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL\$,SA)
76	Reserves tot	FMRRR	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL\$,SA)
77	Reserves nonbor	FMNRBA	DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL\$,SA)
78	C&I loans	FCLNQ	COMMERCIAL & INDUSTRIAL LOANS OUTSTANDING IN 1996 DOLLARS (BCI)
79	C&I loans	FCLBMC	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL\$,SAAR)
80	Cons credit	CCINRV	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)
81	Inst cred /PI	A0M095	Ratio, consumer installment credit to personal income (pct.)
82	S&P 500	FSPCOM	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
83	S&P: indust	FSPIN	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)
84	S&P div yield	FSDXP	S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)
85	S&P PE ratio	FSPXE	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%NSA)
86	FedFunds	FYFF	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
87	Commppaper	CP90	Commercial Paper Rate (AC)
88	3 mo T-bill	FYGM3	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
89	6 mo T-bill	FYGM6	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)
90	1 yr T-bond	FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
91	5 yr T-bond	FYGT5	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)
92	10 yr T-bond	FYGT10	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)
93	Aaabond	FYAAAC	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)
94	Baa bond	FYBAAC	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
95	CP-FF spread	SCP90	cp90-fyff
96	3 mo-FF spread	SFYGM3	fygm3-fyff
97	6 mo-FF spread	SFYGM6	fygm6-fyff
98	1 yr-FF spread	SFYGT1	fygt1-fyff
99	5 yr-FFspread	SFYGT5	fygt5-fyff
100	10yr-FF spread	SFYGT10	fygt10-fyff
101	Aaa-FF spread	SFYAAAC	fyaaac-fyff
102	Baa-FF spread	SFYBAAC	fybaac-fyff
103	Ex rate: avg	EXRUS	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)
104	Ex rate: Switz	EXRSW	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.\$)
105	Ex rate: Japan	EXRJAN	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.\$)
106	Ex rate: UK	EXRUK	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
107	EX rate: Canada	EXRCAN	FOREIGN EXCHANGE RATE: CANADA (CANADIAN \$ PER U.S.\$)
108	PPI: fin gds	PWFSA	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)
109	PPI: cons gds	PWFCSA	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)
110	PPI: int mat'sls	PWIMSA	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)
111	PPI: crude mat'sls	PWCMSA	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)
112	Commod: spot price	PSCCOM	SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)
113	Sens mat'sls price	PSM99Q	INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)
114	NAPM com price	PMCP	NAPM COMMODITY PRICES INDEX (PERCENT)
115	CPI-U: all	PUNEW	CPI-U: ALL ITEMS (82-84=100,SA)
116	CPI-U: apparel	PU83	CPI-U: APPAREL & UPKEEP (82-84=100,SA)
117	CPI-U: transp	PU84	CPI-U: TRANSPORTATION (82-84=100,SA)
118	CPI-U: medical	PU85	CPI-U: MEDICAL CARE (82-84=100,SA)
119	CPI-U: comm.	PUC	CPI-U: COMMODITIES (82-84=100,SA)
120	CPI-U: dbles	PUCD	CPI-U: DURABLES (82-84=100,SA)
121	CPI-U: services	PUS	CPI-U: SERVICES (82-84=100,SA)
122	CPI-U: ex food	PUXF	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)
123	CPI-U: ex shelter	PUXHS	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)
124	CPI-U: ex med	PUXM	CPI-U: ALL ITEMS LESS MEDICAL CARE (82-84=100,SA)
125	PCE defl	GMDC	PCE,IMPL PR DEFL:PCE (1987=100)
126	PCE defl: dlbes	GMDCD	PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)
127	PCE defl: nondble	GMDCN	PCE,IMPL PR DEFL:PCE; NONDURABLES (1996=100)
128	PCE defl: services	GMDCS	PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)
129	AHE: goods	CES275	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFARM PAYROLLS - GOODS PRODUCING
130	AHE: const	CES277	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFARM PAYROLLS - CONSTRUCTION
131	AHE: mfg	CES278	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFARM PAYROLLS - MANUFACTURING
132	Consumer expect	HHSNTN	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)

Chapter 8

Summary and Conclusions

The large collection of available leading indicator variables can be employed to forecast, particularly Dutch, key macroeconomic variables such as inflation, GDP, staffing employment and the business cycle on a short- to medium term horizon. The forecasting models are based on an explicitly selected small collection of targeted indicators or implicitly weighing of all the available indicators as characterized by diffusion indexes or common factors. In this section, we firstly summarize the resulting forecasting models for the different macroeconomic variables. Secondly, we address to what extent the observed aggregate macroeconomic variable can be forecasted by its underlying disaggregates. This thesis employs common factors that are estimated as weighted averages of the available large collection of underlying disaggregate indicator variables. Thirdly, we analyze the importance of the size and structure of the data set in relation to a target variable of interest. Finally, we provide some overall conclusions and directions for future research.

8.1 Inflation

The inflation forecasting models are linear vector autoregressive or error correction models, possibly including exogenous variables. In order to select the appropriate models, the price indices that show a clear changing seasonal patterns are modelled in both first and twelve month differences including an error correction term representing long run equilibrium relationships between inflation and other variables. The price indices unprocessed food and energy don't show clear structural breaks in seasonal pattern and are modelled in first differences without error correction. The second step involves explicit model selection using a small set of exogenous and endogenous variables. We select the best models according to nine different statistical selection criteria, using both in-sample goodness-of-fit, parsimony and out-of-sample forecasting accuracy. In the third step, the optimal models are chosen, based both on the statistical criteria and economic evaluation. Especially if they are included in the long run relationship, exogenous variables provide a solid anchor for

future inflation. Due to its imposed prominence as a built-in channel of self-fulfilling inflation expectations, wage developments form a model anchor as they are well predictable due to sluggishness in the wage formation process.

According to a recursive forecast error evaluation exercise, all models outperform the naive forecast and optimal autoregressive models on most forecast horizons. The inflation forecast errors are to a certain extent explained by unexpected shocks like, to name a few, the animal diseases Mad-Cow-disease and Foot-and-Mouth disease and the erratic developments of oil prices. While forecasting models cannot foresee unexpected shocks, foreseeable institutional changes or base effects, i.e. exceptional shocks during the current year, are incorporated. For example, the introduction of the euro notes and coins in January 2002 and the beginning of a fierce competition between the supermarkets in the Netherlands in November 2002 exhibited a substantial impact on the price indices. Quantifying such an idiosyncratic impact makes forecasting inflation an art as well as a science.

8.2 Business cycle

Policy institutions operate a business cycle indicator as an instrument to measure and forecast the business cycle and its turning points. This thesis constructs business cycle indicators along the lines of the NBER that consists of a reference index, which represents the current stage of the business cycle, and the indicator, which represents the developments of the cycle in the near future. We apply an approximate band-pass filter to isolate from a time series all the intrinsic cyclical motion consisting of waves with lengths longer than a lower bound of, say 18 months, and shorter than an upper bound of, say 10 years. We focus and compare the cycles for nine OECD-countries: the Netherlands, Belgium, Germany, France, Italy, Spain, the United Kingdom, the United States and Japan.

8.2.1 Business cycle measurement

The composite reference index is based on several coincident indicators that measure real economic activity. For an indicator to be useful in practice, a timely update and therefore a limited publication delay for new observations of the source data is a crucial condition to be met. The business cycle indicator for the Dutch economy is based on key variables analyzed by business cycle dating committees that are available with a limited publication delay being industrial production, staffing employment and consumption expenditures.

In an international comparison context, we study the cycles in the manufacturing industry of the nine OECD-countries that are identified at pre-specified business cycle frequencies. We derive the cyclical turning points, low- and high-growth periods and summary statistics describing features like amplitude, steepness and duration of the cycle for each country. Almost all countries tend to display kurtosis indicating that large cyclical movements in the manufacturing industry are relatively common. The manufacturing industries in the Netherlands and, to a lesser extent, Belgium move quickly from low- to high-growth phases and vice versa. The U.S. reveals a pattern of smooth rounded peaks and pointed deep troughs. The hypothesis of no duration dependence is rejected for nearly all deviation cycles in the manufacturing industries for all countries in both high- and low-growth phases, so an ongoing phase becomes more likely to end. The international linkage between the manufacturing industries is explored by calculating the fraction of time two countries are both in the same phase. This statistic shows that manufacturing industry in the United Kingdom is more synchronised with the one in the United States than with the one in the euro area. Moreover, the average synchronisation between the United Kingdom and the euro area countries is lower than the average synchronisation of the euro area countries with one another.

8.2.2 Composite leading business cycle indicator

Since economic fluctuations originate from different sources, we combine a number of relevant leading indicators into a single composite index. Hereby, we explicitly select a small collection of single leading indicators based on economic and statistical criteria. Economic plausibility requires that the leading indicators should either cause business cycle fluctuations or quickly react to positive or negative shocks. Consistent timing requires the leading indicator to systematically anticipate peaks and troughs with a rather constant lead time. Moreover, conformity to the general cycle requires the leading indicator to possess good overall forecasting performance, not only at peaks and troughs. The statistical criteria require the leading indicators to be promptly available without publication delay and be subject to only minor data revisions. For each of the nine OECD countries, we constructed a composite leading indicator that replicates and predicts the deviation cycle in the manufacturing industry. The lead of the composite leading indicator is determined such that the dates of the turning points match most closely with the corresponding turning point dates of the deviation cycle indicator of the manufacturing industry. The underlying leading indicators are selected from a set of

candidate variables for each country. The variables that have been selected for five or more countries are the short term and long term interest rates, the storage of final products, the hourly wages, the domestic sales prices, the IFO-indicator for Germany and the consumption price index.

The Dutch composite leading indicator consists of three financial series, four business and consumer survey results and four real activity variables, of which two supply and two demand related. The two composite indices are presented as time series that view the current stance of the business cycle in a historical perspective. Alternatively, the underlying single indicators are also presented as a cloud of points showing a multivariate assessment of the current phase of the cycle. A concentrated cloud of points indicates that all individual indicators convey a similar message. Moreover, the development of the cloud of points over time clearly signal a materialising turning point.

8.3 GDP

Marcellino (2006b, Chapter 4) reviews the literature and practices on the choice and evaluation of leading indicator models, possibly combined into composite indexes, and the development of more sophisticated methods to relate the indicators to a forecast for a target variable. The objective is to compare the forecasts of quarterly growth rates for GDP of the Netherlands based on the 14 single monthly indicator variables with the ones based on the composite indices. We employ a bivariate VAR containing GDP growth and the quarterly aggregate of the monthly composite leading index with the number of lags determined by the Schwartz Information Criterion. The forecasting performance is confronted with the one based on the unweighted average of the 14 bivariate VARs containing each of the underlying single leading indicator variables. The results show that the composite index performs better than its underlying indicator variables separately. Moreover, the gain in performance results mostly from the trend-cycle decomposition.

Rünstler et al. (2009) perform a more elaborate GDP forecasting exercise using identical procedures across small and large collections of available leading indicator variables for different European countries and the euro area aggregate. The goal of the exercise is in line with this thesis as it compares the forecast performance based on a small collection of explicitly selected leading indicator variables typically seen on the radar screens of business cycle watchers and a large collection of available leading indicators represented by a few common factors. The Dutch results of Rünstler et al.'s (2009) analysis are based on the 14 single indicator variables that comprise the composite indices. The results that are robust across the different country data sets show

that factor-based estimates are in general more accurate than forecasts based on univariate equations. The three factor models rank ahead of the alternative models indicating that employing factors extracted from many monthly time series is preferable to the average of many small-scale equations each constructed with individual leading indicator variables.

The Dutch GDP factor forecasting exercise in this thesis is based on a similar data set sampled at a quarterly frequency. The analysis is performed following the best practices as documented in Eickmeier and Ziegler's (2008) meta-analysis of the empirical factor-based forecast literature and focusses on comparing different specifications of the factors and of the forecast equation relating the factors to the target variable. The forecast simulation design involves in-sample model selection, factor estimation, parameter estimation and, finally, generating factor forecasts and factor augmented autoregressive forecasts. Of all the possible specifications considered, the factor augmented autoregressive specification imposes the factor structure to the least extent and therefore allows the forecast equation most flexibility to adapt to the data. The results indicate that imposing the factor structure, despite poorer factor model diagnostics, improves upon the forecasting performance. The factor forecasts outperform the factor augmented autoregressive specifications consistently for all specifications and all horizons, significantly so at the first horizon. According to the statistic of forecast accuracy over time, the best performing specification is the restricted cyclical dynamic factor forecast. This specification rests upon the most comprehensive factor design, which encompasses both dynamics and cyclicalities, and moreover, imposes the factor structure on the forecast equation. Taking the large collection of indicator variables as the reference for the restricted factor forecast specifications, then the results show that diagnostics matter as the marginal improvement of the factor diagnostics resulting from excluding a particular group of indicator variables progresses upon the forecasting performance.

8.4 Staffing employment

This thesis analyses the developments of the staffing labour cycle in the Netherlands at a disaggregate level using the data set from Randstad. We create a balanced data set that describes the number of hours of staffing employment for 15 different regions and 58 different sectors. We apply factor models to extract low-dimensional common information from the data set and show that the extracted signal resembles the year-on-year growth rate of aggregate staffing employment. The common signal, which excludes the effects of sector or region specific shocks, is also extracted at the disaggregate level. We analyse

the correlation structure and classify the disaggregate cycles as leading and lagging according to eight empirical measures. Almost all regions lead or lag the staffing labour cycle by less than half a year. The regions, whose modest leading characteristics are robust across four different empirical measures, are Gelderland and Overijssel.

Almost all sectors show a lead that lies between -2 years and +1.5 years. The differences across the sectors are more pronounced than across the regions. Three leading sectors, whose leading characteristics are robust across four different empirical measures, are: supporting and auxiliary transport activities; sale, maintenance and repair of motor vehicles and motorcycles; and retail trade. The turnover in the latter two sectors are known to be stylized business cycle leading indicators.

We then explored how the identified leading indicators at the disaggregate level can be exploited to forecast the country aggregate of the staffing employment. We compare different model specifications that employ static factors, dynamic factors, cyclical dynamic factors, a second order autoregressive model and the first moment of the time series. The performance is measured by the forecast bias and the mean squared forecast error. Detailed information does not necessarily improve upon the forecasting performance as the static factor model does not outperform the benchmark autoregressive model. Due to substantial temporal and seasonal variation in the staffing labour market, the dynamic factor model manages to outperform the univariate benchmark forecasting models.

8.5 Forecasting the aggregate using disaggregates

This thesis aims to determine the extent to which disaggregate variables can fruitfully be incorporated in the forecasting models for the aggregate. Disaggregation can be regarded in three dimensions, that is 1) across components: subsets that constitute a variable like sub-indices of a price measure or different sectors composing GDP; 2) temporal: sampling of a specific aggregate at higher frequencies like monthly observations of quarterly GDP and 3) across space: different regions of an economic area such as the member countries that constitute the euro area. Econometric theory shows that aggregating disaggregate forecasts improves upon directly forecasting the aggregate if the data generating process is known. In practice however, the unknown data generating process and the unknown model parameters, which have to be estimated, turn the preeminence of the disaggregate approach into an empirical question. Instead of forecasting the disaggregate variables separately and aggregating

those forecasts, an alternative approach is to include the disaggregate variables directly in the forecasting model for the aggregate.

8.5.1 Inflation

In the European context there is the aggregation of the forecasts of individual countries to a euro area level, see for instance Espasa et al. (2002), Marcelino et al. (2003) and Benalal et al. (2004). In this thesis, we will only address the aggregation of HICP component forecasts as we focus only on monthly Dutch and euro area HICP inflation. The HICP prediction is constructed by aggregating forecasts for the five HICP sub-indices unprocessed food, processed food, non-energy industrial goods, energy and services, whereas total HICP is also modelled directly. Aggregating forecasts of component models is potentially beneficial as forecast errors might cancel between components. Moreover, the disaggregate components can be better modelled by choosing a more suitable model for each component separately and by possibly incorporating additional explanatory variables. However, Hubrich (2005) and Benalal et al. (2004) find empirical evidence for euro area data and across various specifications that directly forecasting the aggregate HICP performs better than aggregating the forecasted components, especially for a forecast horizon up to 12 months ahead. Our evaluation based on perfect knowledge of future variables for exogenous variables shows a clear preference for aggregating the forecasts of the components for the euro area. For the Netherlands for short forecast horizons aggregating is better, but for longer horizons the direct approach is to be preferred.

8.5.2 GDP

The GDP forecast comparison analysis is based on a large collection of available indicators consisting of 124 quarterly time series. The data set aims to provide a balanced representation of the Dutch economy and of the forces it is exposed to. The data cover the Dutch national accounts on the expenditure components of GDP and describe the behavior of the macro actors in the economy: households, firms, monetary financial sector, government and foreign sector. The data set is supplemented with a more detailed extension of macro-wide variables and leading indicators. The data set can be divided into six different categories, which are labeled GDP, industrial production, prices, financial, external and surveys. The data set covers the production, expenditure and income, whose respective aggregates constitute GDP according to the macroeconomic accounting identity.

The results favor factor based forecasts employing the weighted average of the large collection of available disaggregates. Size matters as the entire collection of 124 series outperforms for all horizons and almost all specifications the data configurations consisting of each of the six individual groups sized of around 20 variables. The short term forecasting potential is clearly determined by the cyclical component. As the cyclical fluctuation of the underlying components of GDP don't always move synchronously and can even move anti-cyclically to the equivalent of GDP, directly forecasting the aggregate of GDP using suitably weighted underlying disaggregate components proves optimal.

8.5.3 Staffing employment

As the analysis of the staffing employment in the Netherlands identifies the regions and sectors that show leading properties, the second question is then how the disaggregate information can be exploited to forecast the country aggregate of staffing employment. The available data set is closed in the sense that the country aggregate is by definition the sum of the 870 possible combinations of 15 regions and 58 sectors, of which only 536 combinations contain observations. The closed data allows us to analyse the forecasting performance from the aggregating forecasts versus forecasting aggregates viewpoint. For the recursive estimation window method, forecasting the aggregate based on disaggregate information outperforms aggregating the forecasted disaggregates (based on disaggregate information). The rolling window method shows no distinctive differences. One explanation for the result is that the time-varying and for outlier-corrected aggregation weights might be biased. Note that forecasting the aggregate employing the factor models incorporates the disaggregate information as contained by the factors. The best performing model directly links the disaggregate information to the aggregate target variable. However, detailed information does not necessarily improve the forecasting performance as only the dynamic factor model is able to exploit the substantial temporal and seasonal variation in the staffing labour market.

8.6 Data dimension and factor structure

The underlying notion that economic motions are captured by a few aggregate driving forces implies that the information contained in each available economic key variable at an aggregate level is less informative about macroeconomic behavior than the information contained in the available variables at a disaggregate level. A diffusion index captures the common motion that

is most widespread across the large collection of available indicators, thereby representing the unobserved common shock, that is the common factor. Factor models are a tool to cope with many variables without running into degree of freedom problems often faced in regression based analysis.

The extracted factors represent the underlying specific data set, which then should capture correctly the main forces that drive the target variable of interest. Oversampling refers to the situation in which the data are more informative about some factors than about other ones. Including more variables in an oversampled data set could result in more precise factor estimates, which do however not improve the forecasting performance for the target variables that depend on the less dominant factors. In absence of oversampling, the factor diagnostics that improve the precision of the estimates relate to the importance of the common component, i.e. the common-to-idiosyncratic ratio and its dispersion across the variables in the data set. So, a large data set should not be extended if the additional candidate variable exhibits a low, and highly correlated, common-to-idiosyncratic variance ratio.

In order to empirically determine the importance of the size and the structure of the data set, we generate GDP factor forecasts for different configurations of the data set. Taking the large sized complete data set configuration as the benchmark for the restricted factor forecast specifications, then the results show that diagnostics matter as the marginal improvement of the factor diagnostics resulting from excluding a particular group progresses upon the forecasting performance. For example, excluding the category consisting of surveys improves upon the forecasting performance. Apparently, the survey variables expose the data compilation to oversampling, thereby hampering forecast improvement. The surveys convey an idiosyncratically confined signal, which is complementary to the common variation in the complete data set as captured by the common factors.

Finally, this thesis aims to formalize the relationship between the choice and amount of variables to include in data set in relation to a target variable. We exploit concepts from information theory, in particular Kullback-Leibler criteria, to quantify information, which depends on the size and the correlation structure of the data set. We propose two relative information measures, one based on eigenvalues, the other based on Gaussian distributed data. The second measure is discussed in more detail assuming an approximate factor structure in the data; a test procedure is given whether an additional variable adds information. Ordering the series of the data set according to these measures enables us to identify a subset of the data set that is most informative to modelling a variable of interest. We illustrate the concepts with the macroeconomic data set consisting of 132 monthly U.S. variables, which can be divided

into different categories. We find that relative information is maximized for 40–50 series if we are interested in modelling industrial production and inflation. Allowing for pure leads only or both leads and lags does not affect the size of the subset to a great extent. For industrial production, the relative information measure first selects a group of industrial production variables, followed by employment series, interest rates and spreads, and housing starts and sales. With inflation as target variable, the relative information measure starts with picking price indices, followed by employment, orders, interest rates and spreads, housing starts and sales, and employment. As an empirical result, variables from all categories are included in the optimal data set.

8.7 Concluding remarks

As some leading indicators possess significant forecasting power for inflation and GDP growth for some countries during some time periods, it is though impossible to identify one single indicator that shows a consistently good forecasting performance for all countries over all time periods. The approach of the medium scaled multivariate VAR models is most convenient in case out of a large set of candidate variables, only not too many aggregate key variables exhibit significant relationships. In case of inflation forecasting, the exogenous wage developments form a prominent long run model anchor for future inflation as the sluggishness of the wage formation process makes near future wages well predictable and even partly observable. For the disaggregate modelling approach of the five subcomponents, wages are only important for the core inflation components. Model specification and selection are the result of the interplay between economic and statistical criteria.

The business cycle is an unobserved phenomenon that is present in many macroeconomic series. The underlying notion is that the available variables at a disaggregate level contain more information than each available economic key variable at an aggregate level. Combining leading indicators into a single composite index is therefore useful to pick up signals originating from different sectors in the economy. The selection of the underlying single indicator variables is again based on both economic and statistical criteria. Economic plausibility requires that the leading indicators should either cause business cycle fluctuations or quickly react to positive or negative shocks.

The composite leading index consists of the average of the cyclical components of 11 leading indicator variables after synchronisation and standardisation. The composite index provides better forecasts for the growth rates of GDP than its underlying single indicator variables. The three transformations consisting of i) the signal extraction procedure that results in the trend-

cycle decomposition; ii) the synchronisation of the cycles based on the dynamic correlation structure; and iii) averaging the synchronised cycles using their inverse standard deviations as weights improve upon the forecasting performance. This three transformations procedure is further generalized by the cyclical dynamic factor method, where the weights after standardisation are additionally determined by the dynamic eigenvectors. The factor forecasting results for GDP growth are best for the most restricted factor specification, which encompasses both dynamics and cyclicality, and moreover, imposes the factor structure on the forecast equation, even though less restricted factor specifications show better model diagnostics. The marginal improvement of the factor diagnostics resulting from adding (or deleting) an additional variable to the data set progresses upon the forecasting performance. Finally, qualitatively similar results hold for forecasting the growth rates of staffing labor turnover. The underlying data set of disaggregate turnover of staffing labor is closed in the sense that the aggregate is by definition the sum of the disaggregates. Forecasting the aggregate directly using disaggregate information summarized by the common factors turns out to outperform aggregating disaggregate forecasts.

Bibliography

- Achuthan, L., and A. Banerji 2004. *Beating the business cycle: How to predict and profit from turning points in the economy*. Doubleday. ISBN 0-385-50953-7.
- Agresti, A.M., and B. Mojon 2003. "Some Stylised Facts on the Euro Area Business Cycle" In *Monetary Policy Transmission in the Euro Area*, ed. A Kashyap I. Angeloni and B. Mojon Cambridge University Press pp. 15–35.
- Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin, and G. Veronese 2001. "A Real Time Coincident Indicator of the Euro Area Business Cycle" *Discussion Paper* 3108, Center for Economic Policy Research.
- Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese 2006. "New Eurocoin: Tracking Economic Growth in Real Time" *Discussion Paper* 5633, Center for Economic Policy Research.
- Artis, M.J., A. Banerjee, and M. Marcellino 2002. "Factor Forecasts for the UK" *Discussion Paper* 3119, Center for Economic Policy Research.
- Bai, J. 2003. "Inferential Theory for Factor Models of Large Dimensions" *Econometrica* 71, 135–171.
- Bai, J., and S. Ng 2002. "Determining the Number of Factors in Approximate Factor Models" *Econometrica* 70, 191–221.
- 2004. "A PANIC Attack on Unit Roots and Cointegration" *Econometrica* 72 (1), 1127–1177.
- 2008. "Large Dimensional Factor Analysis" *Foundations and Trends in Econometrics* 3, 89–163.
- Banerjee, A., M. Marcellino, and I. Masten 2005. "Leading Indicators for Euro Area Inflation and GDP Growth" *Oxford Bulletin of Economics and Statistics* 67, 785–813.
- Bates, J.M., and C.W.J. Granger 1969. "The Combination of Forecasts" *The Operational Research Quarterly* 20, 451–468.
- Baxter, M., and R.G. King 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series" *The Review of Economics and Statistics* 81, 575–593.

- Benalal, N., B. Diaz del Hoyo, M. Landau, F. Roma, and F. Skudelny 2004. "To Aggregate or not to Aggregate? Euro Area Inflation Forecasting" *Working Paper* 374, European Central Bank.
- Berk, J.M. 1998. "The Information Content of the Yield Curve for Monetary Policy: A Survey" *De Economist* 146, 303–320.
- Berk, J.M., and J.A. Bikker 1995. "International Interdependence of Business Cycles in the Manufacturing Industry: the Use of Leading Indicators for Forecasting and Analysis" *Journal of Forecasting* 14, 1–23.
- Berkhout, E.E., and M.J. Van Leeuwen 2004. "International Database on Employment and Adaptable Labour (IDEAL)" *Technical Report* 642, SEO, Research commissioned by Randstad Holding.
- Bernanke, B.S., and J. Boivin 2003. "Monetary Policy in a Data-Rich Environment" *Journal of Monetary Economics* 50 (3), 525–546.
- Boivin, J., and S. Ng 2005. "Understanding and Comparing Factor-Based Forecasts" *International Journal of Central Banking* 1, 117–151.
- 2006. "Are More Data Always Better for Factor Analysis" *Journal of Econometrics* 132, 169–194.
- Boschan, C., and W. Ebanks 1978. "The Phase-average Trend: a New Way of Measuring Growth" In *1978 Proceedings of the Business and Economic Statistics Section*, ed. G. Elliot, C.W.J. Granger, and A. Timmerman American Statistical Association, Washington D.C.
- Breitung, J., and S. Eickmeier 2006. "Dynamic Factor Models" In *Modern Econometric Analysis, Chapter 3*, ed. G. Huebler and J. Frohn.
- Brillinger, D.R. 1981. *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco. ISBN 0-89871-501-6.
- Bruneau, C., O. De Bandt, A. Flageollet, and E. Michaux 2007. "Forecasting Inflation using Economic Indicators: the Case of France" *Journal of Forecasting* 26 (1), 1–22.
- Bruneau, C., O. De Bandt, and A. Flageollet 2003. "Forecasting Inflation in the Euro Area" *Working Paper* 102, Banque de France.
- Bry, G., and C. Boschan 1971. "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs" *Working Paper* 20, National Bureau of Economic Research.
- Burnham, K.P., and D.R. Anderson 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. New York: Springer.
- Burns, A.C., and W.C. Mitchell 1946. *Measuring Business Cycles*. National Bureau of Economic Research.
- Camacho, M., G. Perez-Quiros, and L. Saiz 2008. "Do European business cycles look like one?" *Journal of Economic Dynamics and Control* 32, 2165–2190.

- Candelon, B., J. Piplack, and S. Straetmans 2009. "Multivariate business cycle synchronisation in small samples" *Oxford Bulletin of Economics and Statistics* 71 (5), 715–737.
- Canova, F. 1998. "Detrending and Business Cycle Facts" *Journal of Monetary Economics* 41, 475–512.
- 1999. "Does Detrending Matter for the Determination of the Reference Cycle and the Selection of Turning Points?" *Economic Journal* 109, 126–150.
- Christiano, L.J., and T.J. Fitzgerald 2003. "The Band Pass Filter" *International Economic Review* 44, 435–465.
- Christoffersen, P.F., and F.X. Diebold 1998. "Cointegration and Long-Horizon Forecasting" *Journal of Business and Economic Statistics* 16, 450–458.
- Clements, M.P., and D.F. Hendry 1998. *Forecasting Economic Time Series*. Cambridge University Press. ISBN 0-19-828700-3.
- Cogley, T., and J.M. Nason 1995. "Effects of Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research" *Journal of Economic Dynamics and Control* 19, 253–278.
- de Groot, E.A., and P.H.B.F. Franses 2005. "Real-Time Estimates of GDP Growth, Based on Two-Regime Models" *Econometric Institute Report 32*, Erasmus University Rotterdam.
- den Reijer, A.H.J. 2006. "The Dutch business cycle: a Finite Sample Approximation of Selected Indicators." *Working Paper 100*, De Nederlandsche Bank.
- Diebold, F., and R. Mariano 1995. "Comparing Predictive Accuracy" *Journal of Business and Economic Statistics* 13, 253–263.
- Diebold, F.X., and G.D. Rudebusch 2003. "A Nonparametric Investigation of Duration Dependence in the American Business Cycle" *Journal of Political Economy* 111, 435–465.
- Dunnewijk, T. 2001. "Temporary Work Agencies in the Netherlands: Emergence and Perspective" *Technical Report 1*, Central Bureau of Policy Analysis.
- Eichler, M., G. Motta, and R. von Sachs 2009. "Fitting Dynamic Factor Models to Non-stationary Time Series" *Research Memorandum 2*, Maastricht Research School of Economics of Technology and Organization.
- Eickmeier, S., and C. Ziegler 2008. "How Successful are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach" *Journal of Forecasting* 27 (3), 237–265.
- Engle, R.F., and S. Kozicki 1993. "Testing for Common Features" *Journal of Business and Economic Statistics* 11 (4), 393–395.
- Espasa, A., E. Senra, and R. Albacete 2002. "Forecasting Inflation in the European Monetary Union: a Disaggregated Approach by Countries and Sectors" *The European Journal of Finance* 8, 402–421.

- Estrella, A., and F.S. Mishkin 1998. "Financial Variables as Leading Indicators Predicting U.S. Recessions" *The Review of Economics and Statistics* 80, 45–61.
- Fase, M.M.G., and J.A. Bikker 1985. "De datering van economische fluctuaties: proeve van een conjunctuurspiegel voor Nederland 1965-1984" *Maandschrift Economie* 49, 299–332.
- Fase, M.M.G., P. Kramer, and W.C. Boeschoten 1992. "MORKMON II, The Nederlandse Bank's Quarterly Model of the Netherlands Economy" *Economic Modelling* 9 (2), 146–204.
- Forni, M., and M. Lippi 2001a. "The Generalized Dynamic Factor Model: Representation Theory" *The Economic Journal* 17, 1113–1141.
- Forni, M., and M. Lippi 2001b. "The Generalized Factor Model: Representation Theory" *Econometric Theory* 17, 1113–1141.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin 2000. "The Generalized Factor Model: Identification and Estimation" *The Review of Economics and Statistics* 82 (4), 540–554.
- 2001. "Coincident and Leading Indicators for the Euro Area" *Economic Journal* 111, 62–85.
- 2004. "The Generalized Factor Model: Consistency and Rates" *Journal of Econometrics* 119 (2), 231–255.
- 2005. "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting" *Journal of the American Statistical Association* 100 (471), 830–840.
- Franses, P.H.B.F., and E.A. de Groot 2005a. "Real Time Estimates of GDP Growth" *Econometric Institute Report 1*, Erasmus University Rotterdam.
- 2005b. "Real Time Estimates of GDP Growth, Based on Two-Regime Models" *Econometric Institute Report 32*, Erasmus University Rotterdam.
- Fritzer, F., G. Moser, and J. Scharler 2002. "Forecasting Austrian HICP and its Components using VAR and ARIMA Models" *Working Paper 73*, Oesterreichische Nationalbank.
- Garcia-Ferrer, A., F. Highfield, F.C. Palm, and A. Zellner 1987. "Macroeconomic Forecasting using Pooled International Data" *Journal of Business and Economic Statistics* 5, 53–67.
- Gengenbach, C., F.C. Palm, and J.-P. Urbain 2006. "Cointegration Testing in Panels with Common Factors" *Oxford Bulletin of Economics and Statistics* 68 (1), 683–719.
- Ghysels, E., A. Sinko, and R. Valkanov 2006. "MIDAS Regressions: Further results and new directions" *Econometric Reviews* 26, 53–90.
- Goldschmeding, F.J.D. 2003. "Ervaringen in de Uitzendbranche. Een casus: Randstad." In *Motoriek van de Economie. Hoe kan het Bedrijfsleven Inspelen op Economische Ontwikkelingen?*, ed. E.A. de Groot.
- Goodwin, G.C., and R.L. Payne 1977. *Dynamic System Identification: Experiment Design and Data Analysis*. New York, London: Academic Press.

- Gottfried, H. 1992. "In the Margins: Flexibility as a Mode of Regulation in the Temporary Help Service Industry" *Work, Employment and Society* 6, 443–460.
- Granger, C.W.J. 1966. "The Typical Spectral Shape of an Economic Variable" *Econometrica* 34 (1), 150–161.
- Grenouilleau, D. 2004. "A Sorted Leading Indicators Dynamic (SLID) Factor Model for Short Run Euro Area GDP Forecasting" *Economic Paper* 219, European Commission.
- Guay, A., and P. St.-Amant 2005. "DO the Hodrick-Prescott and Baxter-King filters provide a good approximation of business cycles?" *Annales d'Économie et de Statistique* 77, 133–155.
- Gómez, V. 2001. "The Use of Butterworth Filter for Trend and Cycle Estimation in Economic Time Series" *Journal of Business and Economic Statistics* 19, 365–373.
- Gómez, V., and A. Maravall 1996. "Programs TRAMO and SEATS, Instructions for the user (Beta Version September 1997)" *Working Paper* 9628, Bank of Spain.
- Hall, R.E., M. Feldstein, J. Frankel, R. Gordon, C. Romer, D. Romer, and V. Zarnowitz 2003. "The NBER's Recession Dating Procedure" Business Cycle Dating Committee, National Bureau of Economic Research.
- Hamilton, J.D. 1994. *Time Series Analysis*. Princeton University Press. ISBN 0-691-04289-6.
- Harding, D., and A.R. Pagan 2001. "Extracting, Analysing and Using Cyclical Information" mimeo.
- 2002. "Dissecting the Cycle: a Methodological Investigation" *Journal of Monetary Economics* 49, 365–381.
- 2005a. "A Suggested Framework for Classifying the Modes of Cycle Research" *Journal of Applied Econometrics* 20, 151–159.
- 2005b. "Synchronisation of Cycles" *Journal of Econometrics* 127 (1), 59–79.
- Harvey, A.C., and A. Jaeger 1993. "Detrending, Stylized Facts and the Business Cycle" *Journal of Applied Econometrics* 8 (3), 231–247.
- Harvey, A.C., and T.M. Trimbur 2003. "General Model-Based Filters for Extracting Cycles and Trends in Economic Time Series" *The Review of Economics and Statistics* 85, 244–255.
- Harvey, D., S. Leybourne, and P. Newbold 1997. "Testing the Equality of Prediction Mean Squared Errors" *International Journal of Forecasting* 13, 281–291.
- Hecq, A., F.C. Palm, and J.-P. Urbain 2002. "Permanent-Transitory Decomposition in VAR Models with Cointegration and Common Cycles" *Oxford Bulletin of Economics and Statistics* 21 (3), 273–307.
- Hendry, D.F., and G.E. Mizon 2000. "On Selecting Policy Analysis Models by Forecast Accuracy" In *Putting Economics to Work: Volume in Honour of Michio Morishima*, ed. H. Glennerster A.B. Atkinson and N.H. Stern Elsevier - North Holland.

- Hendry, D.F., and K.S.E.M. Hubrich 2006. "Forecasting Economic Aggregates by Disaggregates" *Working Paper* 589, European Central Bank.
- Hodrick, R.J., and E.C. Prescott 1997. "Post-War U.S. Business Cycles: An Empirical Investigation" *Journal of Money, Credit, and Banking* 29, 1–16.
- Hoogstrate, A.J., G.A. Pfann, and F.C. Palm 2000. "Pooling in Dynamic Panel Data Models" *Journal of Business and Economic Statistics* 18, 274–283.
- Horowitz, J.L. 2001. "The Bootstrap" In *Handbook of Econometrics*, ed. J.J. Heckman and E.E. Leamer Elsevier - North Holland pp. 3159–3228.
- Hubrich, K.S.E.M. 2005. "Forecasting Euro Area Inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy?" *International Journal of Forecasting* 21 (1), 119–136.
- Inklaar, R.C., J.P.A.M. Jacobs, and W.E. Romp 2004. "Business cycle indexes: Does a heap of data help?" *Journal of Business Cycle Measurement and Analysis* 1, 309–336.
- Jacobs, J.P.A.M. 1998. *Econometric Business Cycle Research*. Kluwer Academic Publishers.
- Jacobs, J.P.A.M., and P.W. Otter 2008. "Determining the Number of Factors and Lag Order in Dynamic Factor Models: A Minimum Entropy Approach" *Econometric Reviews* 27, 385–397.
- Jansen, W.J., and N.J. Nahuis 2003. "The Stock Market and Consumer Confidence: European Evidence" *Economic Letters* 79, 89–98.
- Katz, L.F., and A.B. Krueger 1999. "The High-Pressure U.S. Labor Market of the 1990s" *Brookings Papers on Economic Activity* 1, 1–65.
- Kilian, L. 1998. "Small-Sample Confidence Intervals for Impulse Response Functions" *The Review of Economics and Statistics* 80 (2), 218–230.
- King, R.G., and S.T. Rebelo 1993. "Low Frequency Filtering and Business Cycles" *Journal of Economic Dynamics and Control* 17, 207–232.
- Koopmans, T.C. 1947. "Measurement without Theory" *The Review of Economics and Statistics* 29, 161–179.
- Kranendonk, H., J. Bonenkamp, and J. Verbruggen 2005. "A Leading Indicator for the Dutch Economy: Methodological and Empirical Revision of the CPB System" In *IFO Survey Data in Business Cycle and Monetary Policy Analysis*, ed. J.E. Sturm and T. Wollmershauser Springer-Verlag.
- Kvasnicka, M. 2003. "Inside the Black Box of Temporary Help Agencies" *Discussion Paper*, Humboldt University Berlin.
- Leser, C.E.V. 1961. "A simple method for trend construction" *Journal of the Royal Statistical Society* 23 (1), 91–107.
- Lütkepohl, H. 1987. *Forecasting Aggregated Vector ARMA Processes*. Springer-Verlag.
- Maravall, A., and A. del Rio 2001. "Time aggregation and the Hodrick-Prescott filter" *Working Paper* 8, Bank of Spain.

- Marcellino, M. 2006a. "Handbook of Economic Forecasting" In *Leading Indicators: What Have We Learned?*, ed. C.W.J. Granger G. Elliot and A. Timmerman Elsevier - North Holland.
- 2006b. "Leading Indicators: What Have We Learned?" In *Handbook of Economic Forecasting*, ed. G. Elliot, C.W.J. Granger, and A. Timmerman Elsevier - North Holland.
- Marcellino, M., J.H. Stock, and M.W. Watson 2003. "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area Wide Information" *European Economic Review* 47, 1–18.
- McGuckin, R.H. 2001. *Business Cycle Indicators Handbook*. The Conference Board.
- Mohr, M. 2005. "A Trend-Cycle(-Season) Filter" *Working Paper* 499, European Central Bank.
- Moore, G.H., and J. Shiskin 1967. "Indicators of Business Expansions and Contractions" *Occasional Study* 103, National Bureau of Economic Research.
- Moser, G., F. Rumler, and J. Scharler 2007. "Forecasting Austrian Inflation" *Economic Modelling* 24 (3), 470–480.
- Muirhead, R.J. 1982. *Aspects of Multivariate Statistical Theory*. Mew York: Wiley.
- Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, and E. Giovannini 2008. *Handbook on constructing composite indicators: methodology and user guide*. Organisation for Economic Co-operation and Development. ISBN 978-92-64-04345-9.
- Nilsson, R., and G. Gyomai 2008. "Cycle extraction: A comparison of the Phase-Average Trend method, the Hodrick-Prescott and Christiano-Fitzgerald filters" *Working Paper*, Organisation for Economic Co-operation and Development.
- Ohn, J., L.W. Taylor, and A.R. Pagan 2004. "Testing for Duration Dependence in Economic Cycles" *Econometrics Journal* 7, 569–583.
- Orphanides, A., and S. van Norden 2002. "The Unreliability of Output-Gap Estimates in Real Time" *The Review of Economics and Statistics* 84, 569–583.
- Osborn, D.R., P.J. Perez, and M. Sensier 2005. "Business Cycle Linkages for the G7 Countries: Does the US lead the World?" *Discussion Paper* 50, Centre for Growth and Business Cycle Research, University of Manchester.
- Otter, P.W., and J.P.A.M. Jacobs 2006. "On Information in Static and Dynamic Factor Models" *Working Paper* #2006/5, CCSO. presented at the Far Eastern Meeting of the Econometric Society, July 2006, Beijing, China.
- Palm, F.C., and A. Zellner 1992. "To Combine or not to Combine? Issues of Combining Forecasts" *Journal of Forecasting* 11, 687–701.
- Peck, J., and N. Theodore 2007. "Flexible Recession: the Temporary Staffing Industry and Mediated Work in the United States" *Cambridge Journal of Economics* 31 (2), 171–192.
- Pedersen, T.M. 2001. "The Hodrick-Prescott Filter, the Slutsky Effect and the Distortary Effect of Filters" *Journal of Economic Dynamics and Control* 25, 1081–1101.

- Pollock, D.S.G. 2000. "Trend Estimation and De-trending via Rational Square-Wave Filters" *Journal of Econometrics* 99, 317–334.
- Priestley, M. 1982. *Spectral Analysis and Time Series*. Elsevier - North Holland. ISBN 0-12-564922-3.
- Ravn, M.R., and H. Uhlig 2002. "On Adjusting the HP-Filter for the Frequency of Observations" *The Review of Economics and Statistics* 84, 371–380.
- Rünstler, G., S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze 2009. "Short-term Forecasting of GDP Using Large Scale Monthly Datasets: a Pseudo Real-Time Forecast Evaluation Exercise" *Journal of Forecasting* 28 (7), 595–611.
- Schleicher, C. 2004. "Kolmogorov-Wiener Filters for Finite Time-Series" *Mimeo*, University of British Columbia.
- Schneider, M., and M. Spitzer 2004. "Forecasting Austrian GDP using the Generalized Dynamic Factor Model" *Working Paper* 89, Oesterreichische Nationalbank.
- Schumacher, C. 2007. "Forecasting German GDP using Alternative Factor Models based on Large Datasets" *Journal of Forecasting* 26 (4), 271–302.
- Schumacher, C., and C. Dreger 2004. "Estimating Large-Scale Factor Models for Economic Activity in Germany: Do They Outperform Simpler Models?" *Jahrbücher für Nationalökonomie und Statistik* 224, 732–750.
- Segal, L.M., and D.G. Sullivan 1997. "The Growth of Temporary Services Work" *Journal of Economic Perspectives* 11, 117–136.
- Stock, J.H., and M.W. Watson 1989. "New Indexes of Coincident and Leading Economic Indicators" In *NBER Macroeconomics Annual*, ed. O. Blanchard and S. Fischer MIT Press pp. 351–394.
- 1999. "Forecasting Inflation" *Journal of Monetary Economics* 44, 293–335.
- 2000. "Business Cycle Fluctuations in US Macroeconomic Time Series" In *Handbook of Macroeconomics, volume 1A*, ed. B. Taylor and M. Woodford.
- 2002a. "Forecasting Using Principal Components From a Large Number of Time Predictors" *Journal of the American Statistical Association* 97, 1167–1179.
- 2002b. "Macroeconomic Forecasting using Diffusion Indexes" *Journal of Business and Economic Statistics* 20, 147–162.
- 2003. "Forecasting Output and Inflation: the Role of Asset Prices" *Journal of Economic Literature* 41, 788–829.
- 2005. "Implications of Dynamic Factor Models for VAR Analysis" *Working Paper* 11467, National Bureau of Economic Research.
- Svensson, L.E.O. 1997. "Inflation forecast targeting: Implementing and monitoring inflation targets" *European Economic Review* 41, 1111–1146.
- Theodore, N., and J. Peck 2002. "The Temporary Staffing Industry: Growth Imperatives and Limits to Contingency" *Economic Geography* 78 (4), 463–494.

- Vahid, F., and R.F. Engle 1993. "Common Trends and Common Cycles" *Journal of Applied Econometrics* 8 (4), 341–360.
- van Norden, S. 2002. "Filtering for Current Analysis" *Working Paper* 28, Bank of Canada.
- van Ruth, F., and B. Schouten 2004. "Een Vergelijking van Verschillende Trend-Cyclus Decomposities" *Mimeo*, Statistics Netherlands.
- van Ruth, F., B. Schouten, and R. Wekker 2005. "The Statistics Netherlands' Business Cycle Tracer. Methodological Aspects; Concept, Cycle Computation and Indicator Selection" *Mimeo* 2005-MIC-44, Statistics Netherlands.
- Wallis, K.F. 1999. "Asymmetric Density Forecasts of Inflation and the Bank of England's Fan Chart" *National Institute Economic Review* 167, 106–112.
- Whittaker, E.T. 1923. "On a new method of graduations" *Proceedings of the Edinburgh Mathematical Society* 41, 63–75.
- Wildi, M. 2008. *Real-time signal extraction: beyond maximum likelihood principles*.
- Young, T.Y., and T.W. Calvert 1974. *Classification, Estimation and Pattern Recognition*. New York; London: Elsevier.
- Zarnowitz, V., and A. Ozyildirim 2002. "Time Series Decomposition and Measurement of Business Cycles, Trends and Growth Cycles" *Working Paper* 8736, National Bureau of Economic Research.
- Zellner, A., and F.C. Palm 2004. *The Structural Econometric Time Series Analysis Approach*. Cambridge University Press.

Nederlandse Samenvatting

Het kerndoel van de economische wetenschap is de analyse van de interactie tussen afzonderlijke eenheden, individuen, spelers, organisaties in de tijdsdimensie, seizoenen, jaren, conjunctuurcyclus, levensloop, intergenerationeel en in de ruimtedimensie, steden, agglomeraties, landen, monetaire gebieden en continenten. Voor dit doel zijn grote hoeveelheden aan dataverzamelingen beschikbaar. Denk hierbij aan data van de Nationale Rekeningen voor alle landen, de prijzen van duizenden producten, goederen en diensten, aankopen door individuele consumenten, zoals beschikbaar door scanner data van supermarkt- en warenhuisketens, arbeidsmarktdata voor honderden verschillende soorten banen, zoals omzetdata voor de uitzendmarkt, transactiedata van diverse financiële markten, milieudata omtrent CO₂-uitstoot en klimaatverandering, gezondheidszorg, onderwijs, enzovoorts.

Het huidige informatietijdperk kan worden gekarakteriseerd door de explosie in hoeveelheid en gedetailleerdheid van potentieel beschikbare data, die voornamelijk mogelijk gemaakt wordt door de vooruitgang in de registratie- en opslagtechnologie. Statistische steekproefgrootheden worden niet langer gemeten door het aantal observaties, maar door het aantal (giga)bytes. Bovendien worden observaties met steeds hogere meetfrequenties, maandelijks, wekelijks, dagelijks, per minuut, geregistreerd. Verhandelde financiële titels zijn tegenwoordig al met slechts enkele seconden vertraging volledig publiek beschikbaar. Het informatietijdperk stelt centrale banken voor de uitdaging het monetaire beleid te baseren op de analyse van grote hoeveelheden beschikbare data. Een precieze inschatting van de macro-economische toestand en de ontwikkeling ervan in de afzienbare toekomst is cruciaal in het beslissingsproces van beleidsmakers. Het feit dat centrale banken de moeite nemen om een breed scala aan variabelen te bestuderen toont aan dat beleidsmakers en economische voorspellers

waardevolle informatie destilleren uit de veelheid van tijdreeksen die de economische activiteit op een meer gedetailleerd niveau beschrijven. Het idee is om gebruik te maken van voorlopende conjunctuurindicatoren, die ofwel cyclische beweging veroorzaken ofwel in een vroeg stadium reageren op schokken. Omdat elke conjuncturele neergang en recessie uniek zijn en voortkomen uit een specifiek palet aan oorzaken, tonen sommige voorlopende indicatoren significante voorspelkracht voor specifieke perioden en landen. In het algemeen is het echter ondoenlijk een enkele voorlopende indicator te identificeren, die een robuuste voorspelkracht toont voor meerdere landen en tijdsperioden. Om de signalen, die voortkomen uit verschillende delen van de economie, op te vangen wordt er een verzameling van voorlopende indicatoren samengesteld, welke het gedrag van diverse actoren weergeeft: gezinnen, bedrijven, monetaire en budgettaire autoriteiten en de buitenlandse sector. Het kerndoel van dit proefschrift is om op basis van grote en kleine, specifiek geselecteerde verzamelingen van voorlopende indicatoren voorspelmodellen voor de korte tot middellange termijn te ontwikkelen, die betrekking hebben op voornamelijk Nederlandse macro-economische kernvariabelen, zoals de conjunctuurencyclus, inflatie, de binnenlandse productie en de flexibele uitzendarbeidsmarkt.

Inflatie

Indien het algemene prijspeil stijgt, kan met de functionele munteenheid van een economie minder goederen en diensten worden gekocht. De inflatievoet is de stijging van het prijspeil en leidt tot een daling van de reële waarde van de munteenheid en daarmee tot een verlies aan koopkracht van het ruilmiddel. In de Europese Unie wordt het algemene prijspeil gemeten door de geharmoniseerde index van consumentenprijzen (*Harmonised index of Consumer Prices*, HICP). Het doel van een prijsindex is om de prijzen van een representatief mandje van goederen en diensten te meten, waarvan het gewogen gemiddelde een maatstaf is voor het algemene prijsniveau. De totale HICP index kan worden onderverdeeld in de vijf subcomponenten onbewerkte voedingsmiddelen, bewerkte voedingsmiddelen, industriële producten exclusief energie, energie en diensten. Een beleidsmaker is mede geïnteresseerd in een maatstaf voor kerninflatie, welke kan worden gedefinieerd als de totale HICP index exclusief de componenten onbewerkte voedingsmiddelen en energie. Deze laatste twee volatiele componenten worden op korte termijn meer bepaald door vraag- en aanbodverhoudingen op specifieke markten en zijn daarmee minder ontvankelijk voor monetair beleid.

De voorlopende indicatoren voor inflatoire ontwikkelingen kunnen wor-

den onderverdeeld in: 1) overmatig vraaggedreven: inflatie veroorzaakt door overmatige totale vraag, die voortkomt uit stimulansen van zowel publieke als private consumptie en investeringen; 2) kostendoordrukkend; schoksgewijze inflatie veroorzaakt door aanbodschokken op specifieke grondstoffenmarkten, welke tot uitdrukking komen in producentenprijzen, zoals bijvoorbeeld olie, metalen en graan; 3) zelfvervullende verwachtingen: inflatie veroorzaakt door verwachtingen van economische actoren omtrent toekomstige inflatieontwikkelingen, die zelfrealiserend kunnen worden tijdens bijvoorbeeld loononderhandelingen, de zogenoemde loon-prijsspiraal. In hoofdstuk 2 worden voorspelmodellen ontwikkeld voor de inflatie in Nederland en het eurogebied, zoals gemeten door de totale HICP index en de vijf subcomponenten met een voorspelhorizon tot 18 maanden vooruit. De voorspelmodellen bestaan uit lineaire vector autoregressieve modellen en foutencorrectiemodellen. De prijsindices die een veranderend seizoenspatroon vertonen worden gemodelleerd als maand-op-maand veranderingen van de jaar-op-jaar inflatievoet, waarbij een foutencorrectieterm het mogelijke lange termijn verband tussen inflatie en andere variabelen oplegt. De seizoenspatronen van de prijsindices onbewerkte voedingsmiddelen en energie tonen geen structurele breuk en worden gemodelleerd als maand-op-maand verschillen zonder dat lange termijn relaties worden toegestaan. Vervolgens worden de mogelijke modelconfiguraties, die zijn gebaseerd op een beperkte verzameling van endogene en exogene variabelen doorgerekend. De verschillende modelspecificaties worden beoordeeld op basis van negen verschillende statistische criteria, die zowel de statistische passendheid van het specifieke model meten als afhangen van de gerealiseerde voorspelkracht over een kleine, vooraf bepaalde tijdsperiode. Tenslotte worden de optimale modellen geselecteerd op basis van de statistische criteria en economische consistentie. Zo vormen exogene variabelen een stabiel anker voor de inflatievoorspellingen. De loonontwikkeling is goed exogeen voorspelbaar vanwege het stroperige collectieve onderhandelingsproces en zelfs gedeeltelijk observeerbaar vanwege langlopende en elkaar overlappende contracten voor de verschillende sectoren. De loonontwikkeling wordt daarom als solide anker voor toekomstige inflatie opgenomen, mits statistisch gevalideerd, in de lange termijn relatie met de prijsindices. De geselecteerde modellen voor de verschillende prijsindices vertonen voor de meeste voorspelhorizons een betere voorspelkracht dan de naïeve voorspelling, welke bestaat uit de laatst waargenomen observatie en het optimale autoregressieve model. De voorspelfouten voor de inflatie worden voor een deel verklaard door de realisatie van onverwachte schokken zoals bijvoorbeeld de dierziektes mond- en klauwzeer en de gekke koeien-ziekte en het onregelmatige verloop

van de olieprijs. Hoewel voorspelmodellen niet op onverwachte schokken kunnen anticiperen, worden voorzienbare institutionele veranderingen en basiseffecten, dat wil zeggen uitzonderlijke schokken gedurende het lopende jaar, meegewogen in de uiteindelijke voorspellingen. Zo hadden bijvoorbeeld de introductie van de euromunten en biljetten in januari 2002 en het begin van een omvangrijke prijzenslag in de Nederlandse supermarkten in november 2002 een substantiële invloed op de prijsindices. Het kwantificeren van dusdanig unieke of uitzonderlijke gebeurtenissen maakt van het voorspellen van inflatie zowel een ambacht als een wetenschap.

Conjunctuurcyclus

Conjunctuurcycli zijn oscillerende bewegingen in economische activiteit, welke zichtbaar zijn als eenduidig fluctuerende patronen in macro-economische variabelen zoals productie, werkgelegenheid, financiële markten, consumptie, prijzen en rentevoeten. De cyclische beweging van de variabelen verloopt grotendeels gelijktijdig of in snelle opeenvolging van elkaar. De vier opeenvolgende fases van de conjunctuurgolf volgend op een top bestaan uit i) afkoeling: de groeivoet is bovengemiddeld, maar afnemend; ii) recessie: de groeivoet is benedengemiddeld en afnemend, welke eindigt in een dal; iii) herstel: de groeivoet is benedengemiddeld, maar toenemend; iv) hoogconjunctuur: de groeivoet is bovengemiddeld en toenemend. Opeenvolgende conjunctuurcycli worden van elkaar gescheiden door de omslagpunten: toppen en dalen. Het doel van het derde hoofdstuk is om een conjunctuurindicator voor de Nederlandse economie te construeren¹. De referentiecycclus is gebaseerd op de cyclische beweging met een looptijd langer dan drie jaar en korter dan elf jaar, die zichtbaar is in industriële productie, consumptie door huishoudens en werkgelegenheid zoals vertegenwoordigd door de omzetontwikkelingen in de uitzendsector. Deze drie variabelen staan prominent op het radarscherm van de Amerikaanse en Europese conjunctuurdateringscomités.

Omdat aan opeenvolgende economische fluctuaties verschillende oorzaken ten grondslag liggen, wordt er een enkelvoudige index samengesteld uit een diverse variëteit aan voorlopende indicatoren. De geselecteerde deelindicatoren anticiperen systematisch op conjuncturele omslagpunten met een constante voorlooptijd voor zowel toppen als dalen. De geselecteerde indicatoren reflecteren bovendien het hele patroon van de referentiecycclus.

¹De conjunctuurindicator wordt maandelijks op basis van de nieuwe beschikbaar gekomen data gepubliceerd op de website van De Nederlandsche Bank, zie <http://www.dnb.nl>

De samengestelde voorlopende index is gebaseerd op elf geselecteerde voorlopende deelindicatoren bestaande uit drie financiële reeksen, vier consumenten- en bedrijfsenquêtereeksen en vier variabelen omtrent reële activiteit, waarvan twee aanbod en twee vraag georiënteerd. Het doel van het vierde hoofdstuk is om de cyclus in de verwerkende industrie voor negen landen, te weten België, Duitsland, Frankrijk, Italië, Japan, Nederland, Spanje, VK en VS, te meten, te vergelijken en te voorspellen. De cyclische omslagpunten, periodes van lage en hoge groei en beschrijvende statistiek betreffende eigenschappen als amplitude, steilheid and looptijd van de cyclus worden voor elk land gedocumenteerd en onderling vergeleken. Scherpe cyclische uitslagen, zoals gemeten in overmatige kurtosis, komen in bijna alle landen relatief vaak voor. De cyclische beweging in de verwerkende industrie in Nederland, en in mindere mate in België verloopt relatief snel van lage naar hoge groeiperiodes en omgekeerd. De VS vertonen een patroon van effen ronde toppen en puntige diepe dalen. Looptijdafhankelijkheid betekent dat de kans dat de hoge of lage groeiperiode de volgende periode eindigt groeit naarmate de betreffende periode langer duurt. De hypthese van looptijdafhankelijkheid wordt voor bijna alle landen voor zowel hoge als lage groeiperiodes verworpen. De internationale verbanden tussen de cycli in de verwerkende industrie van de verschillende landen is onderzocht door de fractie van de tijd te bepalen waarin twee landen zich allebei in een hoge of in een lage groeiperiode bevinden. De gemeten fractie laat zien dat de verwerkende industrie in de VK meer gesynchroniseerd is met die in de VS dan met het samengesteld equivalent van de vijf grootste landen van het eurogebied. Bovendien is de gemiddelde synchronisatie tussen het VK en de landen van het eurogebied lager dan van de landen van het eurogebied onderling.

Bruto binnenlands product

Het bruto binnenlands product (bbp) is de standaard maatstaf voor het aggregaat van alle economische activiteit. Het bbp voldoet aan de macro-economische boekhoudkundige identiteit en kan op drie manieren worden samengesteld: 1) uitgaven benadering: som van finale consumptie, kapitaalvorming en uitvoer minus invoer; 2) productiebenadering: som van toegevoegde waarden, dat is de productie minus inkoop, voor elk stadium in de productieketen voor alle industrieën; 3) inkomensbenadering: de som van beloningen van werknemers, bruto winsten van bedrijven en belastingen minus subsidies op productie en invoer. Hoewel het bbp de geaggregeerde economische activiteit vertegenwoordigt, verloopt de cyclische beweging niet

altijd synchroon aan die van de onderliggende componenten, welke zelfs anti-cyclisch kunnen zijn. Het doel van hoofdstuk 5 is om de voorlopende karakteristieken van de onderliggende componenten aan te wenden om het Nederlandse bbp te voorspellen voor een voorspelhorizon tot 4 kwartalen vooruit. Een algemeen aanvaard axioma in de macro-economie is dat de gezamenlijke beweeglijkheid van economische variabelen voor een groot gedeelte wordt veroorzaakt door een relatief beperkt aantal structurele schokken, zoals bijvoorbeeld technologie, monetair beleid, grondstoffenprijzen, enz. De onderliggende notie dat de economische beweging door een beperkt aantal krachten wordt gedreven impliceert dat de informatie, die besloten ligt in elke economische kernvariabele op geaggregeerd niveau minder informatief is over economisch gedrag dan de informatie, die besloten ligt in de beschikbare variabelen op gedetailleerd niveau. Een diffusieindex bevat de gezamenlijke beweging die het meest wijdverbreid is over de beschikbare indicatorvariabelen op gedetailleerd niveau en vertegenwoordigt daarmee de niet-geobserveerde schok ofwel de gezamenlijke factor. De gezamenlijke factoren kunnen worden geschat met principale componenten. In hoofdstuk 5 worden verschillende specificaties voor de factoren en de voorspelvergelijking die de factoren relateert aan de doelvariabele vergeleken. De voorspelel exercitie bestaat achtereenvolgens uit modelselectie, die is gebaseerd op diagnostische statistiek, het schatten van de niet-geobserveerde factoren, het schatten van de parameters en uiteindelijk het genereren van factormodel voorspellingen en met factoren uitgebreide autoregressieve voorspellingen. Deze laatste specificatie is het minst gebaseerd op de factorstructuur en is daarmee het meest flexibel om zich aan de karakteristieken van de data aan te passen. De resultaten laten zien dat het opleggen van de factorstructuur, ondanks de mindere diagnostiek, leidt tot betere voorspelprestaties. De voorspellingen van het factormodel tonen consistent betere prestaties dan de met factoren uitgebreide autoregressieve voorspellingen voor de verschillende specificaties van de factoren en voor alle voorspelhorizons, met name de eerste. De best presterende specificatie is het gerestricteerde cyclische dynamische factormodel, welke het meest uitvoerig de factorstructuur oplegt en zowel dynamiek als cycliciteit omvat en de factorstructuur oplegt op de voorspelvergelijking.

Uitzendwerk

De uitzendsector wordt gekarakteriseerd door de driehoeksverhouding tussen het werkverschaffende bedrijf, de werknemer en de private arbeidsbemiddelaar. Een uitzendbureau is een private marktpartij die optreedt als

bemiddelaar tussen de tijdelijke vraag en aanbod van arbeid. De flexibele uitzendkracht transformeert voor het werkverschaffende bedrijf arbeid tot een variabele productiefactor. De omzet van de totale uitzendindustrie expandeerde gedurende de laatste dertig jaar krachtig voorafgaand aan macro-economische expansieve periodes, terwijl scherpe omzetsdalingen gevolgd werden door recessie. De fluctuaties in de totale uitzendmarkt zijn daarmee tijdig beschikbare voorlopende indicatoren voor brede conjunctuurbeweging. Het doel van hoofdstuk 6 is om de cyclische ontwikkelingen in de uitzendmarkt op gedesaggregeerd niveau te documenteren en de regio's en sectoren te identificeren die voorlopende karakteristieken vertonen. De analyse is gebaseerd op omzetgegevens van Randstad Nederland, die met een historisch constant aandeel van 40% de grootste speler is op de Nederlandse markt. De observaties bestaan uit vier dimensies: het aantal uitzenduren per administratieve periode van vier weken voor 15 verschillende regio's, bestaande uit de twaalf provincies en de drie grootste steden afzonderlijk, en 58 verschillende sectoren. Het factor-model wordt toegepast om de gemeenschappelijke beweging uit de data te filteren, welke overeenkomt met de jaar-op-jaar groeivoet van het aantal uitzenduren op landelijk geaggregeerd niveau. Het gemeenschappelijke signaal, dat gevrijwaard is van regio- en sectorspecifieke schokken, wordt ook op gedesaggregeerd niveau bepaald. Op basis van de correlatiestructuur worden de cycli geclassificeerd als achterlopend, gelijktijdig en voorlopend aan de hand van acht empirische maatstaven. Bijna alle regio's tonen een voorloop- of een achterlooptijd voor de uitzendcyclus van minder dan een halfjaar. De regio's met de meest robuuste voorlooptijd volgens de verschillende criteria zijn Gelderland en Overijssel. Bijna alle sectoren tonen een voorlooptijd van minder dan 1.5 jaar en een achterlooptijd van minder dan 2 jaar. De verschillen tussen de sectoren onderling zijn scherper dan tussen de regio's. De drie sectoren met de meest robuuste voorlooptijd volgens de verschillende criteria zijn: dienstverlening voor het vervoer; handel in en reparatie van auto's en motorfietsen; detailhandel en reparatie van consumentenartikelen. De omzet van de twee laatstgenoemde sectoren staan bekend als klassieke voorlopende conjunctuurindicatoren. De vervolgvraag is hoe de geïdentificeerde regio's en sectoren die voorlopende karakteristieken tonen aangewend kunnen worden om het aggregaat van het aantal uitzenduren op landelijk niveau te voorspellen. Het beschikbare databestand is volledig in de zin dat het landelijke aggregaat per definitie bestaat uit de som van de 870 mogelijke combinaties van 15 regio's en 58 sectoren. Het volledige databestand maakt het mogelijk om de directe voorspellingen van het aggregaat te vergelijken met het aggregeren van voorspellingen op gedesaggregeerd niveau. De vergelijking toont dat het direct

voorspellen van het aggregaat gebaseerd op gedetailleerde informatie beter resultaten oplevert dan het aggregeren van gedesaggregeerde voorspellingen, die per definitie zijn gebaseerd op gedetailleerde informatie. De directe voorspellingen voor het aggregaat omvat de informatie op gedetailleerd niveau zoals dat door de factoren wordt samengevat. De best presterende modellen leggen daarbij een direct verband tussen de gedesaggregeerde data en het te voorspellen geaggregeerde equivalent op landelijk niveau. De best presterende modelspecificatie is het dynamische factormodel, dat in staat is om de substantiële temporele en seizoensmatige variatie in de uitzenddata te beschrijven.

Datadimensie en factorstructuur

De geëxtraheerde factoren beschrijven het onderliggende databestand, welke geacht wordt de belangrijkste krachten voor de doelvariabele te omvatten. Overbemonstering verwijst naar de situatie waarin de data meer informatie bevatten over sommige factoren dan over andere. Het toevoegen van nieuwe variabelen in een overbemonsterd databestand kan resulteren in preciezere extrahering van de factoren, die echter niet leidt tot een verbeterde voorspelprestatie indien de doelvariabelen afhangen van de minder dominante factoren. De factordiagnostiek voor de extrahering van de factoren is gerelateerd aan de dominantie van de gezamenlijke component, dat is de gezamenlijk-tot-idiosyncratische variantieratio en de spreiding ervan over de variabelen in het databestand. In hoofdstuk 5 wordt de samenhang tussen de op factoren gebaseerde voorspellingen voor verschillende samenstellingen in grootte en structuur van het databestand empirisch onderzocht. De resultaten tonen aan de grootte ertoe doet, omdat het aanwenden van alle beschikbare voorlopende indicatoren tot betere prestaties leidt voor alle voorspelhorizons dan de datasamenstellingen bestaande uit individuele groepen van variabelen met een identiek karakter. De kleinere macro-economische dataverzamelingen kennen een sterkere factorstructuur, welke leidt tot preciezer geëxtraheerde factoren, die echter niet direct samenhangen met de doelvariabele. Overbemonstering, of beter gezegd misbemonstering verwijst naar de misvertegenwoordiging van een klein databestand met een sterke factorstructuur. De resultaten tonen aan dat, uitgaand van een omvangrijk databestand als de referentie voor de gerespecteerde factorvoorspellingen, de marginale verbetering in de factordiagnostiek veroorzaakt door het buitensluiten van een specifieke groep van variabelen, leidt tot verbeterde voorspelprestaties. Het uitsluiten van de categorie bestaande uit enquête variabelen leidt bijvoorbeeld tot betere prestaties. De categorie variabelen bestaande uit enquêtes stellen het

dataverzamelingsproces bloot aan overbemonstering, waardoor de voorspelprestaties afnemen. Deze categorie bevat een idiosyncratisch signaal, welke complementair is aan de gezamenlijke beweging in het volledige databestand zoals geëxtraheerd door de gezamenlijke factoren. Factormodellen zijn ontwikkeld als een instrument om veel variabelen tegelijkertijd te kunnen analyseren zonder in de vrijheidsgradenproblematiek te geraken, die zo karakteristiek is voor regressie-analyse. Een groot databestand moet echter niet verder uitgebreid worden indien de additionele kandidaatvariabele een hoge correlatie paart aan een lage gezamenlijke-tot-idiosyncratische variantieratio of indien de additionele kandidaatvariabele het databestand blootstelt aan overbemonstering. Het aantal en de keuze van de variabelen om op te nemen in een databestand, dat is gerelateerd aan een bepaalde doelvariabele wordt geformaliseerd in hoofdstuk 7. Concepten uit de informatietheorie, in het bijzonder Kullback-Leibler criteria, worden toegepast om de hoeveelheid informatie in een databestand te kwantificeren, welke afhangt van de grootte en de correlatiestructuur van het databestand. Twee relatieve informatiemaatstaven worden geïntroduceerd, waarbij de ene gebaseerd is op eigenwaarden en de andere op Gaussisch verdeelde data. Voor deze tweede maatstaf wordt een toetsprocedure ontwikkeld, die onder de aanname dat de data gekarakteriseerd wordt door een factorstructuur bepaalt of een additionele variabele daadwerkelijk informatie toevoegt. Door de tijdreeksvariabelen in het databestand te ordenen volgens de twee informatiemaatstaven is het mogelijk een subgroep van de volledige dataverzameling te identificeren, die de meeste informatie bevat betreffende een doelvariabele. De procedure kan worden toegepast als een eerste stap voor de constructie van een diffusieindex of een samengestelde voorlopende indicator.

Concluderende opmerkingen

Ondanks dat sommige voorlopende indicatoren significante voorspellende waarde bezitten voor inflatie en de groeivoet van het bbp voor sommige landen gedurende bepaalde tijdsperioden is het desalniettemin onmogelijk om een enkele indicator variabele te identificeren die een consistent goede voorspelkracht toont voor alle landen gedurende alle tijdsperioden. Het toepassen van het vector autoregressieve model is het meest geschikt indien een beperkt aantal macro-economische kernaggregaten significante statistische verbanden tonen. Voor het voorspellen van inflatie spelen de exogene loonontwikkelingen een prominente rol als modelanker voor de middellange termijn. De toekomstige lonen zijn goed voorspelbaar en zelfs gedeeltelijk observeerbaar vanwege de stroperigheid van het loononderhandelingsproces.

De loonvorming is alleen voor de aan kerninflatie gerelateerde onderliggende inflatiecomponenten belangrijk. De specificatie en selectie van het model komt voort uit het samenspel van economische en statistische criteria. De conjunctuurcyclus is een niet geobserveerd fenomeen dat in vele macro-economische tijdreeksen aanwezig is. De onderliggende notie is dat de beschikbare variabelen op gedesaggregeerd niveau meer informatie omvatten dan elk van de beschikbare kernvariabelen op geaggregeerd niveau. Het combineren van voorlopende indicatoren in een samengestelde index is daarmee nuttig om de signalen, die uit verschillende sectoren voortkomen, op te pikken. De selectie van de onderliggende indicatorvariabelen is wederom gebaseerd op zowel economische als statistische criteria. De voorlopende indicator moet een economisch plausibele veroorzaker zijn of snel reageren op positieve en negatieve schokken. De samengestelde index bestaat uit het gemiddelde van de cyclische componenten van 11 voorlopende indicatorvariabelen nadat deze zijn gesynchroniseerd en gestandaardiseerd. De samengestelde index levert betere voorspelprestaties voor de groeivoet van het bbp dan de individuele onderliggende indicatoren. De drie transformaties bestaande uit i) het uitfilteren van het signaal, welke resulteert in de decompositie van de trend en de cyclus; ii) het synchroniseren van de cycli gebaseerd op de dynamische correlatiestructuur; en iii) het gewogen middelen van de gesynchroniseerde cycli, waarbij de gewichten bestaan uit de inverse van de standaardafwijking, verbeteren de voorspelprestaties. Het cyclische dynamische factormodel generaliseert deze drie transformaties, waarbij de gewichten na standaardisatie worden bepaald door de dynamische eigenvectoren. De factorvoorspellingen voor de groeivoet van het bbp zijn het beste voor de meest gerespecteerde specificatie van het factormodel, welke zowel de cyclus uitfiltert als op basis van de dynamische correlatiestructuur synchroniseert en bovendien de factorstructuur oplegt op de voorspelvergelijking. Terwijl minder strikt gespecificeerde factor modellen een betere modeldiagnostiek tonen. De marginale verbetering van de factorstructuur, die voortkomt uit het toevoegen (of verwijderen) van een additionele variabele aan het databestand zorgt voor een verbeterd voorspellend vermogen. De resultaten van het voorspellen van de groeivoeten van het aantal uitzendingen zijn kwalitatief hetzelfde als die voor het bbp. Het onderliggende databestand bestaande uit het aantal uitzendingen op gedetailleerd niveau is volledig in de zin dat het aggregaat per definitie de som van alle gedesaggregeerde waarnemingen is. Het direct voorspellen van het aggregaat, waarbij gebruik gemaakt wordt van gedesaggregeerde informatie zoals samengevat door de gezamenlijke factoren levert betere prestaties dan het aggregeren van gedesaggregeerde voorspellingen.