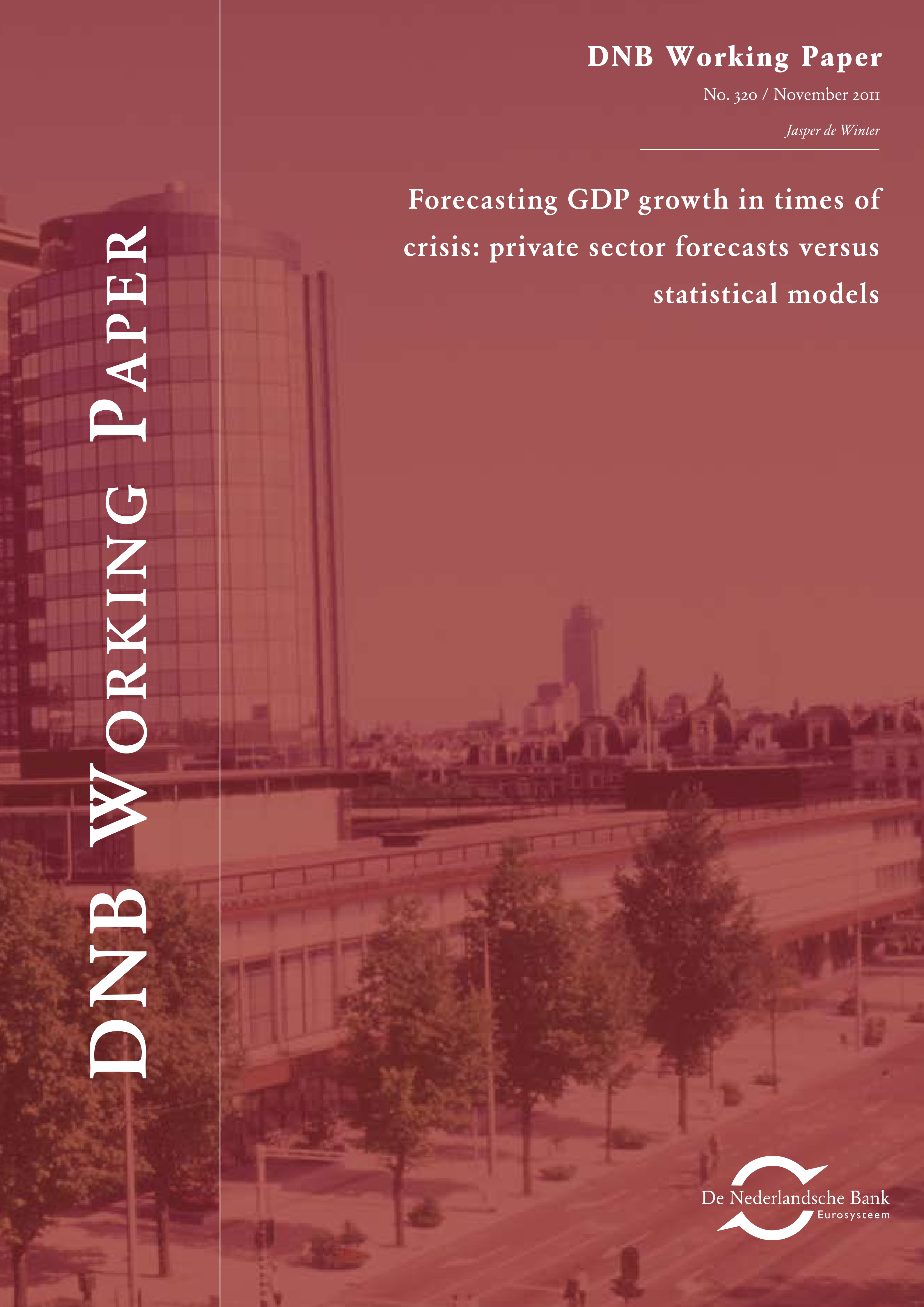


DNB Working Paper

DNB Working Paper

No. 320 / November 2011

Jasper de Winter

Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models

De Nederlandsche Bank
Eurosysteem

Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models

Jasper de Winter *

* Views expressed are those of the author and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 320

November 2011

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models

Jasper de Winter[a]

Version: 7 November 2011

Abstract

This paper examines the accuracy of short run forecasts of Dutch GDP growth by several linear statistical models and private sector analysts. We focus on the financial crisis of 2008-2009 and the dot-com recession of 2001-2002. The dynamic factor model turns out to be the best model. Its forecast accuracy during the crisis deteriorates much less than that of the other linear models and hardly at all when backcasting and nowcasting. Moreover, the dynamic factor model beats the private sector forecasters at nowcasting. This finding suggests that adding judgement to a mechanical model may not improve short-term forecasting performance.

JEL classifications: E52, C53, C33.

Keywords: Nowcasting, Professional Forecasters, Factor Model, Forecasting.

[a] De Nederlandsche Bank, P.O. Box 98, 1000 AB Amsterdam, the Netherlands. Tel: +31-20-524 22 42, Fax: +31-20-524506, E-mail: J.M.de.Winter@dnb.nl. Comments and suggestions by Marta Bańbura, Jos Jansen, Pierre Lafourcade, Job Swank, and seminar participants at the 2011 ECB Expert Meeting on Activity Forecasting and the 2011 *Conjunctuurdag* are gratefully acknowledged. Special thanks to Marta Bańbura for sharing elements of her Matlab-code. The views expressed in this paper do not necessarily reflect those of De Nederlandsche Bank. The usual disclaimer applies.

1. Introduction

The recession of 2008-2009 marked the largest fall in Dutch Gross Domestic Product (GDP) since the Second World War. After peaking in the first quarter of 2008, output declined for five consecutive quarters by 4.8 pp in total. Although some indicators clearly revealed the build up of imbalances that usually precede a crisis (Reinhart and Rogoff, 2008), forecasters across the world failed to forecast the depth and duration of the crisis (e.g. IMF, 2011). This also holds for the Netherlands (De Jong et al. 2010).

This paper examines if short-term forecasting models could have been helpful in forecasting the particular dynamics of the recent crisis and, more generally, recessions. This is a highly relevant issue for policy makers and economic agents alike, as information on where the economy stands and where it is heading in the short run is particularly valuable in times of severe crises. The recent literature on forecasting GDP in the short run, which extensively uses leading indicators, (e.g. Rünstler et al., 2009, Bańbura et al., 2011 and Baffigi et al., 2004) has not explicitly analysed the financial crisis[1]. This paper aims to partially fill this gap in the empirical literature by an in-depth analysis of the Dutch case. Furthermore, it contributes to the literature by including forecasts by professional forecasters into the analysis. This part aims to get an indication of the potential gains of adjusting the mechanical projections of statistical models by subjective judgement.

The first part of the paper conducts a "horse-race" between several well-known linear models. Our empirical application for the Netherlands focuses on the forecast performance during the financial crisis of 2008-2009 and the dot-com recession of 2001-2002. The latter episode serves as a benchmark for a "normal" recession. The second part of the paper evaluates the linear models' performance against forecasts of quarterly GDP by professional forecasters, which are collected by Consensus Economics. To the best of our knowledge, these *quarterly* forecasts have never been analysed in the literature before. Although there is a rich literature on the accuracy of Consensus Forecasts' *annual* forecasts (e.g. Ager et al., 2009, Batchelor, 2001 and Loungani, 2001).

[1] Exception is a recent paper by Lombardi and Maier (2011), that evaluates the forecasting performance of a simple bridge model that only uses the Purchasing Managers Index (PMI) versus a "data-rich" dynamic factor model.

The remainder of the paper is structured as follows. The second section describes the Consensus Forecasts data and their behaviour before and during the financial crisis and the dot-com recession. The third section discusses the forecasting models, while the fourth section outlines the empirical setup of the analysis. The fifth section reports the empirical results. The sixth section concludes.

2. Quarterly Consensus Forecasts

Consensus Economics has been collecting and publishing forecasts on a monthly basis under the name of *Consensus Forecasts*. The Consensus Forecasts are best known for the monthly forecast of private analysts' expectations of output growth for a broad range of countries. The monthly surveys report the forecasts for the current and next year for a set of key macroeconomic figures. About nine institutions participate in the poll for the Netherlands, mainly banks. It is less well known that the panellists also provide quarterly forecast to Consensus Economics, released on the last month of each quarter. The panellists supply their forecasts for six consecutive quarters, starting from the first unpublished quarter. The quarterly GDP forecasts for the Netherlands have been published by Consensus Economics since December 1994.

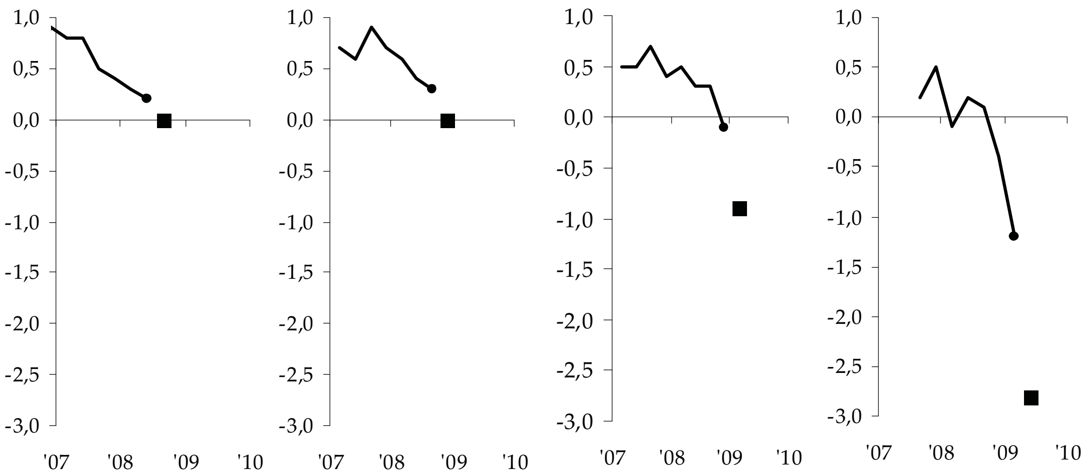
2.1 Consensus Forecasts during recessions

Figure 1 shows the evolving mean or quarterly GDP forecast of the Consensus Forecast for the quarters leading up to the financial crisis of 2008-2009. The horizontal axis indicates the time the forecast was made, while the vertical axis shows the mean forecast[2]. The black circles indicate the last projection before the release of GDP and the black squares indicate the current estimate of GDP growth. The figure clearly shows that the collapse of Lehman Brothers in September 2008 started a sequence of downward adjustments to the quarterly growth rates, but these adjustments ultimately underestimated the size of the downturn, especially in 2008Q4 and 2009Q1. The overestimation of GDP growth seems to be a recurrent feature during downturns as indicated in Figure 2, which plots the quarterly growth forecasts around the dot-com recession of 2001-2002. Again, respondents seem to be sluggish in adjusting their forecast downward.

[2] The quarterly forecasts only report the mean (the "consensus") forecast of the participating panellists. The monthly forecasts also report the forecast of each panellist.

Figure 1: Quarterly forecasts Consensus Economics during the financial crisis*)
Real GDP growth Netherlands, quarter-on-quarter, Netherlands, percent

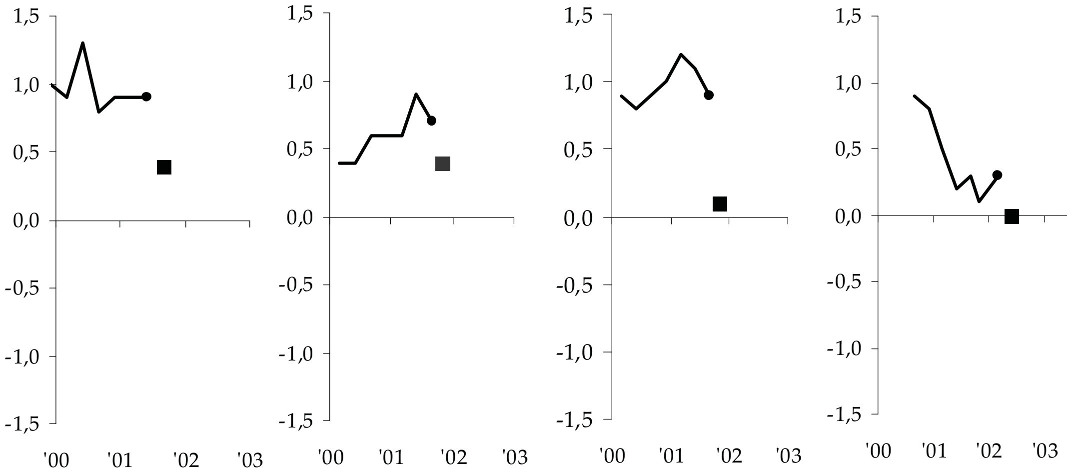
2008Q2	2008Q3	2008Q4	2009Q1



*) The black circles indicate the last Consensus Forecast made before the first release of quarterly GDP growth (black
Source: Consensus Economics.

Figure 2: Quarterly forecasts Consensus Economics during the dot-com recession*)
Real GDP growth Netherlands, quarter-on-quarter, percent

2001Q3	2001Q4	2002Q1	2002Q2



*) The black circles indicate the last Consensus Forecast made before the first release of quarterly GDP growth (black
Source: Consensus Economics.

2.2 Reasons for stickiness in Consensus Forecasts

The stickiness in the *quarterly* Consensus Forecasts is in line with the literature analysing the *annual* Consensus Forecasts. The main conclusion of this literature is that forecasters tend to be excessively cautious and do not revise their forecasts promptly and sufficiently to reflect incoming news (Isiklar et al., 2006), especially news from foreign economies (Loungani et al., 2011). There also appears to be a tendency for herd behaviour in forecasting, possibly owing to forecasters putting a higher weight on the group's shared view than on private priors and incoming information. Loungani et al. (2011) find that the predictive failure in case of crisis-related recessions is larger than during normal recessions. One reason is that forecasters seem to have a tendency to smooth their forecasts, failing to adjust them sufficiently in response to news. All these features also seem to present in the quarterly Consensus Forecast.

Overall, panellists seem to add a considerable part of judgement to incoming information. Unfortunately, it is not known how much judgement they add, but it is probably sizeable. Information from a comparable survey amongst professional forecasters in the euro area (ECB, 2009) indicates panellists regard forty percent of their short-term GDP forecast to be judgement based. These judgemental adjustments are usually added to mechanical model outcomes (mostly simple time-series models), although some respondents indicate that their short-term GDP forecast is completely based on judgement.

Besides these factors, the slow adjustment of the quarterly forecasts can be partly traced back to the difficult task of aggregating all the monthly incoming information in a consistent manner. The problem with the monthly data —such as industrial production, consumer confidence and inflation— is that they only cover a subset of economic activity, and are often "noisy". In addition, the vast availability of these indicators also poses a problem. The main difficulty is that they can —and often do— provide conflicting signals and there is no universally accepted approach to aggregate the data into one single GDP growth figure. Moreover, most monthly indicators have sizeable publication delays, which cause "ragged" edges in the dataset (Bańbura et al., 2011). For instance, at the beginning of 2008Q3, forecasters only had information on GDP growth until 2008Q1. They did have more recent information on industrial production and retail trade, but the latest figure dated back to May 2008. The most recent information stemmed from financial markets, which is available on a daily basis, and survey data, which are promptly available at the end of the month.

An alternative to using the forecaster's judgement-augmented forecasts is to use pure mechanical short-term forecasting models to process the information. The main advantage of these models is that they do not contain any subjective interpretation, but just let the "data speak". Whether this leads to better forecasting results is then an empirical question. The next section presents five linear models often used in the literature to forecast GDP in the short-term.

3. Description of linear models

From an econometric viewpoint, combining all the monthly information is not a straightforward problem. Regression by Ordinary Least Squares does not work because the time-series for GDP are too short and the number of monthly indicators is too large to estimate all coefficients reliably. In the literature, there are basically two approaches to overcome this "curse of dimensionality". One method is to estimate the relationship between GDP and all indicators separately. In a second step, the forecasts are combined (pooled) to produce one aggregate forecast of GDP (e.g. Timmerman, 2006). The second method is to use all monthly indicators simultaneously in one model by extracting the common patterns —or "factors"— out of all monthly indicators. In a second step, GDP is regressed on these factors to forecast GDP. The dynamic factor model approach has been shown to provide relatively accurate forecast in the United States (e.g. Giannone et al., 2008), the euro area (e.g. Bańbura et al., 2011 and Rünstler et al., 2009) and the Netherlands (Den Reijer, 2005). Besides the pooling approach and the dynamic factor model approach, we shortly describe two benchmark models, that only use information contained in GDP[3].

3.1 Extracting information through factors

In this section, we will describe the dynamic factor model specification. For a more comprehensive account of the model structure, we refer to Bańbura and Rünstler (2011), and for more details on the estimation technique to Doz et al. (2011). The model of Bańbura and Rünstler (2011) is tailored to mix the *quarterly* frequency of GDP with the *monthly* frequency of the indicators. Moreover, the model can efficiently take into account the "ragged edge" nature of the data, by using the Kalman filter and smoother. Bańbura and Rünstler (2011) show that it is relatively straightforward to derive the contributions of the indicators to the GDP forecast, based on an algorithm developed by Koopman and Harvey (2003). Consider a vector of monthly series $x_t = (x_{1,t}, x_{2,t} ... x_{n,t})$, $t = 1,.....T$, which have been standardized to mean zero and variance one. Each variable x_t can be represented as the sum of two orthogonal

[3] This section does not aim to give a comprehensive test of *all* linear models. For instance, MIDAS-models (e.g. Andreou et al., 2011) —which can be seen as a special class of bridge equations— have recently gained interest amongst academics and central banks (e.g. Kuzin et al., 2011), but these are outside the scope of this paper.

unobserved components: a common component and an idiosyncratic component. The common component is driven by a small number $r << n$ of unobserved common factors that account for most of the co-movement among the variables. The idiosyncratic component ξ_t is driven by variable-specific shocks. The model can be written as:

(1) $x_t = \Lambda f_t + \xi_t \qquad \xi_t \sim \mathbb{N}(0, \Sigma_\xi)$

where, f_t is a $r \times 1$ vector of common factors, Λ is the $n \times r$ matrix of factor loadings and ξ_t is a $n \times 1$ vector of idiosyncratic components. The idiosyncratic components are assumed to be multivariate white noise, so that the covariance matrix Σ_ξ is diagonal. The factor dynamics are specified as a vector autoregression (VAR) of order p:

(2) $f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + B\eta_t \qquad \eta_t \sim \mathbb{N}(0, I_q)$

The VAR is driven by a q dimensional standardized white noise η_t, where B is a $r \times q$ matrix[4] and A is a square $r \times r$ matrix. Combining the monthly factor model with a forecast equation for mean-adjusted quarterly GDP growth completes the model. For this purpose a latent monthly GDP growth rate y_t^m is introduced, which is related to the monthly factors by the equation:

(3) $y_t^m = \beta' f_t + \varepsilon_t \qquad \varepsilon_t \sim \mathbb{N}(0, \sigma_\varepsilon^2)$

However, GDP is only observed in the third month of each quarter, and has a quarterly frequency, defined as y_t^q. The quarterly growth rates are identical to the 3 month average of the monthly growth rates, but are unobserved in the first two months of the quarter. Therefore:

$y_{3k}^q = \frac{1}{3}(y_{3k}^m + y_{3k-1}^m + y_{3k-2}^m) \qquad y_{3k}^m$ is observed and y_{3k-1}^m and y_{3k-2}^m are unobserved.

Accordingly, the monthly series x_t have been transformed as 3-month growth rates or differences. The description of the model is completed by assuming no cross-correlation at any lead or lag between the innovations ξ_t, η_t and ε_t. After estimation of the coefficients, the model is cast in state-space form as described in Bańbura and Rünstler (2011). The factors are

4 We opted for a specification of $r = 5, q = 4$ and $p = 2$ in equation (2). This choice was based on a specification search over $r \leq 8, q \leq r$ and $p \leq 3$ using the RMSFE as a selection criterion, in line with Matheson (2011) and Rünstler et al. (2009).

then re-estimated using the Kalman filter and smoother. Doz et al. (2011) discusses the estimation process and the consistency properties of the parameters in more detail.

3.2 Combining information by pooling

This section describes two approaches that pool the monthly information by bridge equations and bivariate equation models (VARs).

Pooled bridge equations

The bridge equation is perhaps the most widely used method for forecasting quarterly GDP using monthly indicators. For recent applications see Kitchen and Monaco (2003) and Baffigi, Golinelli and Parigi (2004). The bridge equations are constructed using the following steps: (1) we consider the set x_t of monthly indicators and forecast the individual indicators $x_{i,k}$ over the relevant horizon using a univariate AR(p) model: $x_{i,t} = \sum_{s=1}^{p_i} \rho_s x_{i,t-s} + \varepsilon_{i,t}$; (2) for each indicator $x_{i,t}$, we calculate the quarterly average $x_{i,t}^q$ and estimate the following bridge equation: $y_t^q = c_i + \sum_{s=o}^{q_i} \beta_{i,s} x_{i,t-s}^q + \varepsilon_{i,t}^q$ which relates quarterly GDP to the quarterly growth rate of the monthly indicator. The forecast of GDP growth is obtained by inserting the forecasts of the monthly indicator of the AR(p) model into the bridge equation. The lag length of the independent monthly variables in step (1) and (2) is determined via the Schwarz Information Criterion (SIC). The pooled forecast is a weighted average of the *n* forecasts from the individual bridge equations. The weight of each bridge equation forecast is determined by the relative forecast performance in previous periods, which is measured as the inverse of the out-of-sample Root Mean Squared Forecast Error (RMSFE). This implies that models with a relatively high forecast accuracy have a high weight, while models with low forecasting accuracy have a low weight. The RMSFE is defined as:

$RMSFE_{i,t}^q = \sqrt{\frac{1}{t}\sum_{s=0}^{t}(\hat{y}_{i,s}^q - y_s^q)^2}$ where $\hat{y}_{i,s}^q$ is the estimated quarterly growth rate of variable *i* on time *s*.

Pooled Vector Autoregressive equations

The pooled VAR approach is very similar to the bridge equation approach. As opposed to bridge equations, the VAR models use the information content of GDP growth itself to produce better forecasts of GDP (e.g. Camba-Mendez et al., 2001). We consider the set x_t of monthly indicators and forecast the individual indicators $x_{i,k}$ over the relevant horizon using an AR(p) model, just as we did with the bridge equations. In the second step, we estimate the following quarterly bivariate VAR model on GDP growth and each of the k indicators:

$z_{i,t}^q = c_i + \sum^{p_i} \beta_{i,s} z_{i,t-s}^q + \varepsilon_{i,t}^q$ where $z_{i,t}^q = [y_t^q, x_t^q]$. the pooled forecast of GDP growth is an unweighted average of the k forecasts from the individual VARs. The lag length of the independent monthly variables in both steps is determined via the Schwarz Information Criterion (SIC). Again, the weight of each VAR model forecast is determined by the inverse of the root mean squared forecast error.

3.3 Naïve univariate benchmark models

Beside the dynamic factor model and the pooled approaches we considered two naïve benchmark models. The first naïve benchmark considered is a first-order autoregressive model: $y_t^q = c + \beta y_{t-1}^q + \varepsilon_t^q$, where c is a constant and ε_t^q is quarterly white noise, $\varepsilon_t^q \sim N(0, \sigma_\varepsilon^2)$. The second naïve benchmark is a recursive average GDP growth rate model, $y_t^q = \mu_t$, where μ_t is the average quarterly GDP over the period $1, \ldots. t$;

4. Empirical setup and forecast evaluation

This section describes the dataset that was used to estimate the mechanical linear model (section 4.1), the pseudo real-time setup of the analysis (section 4.2) and the statistical methods to analyse forecasting performance (section 4.3).

4.1 Dataset

The dataset that we used consists of eighty monthly time-series variables that are available from De Nederlandsche Bank's (DNB) internal database, the ECB Statistical Data Warehouse and Statistics Netherlands. The variables are spread over four groups: production & sales (30), surveys (28), information from financial markets (11) and prices (11). Most of the series refer to the Netherlands, but we purposely included series referring to important trading partners of the Netherlands to fully take into account the important role exports have for explaining GDP growth for a small open economy as the Netherlands.

The monthly series are preferably collected on a seasonally- (and calendar effects-) adjusted basis. However, some of the variables are only available in raw format and are seasonally adjusted by applying the US Census' X12-method. All monthly series are made stationary by first-differencing. Series that are expressed as a percentage (interest rates, unemployment) or percentage balance (survey results) are differenced directly. For trending data (such as industrial production, retail sales and monetary aggregates) we take logarithms beforehand, which amounts to calculating rates of change. As is required for estimation of the dynamic factor model, all variables were standardized by subtracting the mean and then dividing by their standard deviation. This normalization is necessary to avoid overweighting of large variance series in the factor estimation. Table A.1 in the appendix provides an overview of all variables, the applied transformations and the publication lag of each indicator.

4.2 Pseudo real-time design

We aim to replicate the availability of the data at the time the forecast was made in order to mimic as closely as possible the real-time flow of information. More precisely, we used a data set downloaded on May 10, 2011 and combined this with the typical data release calendar to reconstruct data for earlier months. We recreated the monthly datasets for the period January 1995 until December 2010. All monthly series start in January 1985, whilst the quarterly GDP series start in 1985Q1[5]. The so-called vintages replicate the availability of the data on the 10th of each month. This approach is called "pseudo" real-time, which means we take into account the publication delays in the data, but not the fact that the data might have been revised.

The parameters of all models are estimated recursively using only the information available at the time of the forecast. This approach is regularly used in empirical studies (e.g. Rünstler et al., 2009, Giannone et al., 2008 and Garratt et al., 2008). More precisely, we consider a sequence of eleven forecasts for GDP growth in a given quarter, obtained in consecutive months. The timing is best explained with an example. Assume that the objective is to forecast GDP growth in the third quarter of 2008. We start forecasting in January 2008. This forecast is called the two-quarter-ahead forecast in month one. We produce a forecast each month, and with the first release of GDP in mid-November, we stop forecasting. Following the usual naming convention *forecasts* refer to one or two quarter ahead forecasts, *nowcasts* refer to current quarter forecasts and *backcast*s refer to "forecasts" for the preceding quarter for which no GDP figures are available. In the context of the above example, we made two

[5] Regarding GDP, two seasonally adjusted quarterly time series are observed; the series starting in 2001Q1 and the "ESA95"-series that covers the period 1977Q1-2004Q4. The full sample series is constructed by backdating the GDP-series over the period 1977Q1-200Q4 using the quarter-on-quarter growth rates of the "ESA95"-series.

quarter ahead forecasts from January to March, one quarter ahead forecasts from April to June, nowcasts from July to September, and backcasts in October and November.

In order to compare the forecasts of the mechanical models to the Consensus Forecast we made one modification to this pseudo real-time set, as the information set available to the forecasters was slightly different because of data revisions. Unfortunately, we cannot replicate the exact dataset available to the forecasters, because we do not have full revision histories for all eighty variables in the sample. However, we do have a full revision history for GDP, since every vintage GDP release has been archived in DNB's database since 1985[6]. Moreover, about fifty percent of the data, i.e. all survey data and financial data, are never revised, so that the effect of revisions on the estimates of the linear model forecasts is mitigated. Lastly, a truly real-time comparison would most likely lead to very comparable outcomes since data revisions across many variables wash out, especially for factor models (Bernanke and Boivin, 2003). In order to compare the Consensus Forecasts to the best linear model, we re-estimated all coefficients and forecasts of the linear model by replacing the *pseudo* real-time vintages for GDP by the real-time vintages. Finally, we calculated the GDP forecast errors of the professional analysts and the linear model against the latest data vintage, as suggested by Orphanides and van Norden (2002).

4.3 Forecast evaluation

We analysed the out-of sample forecasting performance of the linear models and Consensus Forecasts over three sample periods, i.e. (1) the whole sample period (1995Q1:2010Q4), (2) the period excluding the financial crisis (1995Q1:2007Q4), (3) the period excluding the dot-com recession and the financial crisis (1995Q1:2000Q4 and 2003Q4:2007Q4). The latter period leave no recessionary periods in the sample and indicates how well the models perform without having to forecast turning points in the business cycle. The period excluding only the financial crisis sheds light on model performance during "normal" recessions. The period including the financial crisis really puts the linear model to the test, providing insight into which models perform best during exceptional circumstances.

We will use the out-of sample Root Mean Squared Forecast Error (RMSFE) to compare the forecast performance of the models. To assess the statistical significance of differences in forecast performance, we use the Diebold-Mariano test (Diebold-Mariano, 1995). In this

6 See Roodenburg en den Reijer (2006) for a review of the quality of the consecutive Dutch GDP releases.

setup, the null hypothesis of equal mean squared errors is tested against the alternative that the dynamic factor model produces smaller forecast errors. To test if the models are in any way complementary, i.e. if pooling of two model types can increase the forecast accuracy, we will perform an encompassing tests proposed by Fair and Shiller (1990). The procedure is quite straightforward, i.e. estimate the following regression:

$$y_{t+h}^{q} = \alpha_0 + \beta_0^{z}\hat{y}_{t+h,t}^{q} + \beta_0^{DFM}\hat{y}_{t+h,t}^{q} + \varepsilon_0$$

Where *z*= RM, AR, BEQ, VAR and the Consensus Forecasts and *DFM*=dynamic factor model. The alternative models do not contain additional information with respect to the dynamic factor model *if and only if* β_0^{z} is not significantly different from zero and β_0^{DFM} is significantly different from zero (and vice versa for testing the alternative linear model is superior to the dynamic factor model). They both contain valuable information if both coefficients are significantly different from zero. This procedure was used among others by Romer and Romer (2000) to test whether the Fed Greenbook forecasts were superior to private sector forecasts, and more recently by Camacho and Perez-Quiros (2010).

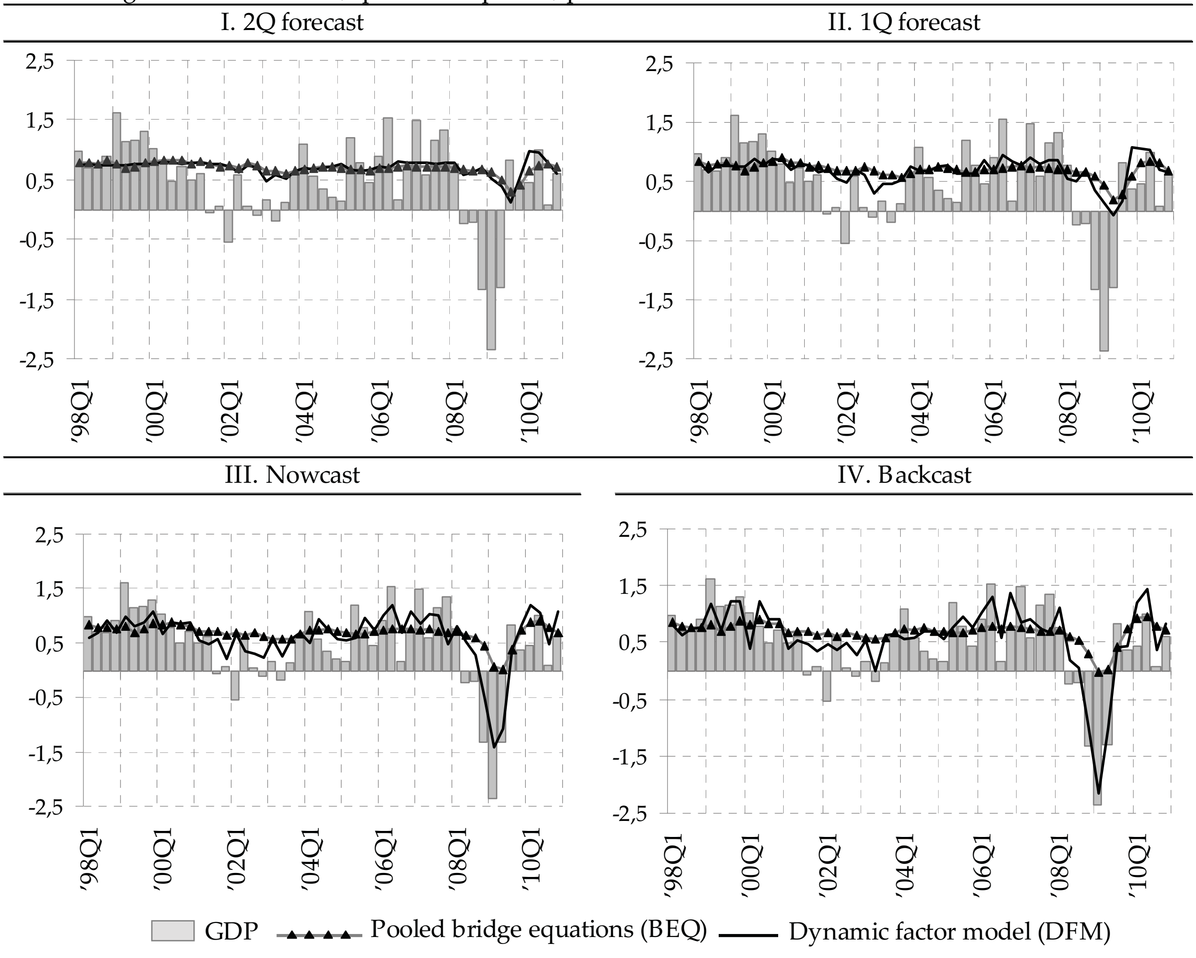
5. Outcome

This section presents the outcome of the "horse race" between the linear models (section 5.1 and 5.2) and the best linear model versus Consensus Forecasts (section 5.3). It concludes with a digression on the dynamic factor model (section 5.4).

5.1 A first look at the linear model forecasts

Figure 3.I to 3.IV depict the forecasting performance of the dynamic factor model and the pooled bridge equations, for different points in time. The top-left panel (3.I) shows the two-quarter-ahead forecast. Moving from left to right and down the forecasting horizon is getting progressively shorter. The horizontal axis shows the forecasted quarter and the vertical axis. shows the GDP forecast and realisation.

Figure 3 GDP forecasts dynamic factor model and pooled bridge equations
Real GDP growth Netherlands, quarter on quarter, percent



Notes: month 3 for nowcasts, 1Q forecast and 2Q forecast and month 2 for backcast.
Source: own calculations

As can be seen the two-quarter-ahead forecast of the pooled bridge equations as well as the dynamic factor model missed the abrupt growth slowdown at the end of 2008. As the forecast horizon shortens (moving from pane 3.I to 3.IV), the forecast performance of the models steadily increases.

Two points stand out. First, starting from the nowcast, the dynamic factor model clearly indicates an abrupt slowdown at the end of 2008, whilst the backcast was very close to the realisation. This stands in sharp contrast to the performance of the pooled bridge equations, that grossly underestimated the depth of the crisis. The same holds true for the other statistical models (Figure A.1 to A.3 in appendix A). Second, the forecasts of the pooled bridge equations shows very little dynamics in the projection period: most forecasts fluctuate only slightly around trend growth (about 0.5 percent quarter on quarter).

Overall, the forecasts of the considered pooled approaches and the naïve benchmarks were only revised down after the very negative GDP growth figures for 2008Q4 were released. In contrast, the dynamic factor model growth forecasts were much quicker in acknowledging the quick deterioration in the economy. This signal became more accurate when more "hard" information was released, like industrial production and retail sales. The graphical analysis indicates the dynamic factor model is more efficient in translating the flow of monthly information to an accurate GDP growth forecast. The next section presents the outcome of formal statistical tests to verify these outcomes.

5.2 The best linear model

Table 1 shows the RMSFEs over the whole sample period, including the financial crisis and the dot-com recession. Shaded areas indicate the dynamic factor model has a higher forecast accuracy *and* the difference is statistically significant according to the Diebold Mariano test. Panels A, B and C show the outcome for respectively the whole sample period, the period excluding the financial crisis and the period excluding both the financial crisis and dot-com recession.

The main message from Table 1 is that the dynamic factor is almost univocally the best forecast model across forecast horizons when considering the whole sample. The dynamic factor model reduces the RMSFE against a naïve recursive mean model (RM in Table 1) by as much as 40 percent (when backcasting). The relative performance of the dynamic factor model is best for now- and backcasting. This outcome is in line with the graphical analysis in section 5.1. The relative performance of the dynamic factor models worsens when the forecast horizon increases. Consequently, the difference in forecast accuracy of the dynamic factor model and the other models is negligible for the two-quarters-ahead forecast, though still significant against some of the models when the financial crisis is included.

Table 1. Forecast accuracy - Mechanical linear models[1]

		RM	AR	BEQ	VAR	DFM
I. Whole sample (1995Q1 -2010Q4)						
2Q Forecast	month 1	0.73	0.73	0.72	0.73 *	0.72
	month 2	0.73 *	0.73 *	0.71	0.73 *	0.71
	month 3	0.73 *	0.73 *	0.71	0.72 *	0.70
1Q Forecast	month 1	0.73 *	0.73 *	0.70	0.72 *	0.69
	month 2	0.73 *	0.73 *	0.69	0.72 *	0.67
	month 3	0.72 *	0.72 *	0.68 **	0.71 **	0.62
Nowcast	month 1	0.72 *	0.72 *	0.67 *	0.71 *	0.59
	month 2	0.72 *	0.72 *	0.66 *	0.71 *	0.53
	month 3	0.72 *	0.71 *	0.63 *	0.62 *	0.48
Backcast	month 1	0.72 *	0.71 **	0.62 *	0.62 **	0.41
	month 2	0.72 *	0.71 **	0.61 **	0.62 **	0.39
I. Sample without financial crisis (1995Q1 -2007Q4						
2Q Forecast	month 1	0.52	0.52	0.52	0.53 *	0.52
	month 2	0.52	0.52	0.52	0.53 *	0.51
	month 3	0.52	0.52	0.51	0.52	0.51
1Q Forecast	month 1	0.52	0.52	0.51	0.52	0.50
	month 2	0.52	0.52	0.50	0.52	0.49
	month 3	0.52 *	0.52 *	0.50 *	0.52 **	0.46
Nowcast	month 1	0.52	0.52	0.50	0.52	0.47
	month 2	0.52	0.52	0.49	0.52 *	0.46
	month 3	0.52 *	0.54 *	0.49	0.50 **	0.45
Backcast	month 1	0.52 ***	0.54 ***	0.48 ***	0.50 ***	0.41
	month 2	0.52 ***	0.54 ***	0.48 ***	0.50 ***	0.39
III. Sample without financial crisis and dot-com recession (1995Q1:2000Q4 and 2003Q4:2007Q4)						
2Q Forecast	month 1	0.45	0.44	0.45	0.46	0.45
	month 2	0.45	0.44	0.45	0.46	0.45
	month 3	0.45	0.44	0.45	0.45	0.46
1Q Forecast	month 1	0.45	0.44	0.45	0.45	0.46
	month 2	0.45	0.44	0.45	0.45	0.46
	month 3	0.44	0.44	0.45	0.46	0.44
Nowcast	month 1	0.44	0.44	0.45	0.46	0.45
	month 2	0.44	0.44	0.45	0.46	0.45
	month 3	0.44	0.45	0.44	0.46	0.44
Backcast	month 1	0.44 *	0.45 **	0.44	0.46 *	0.40
	month 2	0.44 **	0.45 ***	0.43 **	0.46 **	0.38

Note: Figures in bold italics indicate smallest RMSFE. *, **, *** indicate the significance level (10%, 5%, 1%) from a one sided Diebold-Mariano test (Diebold and Mariano, 1995). The null hypothesis is that the difference in forecast accuracy between the dynamic factor model and the alternative model is zero.

[1] RM: recursive mean, AR: autoregressive model, BEQ: weighted average quarterly bridge equations, VAR: weighted average quarterly vector autoregressions, DFM: dynamic factor model.

A striking result concerns the negligible deterioration in the accuracy of nowcasts and backcasts of the dynamic factor model during the financial crisis. This implies that the dynamic factor model is relatively robust to sharp changes in the variables, while the other models clearly had difficulty translating the sharp decline in indicators to a sharp decline in

GDP growth. This might be caused by the relative strong backward-looking nature of these models. In technical terms, the autoregressive component is bigger for these models, leading to a relatively slow reaction to abrupt changes. This outcome reinforces the outcome of the graphical analysis, which showed that the dynamic factor model was reacting relatively quickly to new information, translating into a relatively sharp drop in the GDP growth forecast at the onset of the financial crisis.

A third result that stands out is that the dynamic factor model is especially advantageous in periods of abrupt changes, such as the financial crisis. Excluding the financial crisis, the forecast accuracy of the dynamic factor is only superior to the other models in the very short-term, when backcasting. This is in line with previous research (e.g. Rünstler et al., 2009, Camacho and Perez-Quiros, 2010 and Bańbura et al., 2011) that concludes that the dynamic factor model only outperforms at very short forecast horizons, and in particular when nowcasting and backcasting.

The fact that the dynamic factor model outperforms the other models does not necessarily mean the other forecast do not contain any additional information. To test this we conducted encompassing test. The outcome is shown in Table 2. Shaded areas indicate the other models contain no extra information with regards to the dynamic factor model. In other words, the most accurate forecasts are made by using the dynamic factor model in isolation. The main message from Table 2 is that there is no gain in forecast accuracy from combining the dynamic factor model with any of the other linear models when (severe) recessions are included in the sample. When excluding the financial crisis, there is no gain from any of the linear models over a simple forecast equalling the mean growth rate in the previous quarters[7]. When excluding the financial crisis as well as the dot-com crisis (panel III in table 2) this is also the case for the one quarter ahead forecast and the nowcast. Even when all recessions are excluded from the sample, the dynamic factor still reduces the forecast error when backcasting.

In sum, the main insight from table 1 and 2 is that the relative forecast accuracy of the dynamic factor model increases at turning points in the business cycle, especially when the financial crisis is included.

7 In mathematical terms: if β_0^z and β_0^{DFM} in the encompassing test ($y_{t+h}^q = \alpha_0 + \beta_0^z \hat{y}_{t+h,t}^q + \beta_0^{DFM} \hat{y}_{t+h,t}^q + \varepsilon_0$) are zero, than α_0, which roughly equals the average GDP growth in the previous quarters, is driving the forecast.

Table 2. Forecast combination - Encompassing test[1]

		RM	AR	BEQ	VAR
		I. Whole sample (1995Q1 -2010Q4)			
2Q Forecast	month 1	1.27	1.26	1.13	1.28
	month 2	1.54 **	1.53 **	1.29	1.52 **
	month 3	1.61 **	1.61 **	1.32	1.61 **
1Q Forecast	month 1	1.58 ***	1.59 ***	1.09	1.57 ***
	month 2	1.67 ***	1.69 ***	1.15 *	1.68 ***
	month 3	2.13 ***	2.12 ***	1.93 ***	2.18 ***
Nowcast	month 1	1.60 ***	1.59 ***	1.27 ***	1.61 ***
	month 2	1.25 ***	1.24 ***	1.02 ***	1.24 ***
	month 3	1.19 ***	1.23 ***	1.03 ***	1.22 ***
Backcast	month 1	1.06 ***	1.08 ***	1.08 ***	1.08 ***
	month 2	1.04 ***	1.06 ***	1.00 ***	1.03 ***
		I. Sample without financial crisis (1995Q1 -2007Q4			
2Q Forecast	month 1	0.46	0.47	0.41	0.48
	month 2	0.59	0.59	0.57	0.60
	month 3	0.97	0.95	0.80	0.95
1Q Forecast	month 1	1.16 *	1.14	1.00	1.19 *
	month 2	1.35 **	1.33 **	1.29 *	1.41 **
	month 3	1.98 ***	2.01 ***	1.92 ***	2.04 ***
Nowcast	month 1	1.17 ***	1.19 ***	1.02 **	1.20 ***
	month 2	0.91 ***	0.93 ***	0.77 **	0.93 ***
	month 3	0.99 ***	0.90 ***	0.72 **	0.92 ***
Backcast	month 1	1.06 ***	1.01 ***	0.91 ***	1.03 ***
	month 2	1.01 ***	0.97 ***	0.85 ***	0.98 ***
		III. Sample without financial crisis and dot-com recession (1995Q1:2000Q4 and 2003Q4:2007Q4)			
2Q Forecast	month 1	-0.18	-0.18	-0.19	-0.18
	month 2	-0.25	-0.25	-0.25	-0.24
	month 3	-0.59	-0.57	-0.58	-0.57
1Q Forecast	month 1	-0.55	-0.53	-0.63	-0.57
	month 2	-0.45	-0.44	-0.50	-0.50
	month 3	0.36	0.37	0.38	0.35
Nowcast	month 1	-0.17	-0.16	-0.13	-0.16
	month 2	0.24	0.24	0.24	0.24
	month 3	0.35	0.37	0.31	0.37
Backcast	month 1	0.65 **	0.67 **	0.66 **	0.68 **
	month 2	0.71 ***	0.71 ***	0.72 ***	0.72 ***

Note: *, **, *** indicate significance level (10%, 5%, 1%) tests if β_0^{DFM} is significantly different from zero (two-sided t-test). Shaded areas indicate β_0^{Z} is not statistically different from zero and β_0^{DFM} is statistically different from zero.

[1] RM: recursive mean, AR: autoregressive model, BEQ: weighted average quarterly bridge equations, VAR: weighted average quarterly vector autoregressions.

5.3 The best linear model versus Consensus Forecasts

This section evaluates the best linear model, i.e. the dynamic factor model, against the mean quarterly Consensus Forecasts. Because the quarterly Consensus Forecast are published only once a quarter, the comparison is more limited than the comparison with the mechanical linear models. More precisely, the Consensus Forecasts are released in the second week of the final month of each quarter, based on a survey of panellist's forecasts in the previous two weeks (Batchelor, 2011)[8]. Given the release scheme of Statistics Netherlands, this means that at the time the forecasts are made, panellists have information on GDP growth in the preceding quarter. It is therefore not possible to make a comparison of the forecast accuracy of the backcasts.

Table 3 compares the forecast accuracy of Consensus Forecasts with the dynamic factor model. The main message from this table is that the dynamic factor model beats the Consensus for the nowcasts, and one and two quarter forecasting horizons. Including the crisis, this advantage reduces to the nowcasts.

Table 3. Forecast accuracy - Dynamic Factor Models versus Consensus Forecasts[1]

Root Mean Squared Forecast Error (RMSFE)

	I. Whole sample (1995Q1 -2010Q4)	
	Dynamic Factor Model	Consensus Forecasts
2Q Forecast	0.68	0.67
1Q Forecast	0.60	0.62
Nowcast	0.45 **	0.56
	II. Sample without financial crisis (1995Q1 -2007Q4)	
2Q Forecast	0.40 **	0.52
1Q Forecast	0.40 ***	0.52
Nowcast	0.41 **	0.49
	III. Sample without financial crisis and dot-com recession (1995Q1:2000Q4 and 2003Q4:2007Q4)	
2Q Forecast	0.41 **	0.50
1Q Forecast	0.39 **	0.50
Nowcast	0.40 **	0.52

Note: Forecasts on the third month of a quarter. Figures in bold italics indicate smallest RMSFE. *, **, *** indicate the significance level (10%, 5%, 1%) from a one sided Diebold-Mariano test (Diebold and Mariano, 1995). The null hypothesis is that the difference in forecast accuracy between the dynamic factor model and Consensus Forecasts is zero.

Table 4 presents the outcome of the encompassing test of the dynamic factor model against Consensus Economics. The table indicates that combining the Consensus Forecasts with the dynamic factor model could have slightly increased the forecast accuracy during the financial crisis. The coefficient of Consensus Forecasts is rather low compared to the dynamic factor model forecast, indicating the increase in forecast accuracy from combining the two

8 This timing slightly differs from the data availability for the dynamic factor model, because this is based on the information set available on the 10th of a given month. In general, panellists surveyed in the first week will have slightly less information, whilst panellists surveyed in the second month will have slightly more information. We assumed these information (dis)advantages roughly cancelled each other out.

approaches is rather limited. Excluding the financial crisis there was no value added from combining the two.

Table 4. Encompassing test - Dynamic Factor Model versus Consensus Forecasts[1]

	β_0^{DFM}	$\beta_0^{Consensus}$
I. Whole sample (1995Q1 -2010Q4)		
2Q Forecast	1.22	0.55 **
1Q Forecast	1.60 ***	0.38 *
Nowcast	1.11 ***	0.33 *
II. Sample without financial crisis (1995Q1 -2007Q4)		
2Q Forecast	1.55 **	0.04
1Q Forecast	1.88 ***	0.08
Nowcast	0.80 **	0.26
III. Sample without financial crisis and dot-com recession (1995Q1:2000Q4 and 2003Q4:2007Q4)		
2Q Forecast	0.39	0.07
1Q Forecast	0.97	0.27
Nowcast	1.01	0.21

Note: *, **, *** indicate significance level (10%, 5%, 1%) tests if β_0^{DFM} is significantly different from zero (two-sided t-test). Shaded areas indicate $\beta_0^{Consensus}$ is not statistically different from zero and β_0^{DFM} is statistically different from zero.

We conclude the dynamic factor model always has higher forecast accuracy than the forecasts of Consensus Economics when nowcasting. This outcome is roughly in line with the outcome of the mechanical linear models, that indicates the dynamic factor model is especially well equipped to now- and backcast. Overall, the gain from combining the dynamic factor model forecast with Consensus Forecasts is rather limited. However, some caution is warranted because we were not able to fully replicate the information set available to the professional forecasters at the time they made their forecast, although the differences are presumably rather small (see section 4.2).

5.4 A short digression on the dynamic factor model

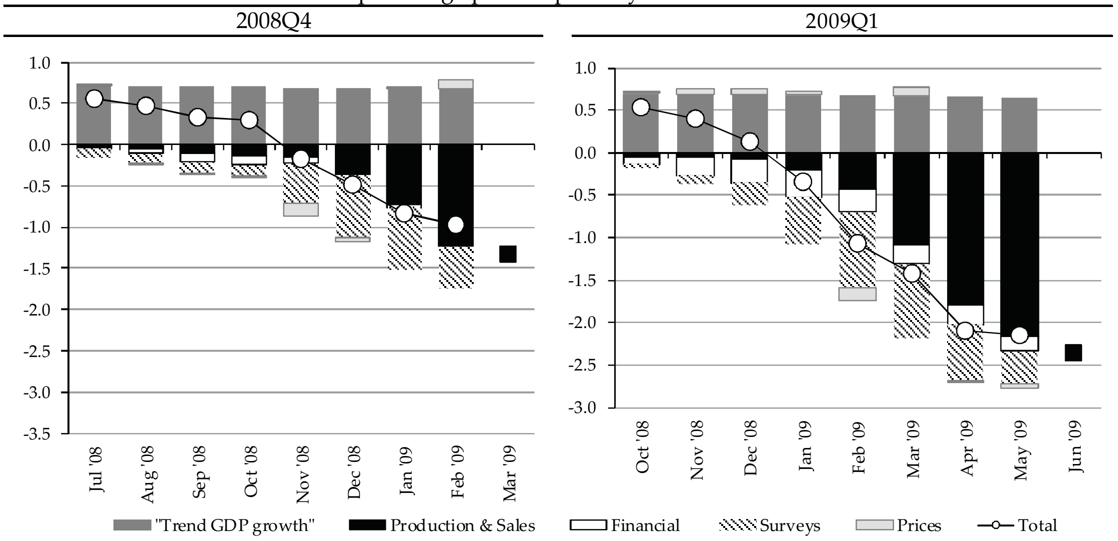
In order to better understand the relatively good forecasting performance of the dynamic factor model, Figure 4 shows a decomposition of the dynamic factor model forecasts[9] for the two most negative growth quarters during the financial crisis (2008Q4 and 2009Q1). The forecast is decomposed into four data blocks, i.e.: production and sales, financial data, surveys and prices. The left hand panel shows the evolution of the forecast for 2008Q4,

[9] The derivation of the contributions is based on an algorithm by Koopman and Harvey (2003). Bańbura and Rünstler (2011) applied the algorithm to a dynamic factor model. We refer to the latter paper for a full description of the algorithm.

starting in July 2008 (the first month of the one quarter ahead forecast). The right hand panel shows the evolution of the dynamic factor model forecast for 2009Q1, starting in October 2009.

Figure 4. Contributions to the GDP forecast for the Netherlands, by group *)

Contribution to the GDP forecast in percentage-point respectively total



*) White circles indicate total growth projection. The release of GDP is indicated by a black square.
Source: own calculations.

The evolving forecast for 2008Q4 figure the dynamic factor model inched down when the first signals on the deterioration in the economy were coming in through financial market s and survey data, during 2008Q2. The signal became progressively clearer when more information measuring the “hard” economic indicators (such as industrial production and retail sales) was released. This is in line with previous research (Giannone et al., 2008 and Bańbura and Rünstler, 2011) that indicates the contribution of surveys and financial data (i.e. the “soft” data) have the highest contributions for the one quarter ahead forecasts and nowcasts. As soon as “hard” information (e.g. exports and industrial production) for the projected quarter is available, the “soft” data lose their forecasting edge over “hard” data.

Figures A.4 to A.6 in the appendix give an overview of the contribution of individual indicators for the —one and two quarter ahead— forecast as well as the now- and backcast for 2009Q1. When backcasting (Figure A.4), the relatively high weight of world trade and industrial production in Germany, Italy and France indicates the strong drop in GDP growth was strongly driven by economic developments on Dutch export markets. When nowcasting, the importance of confidence indicators increases. The one and two quarter forecasts for 2009Q1 clearly display an increasing weight for indicators from financial markets, especially stock markets. These results indicate the importance of "soft" information from financial markets and surveys for early detection of abrupt changes in GDP growth.

6. Conclusions

The financial crisis provides an excellent opportunity to "stress-test" linear forecasting models. Our empirical analysis of the Dutch experience in the period 1995-2010 finds that the dynamic factor model is the only model capable to nowcast the depth and timing of the crisis. It is the only model that offers a reasonably accurate picture of the current state of the economy when it goes through a volatile phase. Moreover, the dynamic factor model generates better forecasts than the judgemental Consensus Forecasts in both crisis and non-crisis times, while it produces better nowcasts during the financial crisis. This finding is in line with other research indicating professional forecasters are sluggish to incorporate new information in their projections.

Overall, our findings suggest that a suitably designed mechanical dynamic factor model is more efficient at processing new information than the other linear statistical models considered as well as professional forecasters, especially during the financial crisis. Our result is quite relevant to policy makers and private economic agents, as it suggests that it is optimal to generate short-term GDP growth forecasts by purely statistical methods. Augmenting the outcomes of the statistical procedure by judgement has hardly any added value.

There are several avenues for further research on this topic. First of all, it is interesting to broaden the analysis to other countries, as data on quarterly Consensus Forecasts are available for thirteen countries. Secondly, in light of the heightened interest in constructing real-time databases in recent research (e.g. Giannone et al., 2010), the empirical analysis could be further refined by using truly real-time data for estimation and forecasting. The current setup takes into account publication delays and revisions in GDP, but was unable to include possible revisions to the data on prices, production and sales. In this regard, the comparison between the Consensus Forecasts and the dynamic factor model should be interpreted with some caution. A third way forward is to enrich the set of statistical models by optimizing the weighting schemes that pool the individual bridge and VAR equations. Drechsel and Scheufele (2011) found that using clever combinations can improve the forecast for industrial production growth, especially during the financial crisis. Finally, the role of judgement in the Consensus Forecasts could be further explored, by investigating how panellists incorporate judgemental elements into their forecasts. Insights from recent case studies on the role of judgement when forecasting firm level sales data (e.g. Franses and Lagerstee, 2010 and Fildes and Goodwin, 2007) can be helpful in this respect.

References

Ager, P., Kappler, M. and S. Osterloh, 2009, The accuracy and efficiency of the Consensus Forecasts: a further application and extension of the pooled approach, *International Journal of Forecasting*, 25, 167-181.

Andreou, A., Ghyssels, E. and A. Kourtellos, 2011, Forecasting with mixed-frequency data In: Clements, M. and D. Hendry, *The Oxford Handbook of Economic Forecasting*, Oxford: Oxford University Press.

Baffigi, A., Golinelli, R. and G. Parigi, 2004, Bridge models to forecast the euro area GDP, *International Journal of Forecasting*, 20, 447-460.

Bańbura, M. and G. Rünstler, 2011, A look into the factor model black box: publication lags and the role of hard and soft in forecasting GDP, *International Journal of Forecasting*, 27, 333-346.

Bańbura, M., Giannone, D. and L. Reichlin, 2011, Nowcasting, In: Clements, M. and D. Hendry, *The Oxford Handbook of Economic Forecasting*, Oxford: Oxford University Press.

Batchelor, R., 2001, How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus, *Applied Economics*, 33, 225-235.

Bernanke, B.S. and J. Boivin, 2003, Monetary policy in a data-rich environment, *Journal of Monetary Economics*, 50, 525-546.

Camacho, M. and G. Perez-Quiros, 2010, Introducing the EURO-STING: Short Term INdicator of Euro Area Growth, *Journal of Applied Econometrics*, 25, 663-694.

Camba-Mendez, G., Kapetianos, G., Smith, M. and R. Weale, 2001, An automatic leading indicator of economic activity: forecasting GDP growth for European countries, *Econometrics Journal*, 4, 56-80.

Diebold, F.X. and R.S. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Doz, C., Giannone, D. and L. Reichlin, 2011, A two-step estimator for large approximate dynamic factor models based on Kalman filtering, *Journal of Econometrics*, 164, 188-205.

Drechsel, K. and R. Scheufele, 2011, The performance of short-term forecasts of the German economy before and during the 2008-2009 recession, *International Journal of Forecasting,* forthcoming.

ECB, 2009, *Results of a special questionnaire for participants in the ECB survey of professional forecasters*, Frankfurt: European Central Bank.

Fair, R.C. and R.J. Shiller, 1990, Comparing information in forecasts from econometric models, *The American Economic Review*, 80, 375-389.

Fildes, R. and P. Goodwin, 2007, Good and bad judgement in forecasting: lessons from four companies. *Foresight: The International Journal of Applied Forecasting*, 8, 5-10.

Franses, P.H. and R. Lagerstee, 2010, Do experts' adjustments on model-based SKU-level forecast improve forecast quality? *Journal of Forecasting*, 29, 331-340.

Garratt, A., Lee, K. and S. Vahey, 2008, Real-time probability forecasts of UK macroeconomic events, *National Institute Economic Review*, 203, 78-90

Giannone, D. Henry, J., Lalik, M. and M. Modugno, 2010, *An area-wide real-time database for the euro area*, ECB Working Paper 1145.

Giannone, D., Reichlin, L. and D. Small, 2008, Nowcasting: the real-time informational content of macroeconomic data, *Journal of Monetary Economics*, 55, 665-676.

IMF, 2011, *IMF performance in the run-up to the financial and economic crisis,* Washington: International Monetary Fund.

Isiklar, G., Lahiri, K. and P. Loungani, 2006, How quickly do forecasters incorporate news? evidence from cross-country surveys, *Journal of Applied Econometrics*, 21, 703-725.

Jong, de, J., Roscam Abbing, M. en J. Verbruggen, 2010, *Forecasting in times of crisis. CPB-forecasts during the great recession*, CPB Document 207.

Kitchen, J. and R. Monaco, 2003, Real-time forecasting in practice: the US Treasury staff's real-time GDP forecast system, *Business Economics*, 38, 10-28.

Koopman, S.J. and A. Harvey, 2003, Computing observation weights for signal extraction and filtering, *Journal of Economic Dynamics and Control*, 27, 1317-1333.

Kuzin, V., Marcellino, M. and C. Schumacher, 2011, MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the Euro Area, *International Journal of Forecasting*, 27, 529-542.

Lombardi, M.J. and P. Maier, 2011, *Forecasting economic growth in the euro area during the great moderation and the great recession*, ECB Working Paper 1379.

Loungani, P. 2001, How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth, *International Journal of Forecasting*, 17, 429-432.

Loungani, P., Stekler, H. and N. Tamirisa, 2011, Information rigidity in growth forecasts: some cross-country evidence, IMF Working Paper 125.

Matheson, T., 2011, *New indicators for tracking growth in real time*, IMF Working Papers 43.

Orphanides A. and S. van Norden, 2002, The unreliability of output-gap estimates in real time. *Review of Economics and Statistics*, 4, 569–583.

Reijer, den A.H.J., 2005, *Forecasting Dutch GDP using large scale factor models*, DNB Working Paper 28.

Reinhart, C.M. and K.S. Rogoff, 2008, Is the 2007 US sub-prime Financial crisis so different? An international historical comparison, *American Economic Review*, 98, 339-344.

Romer, C.D. and D.H. Romer, 2000, Federal Reserve information and the behaviour of interest rates, *The American Economic Review*, 90, 431-457.

Roodenburg, O. and A.H.J. den Reijer, 2006, Dutch GDP Revisions: are they predictable and where do they come from?, *Applied Economics Quarterly*, 52, 337-356.

Rünstler, G., Barhoumi, K., Benk, S. Cristadoro, R., den Reijer, A., Jakaitiene, A. , Jeloniek, P., Rua, A., Ruth, K. and C. van Nieuwenhuyze, 2009, Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise, *Journal of Forecasting*, 28, 595-611.

Timmerman, A., 2006, Forecast Combinations, In: Elliott, G., Granger, C.W.J. and A. Timmerman (eds), 2006, *Handbook of Economic Forecasting*, North Holland, Amsterdam, 135-196.

Appendix

Table A.1 Database description

nr.	description	type	transformation[1] log	diff	filter	lag[1]
1	Industrial production - industry (total)	Production & sales	1	1	3	2
2	Industrial production - manufacturing (total)	Production & sales	1	1	3	2
3	Industrial production - manufacture of wearing apparel	Production & sales	1	1	3	3
4	Industrial production - manufacture of motor vehicles and (semi) trailers	Production & sales	1	1	3	3
5	Industrial production - manufacture of other transport equipment	Production & sales	1	1	3	3
6	Industrial production - manufacture of basic metals and metal products	Production & sales	1	1	3	2
7	Industrial production - treatment and coating of metals, machines and electr. equipment	Production & sales	1	1	3	3
8	Industrial production - electricity, gas, steam, water and air conditioning	Production & sales	1	1	3	2
9	Industrial production - manufacture of textiles	Production & sales	1	1	3	3
10	Industrial production - printing and reproduction of recorded media	Production & sales	1	1	3	3
11	Industrial production - construction	Production & sales	1	1	3	3
12	Industrial production - manufacture of food products and beverages	Production & sales	1	1	3	2
13	Industrial production - capital goods industry	Production & sales	1	1	3	2
14	Industrial production - durable consumer goods industry	Production & sales	1	1	3	2
15	Industrial production - non-durable consumer goods industry	Production & sales	1	1	3	2
16	Industrial production - consumer goods industry	Production & sales	1	1	3	2
17	Industrial production - Belgium (total)	Production & sales	1	1	3	3
18	Industrial production - Germany (total)	Production & sales	1	1	3	2
19	Industrial production - Spain (total)	Production & sales	1	1	3	2
20	Industrial production - France (total)	Production & sales	1	1	3	2
21	Industrial production - United Kingdom (total)	Production & sales	1	1	3	2
22	Industrial production - Italy (total)	Production & sales	1	1	3	2
23	Car registration - new commercial vehicles	Production & sales	1	1	3	3
24	Car registration - new passenger car	Production & sales	1	1	3	2
25	Retail trade turnover (total)	Production & sales	1	1	3	2
26	Consumer expenditure by households	Production & sales	1	1	3	2
27	Unemployment	Production & sales	0	1	3	2
28	World Trade	Production & sales	1	1	3	2
29	Imports	Production & sales	1	1	3	2
30	Exports	Production & sales	1	1	3	2
31	M1	Financial	1	2	3	2
32	M3	Financial	1	2	3	2
33	Interest rate (short term)	Financial	0	1	3	1
34	Interest rate (long term)	Financial	0	1	3	1
35	Exchange rate, USD per EUR	Financial	0	1	3	1
36	Loans to the private sector	Financial	1	1	3	2
37	Loans on mortgage (nominal rate 5 < 10 years)	Financial	0	1	3	2
38	Share index, AEX	Financial	1	1	3	1
39	Share index, Amsterdam Midkap	Financial	1	1	3	1
40	Share index, Dow Jones Euro Stoxx 50	Financial	1	1	3	1
41	Share index, Dow Jones Euro Stoxx Industrials	Financial	1	1	3	1

[1] type: type of data (Production & Sales, Surveys, Prices, Financial and Other), log: 0=no logarithm, 1=logarithm, diff: degree of differencing 1=first difference, 2=second difference, filter: 3= change against the same month of the previous month.publ. lag: publication lag, e.g. publication lag equals 1 if the data are known for the previous month, 2 when the data are known for two months ago etc.

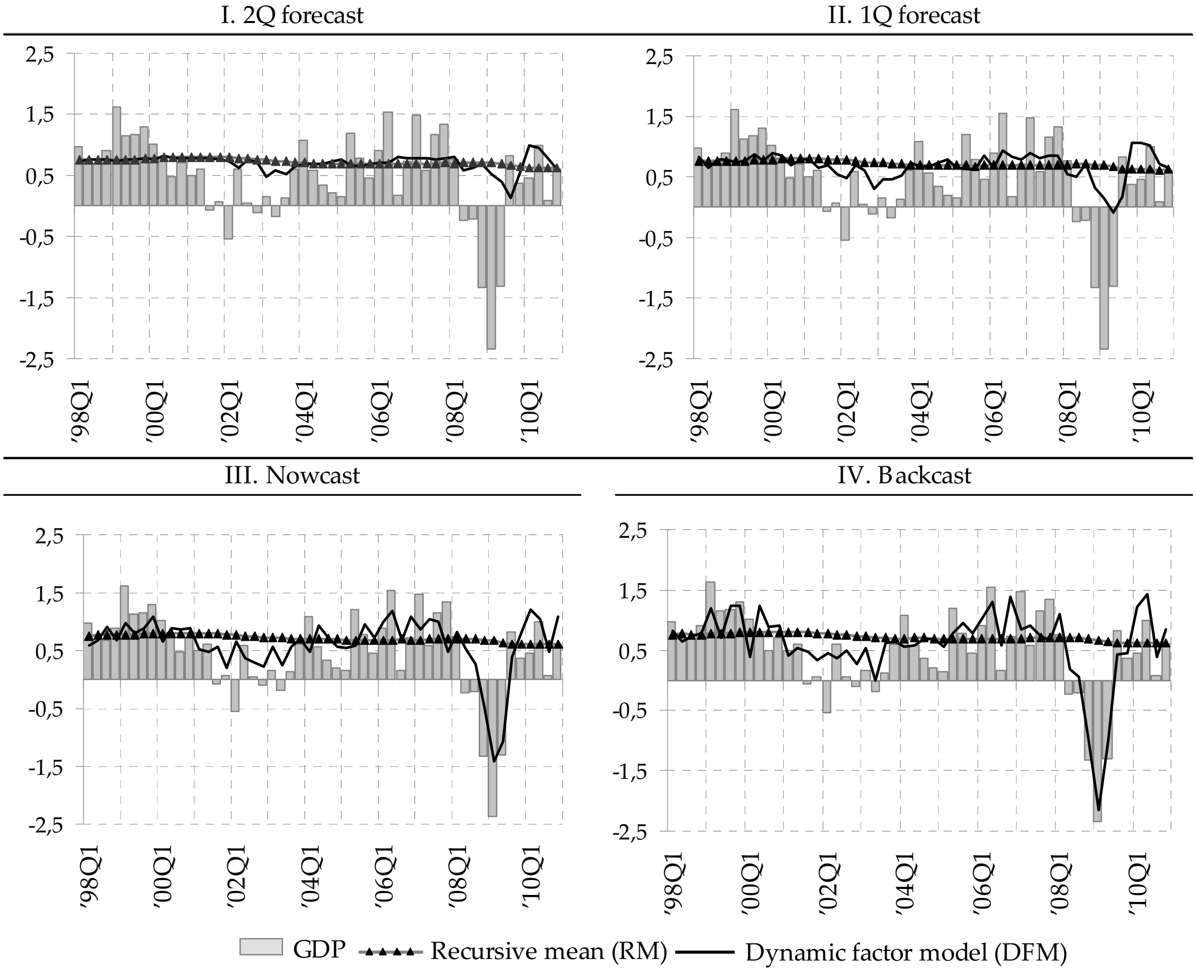
Table A.1 cont'd

nr.	description	type	transformation[1] log	diff	filter	lag[1]
42	Consumerprice index (total)	Prices	1	2	3	1
43	Consumerprice index - underyling	Prices	1	2	3	1
44	World market commodity prices, industrial materials	Prices	1	2	3	1
45	Producer prices, final products, domestic market	Prices	1	2	3	1
46	Producer prices, investment goods, domestic market	Prices	1	2	3	1
47	Producer prices, intermediate goods, domestic market	Prices	1	2	3	1
48	Producer prices, intermediate & final products, foreign market	Prices	1	2	3	1
49	Producer prices, intermediate & final products, domestic market	Prices	1	2	3	1
50	Terms of trade	Prices	1	2	3	2
51	Import prices	Prices	1	2	3	2
52	Export prices	Prices	1	2	3	2
53	Construction Confidence (headline)	Survey	0	1	3	1
54	Construction Confidence - building activity development over the past 3 months	Survey	0	1	3	1
55	Construction Confidence - evolution of your current overall order books	Survey	0	1	3	1
56	Construction Confidence - employment expectations over the next 3 months	Survey	0	1	3	1
57	Industrial Confidence (headline)	Survey	0	1	3	1
58	Industrial Confidence - production trend observed in recent months	Survey	0	1	3	1
59	Industrial Confidence - assessment of order-book levels	Survey	0	1	3	1
60	Industrial Confidence - assessment of export order-book levels	Survey	0	1	3	1
61	Industrial Confidence - assessment of stocks of finished products	Survey	0	1	3	1
62	Industrial Confidence - production expectations for the months ahead	Survey	0	1	3	1
63	Industrial Confidence - employment expectations for the months ahead	Survey	0	1	3	1
64	Consumer Confidence (headline)	Survey	0	1	3	1
65	Consumer Confidence - financial situation over last 12 months	Survey	0	1	3	1
66	Consumer Confidence - financial situation over next 12 months	Survey	0	1	3	1
67	Consumer Confidence - general economic situation over last 12 months	Survey	0	1	3	1
68	Consumer Confidence - general economic situation over next 12 months	Survey	0	1	3	1
69	Consumer Confidence - unemployment expectations over next 12 months	Survey	0	1	3	1
70	Consumer Confidence - major purchases at present	Survey	0	1	3	1
71	Consumer Confidence - major purchases over next 12 months	Survey	0	1	3	1
72	Consumer Confidence - savings at present	Survey	0	1	3	1
73	Consumer Confidence - savings over next 12 months	Survey	0	1	3	1
74	Consumer Confidence - statement on financial situation of household	Survey	0	1	3	1
75	IFO-indicator, expected business-situation	Survey	0	1	3	1
76	BNB-indicator, gross-index	Survey	0	1	3	1
77	Industrial confidence - Spain (headline)	Survey	0	1	3	1
78	Industrial confidence - France (headline)	Survey	0	1	3	1
79	Industrial confidence - Italy (headline)	Survey	0	1	3	1
80	Industrial confidence - United Kingdom (headline)	Survey	0	1	3	1

[1] type: type of data (Production & Sales, Surveys, Prices, Financial and Other), log: 0=no logarithm, 1=logarithm, diff: degree of differencing 1=first difference, 2=second difference, filter: 3= change against the same month of the previous month.publ. lag: publication lag, e.g. publication lag equals 1 if the data are known for the previous month, 2 when the data are known for two months ago etc.

Figure A.1. GDP forecasts dynamic factor model and recursive mean

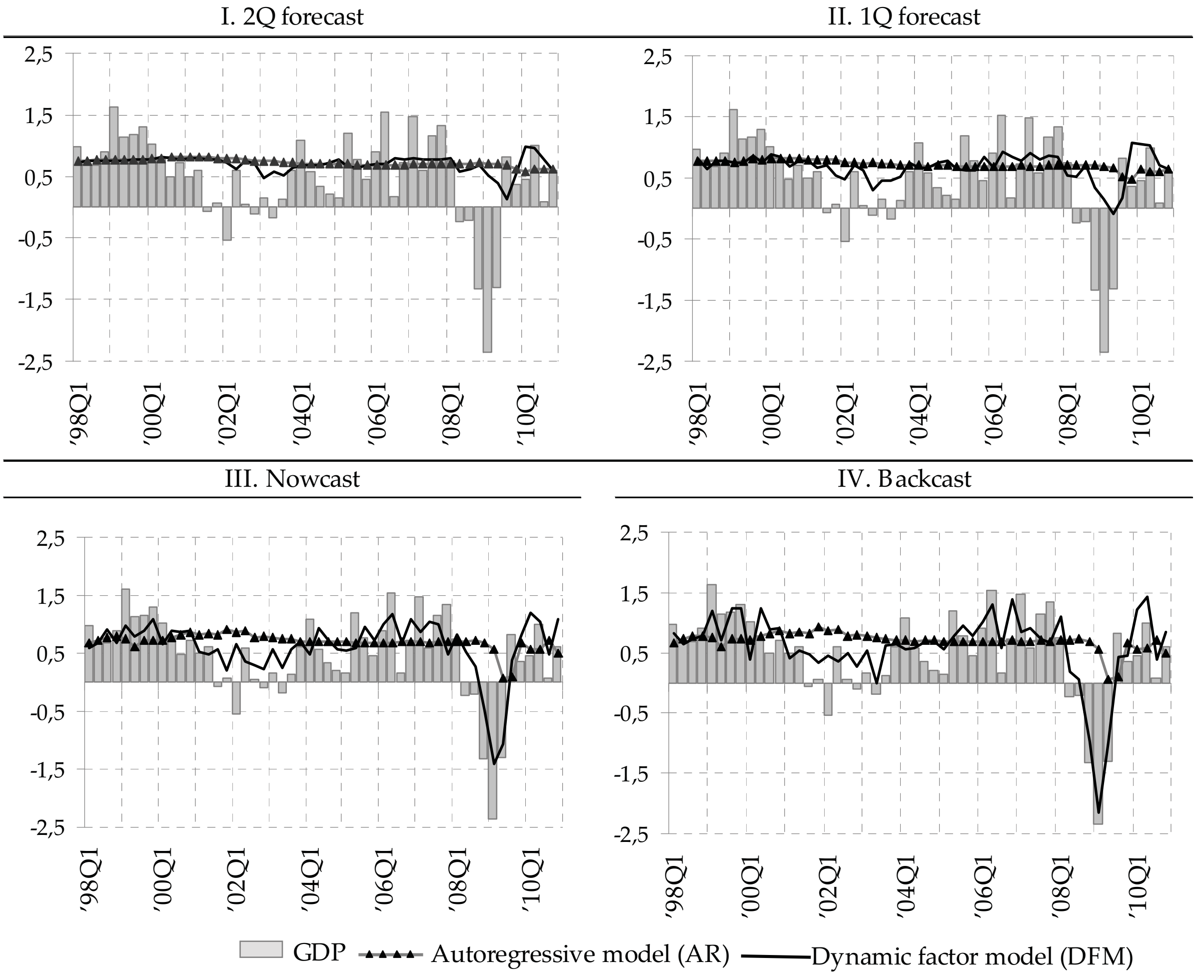
Real GDP growth Netherlands, quarter on quarter, percent



Notes: month 3 for nowcasts, 1Q forecast and 2Q forecast and month 2 for backcast.

Source: own calculations

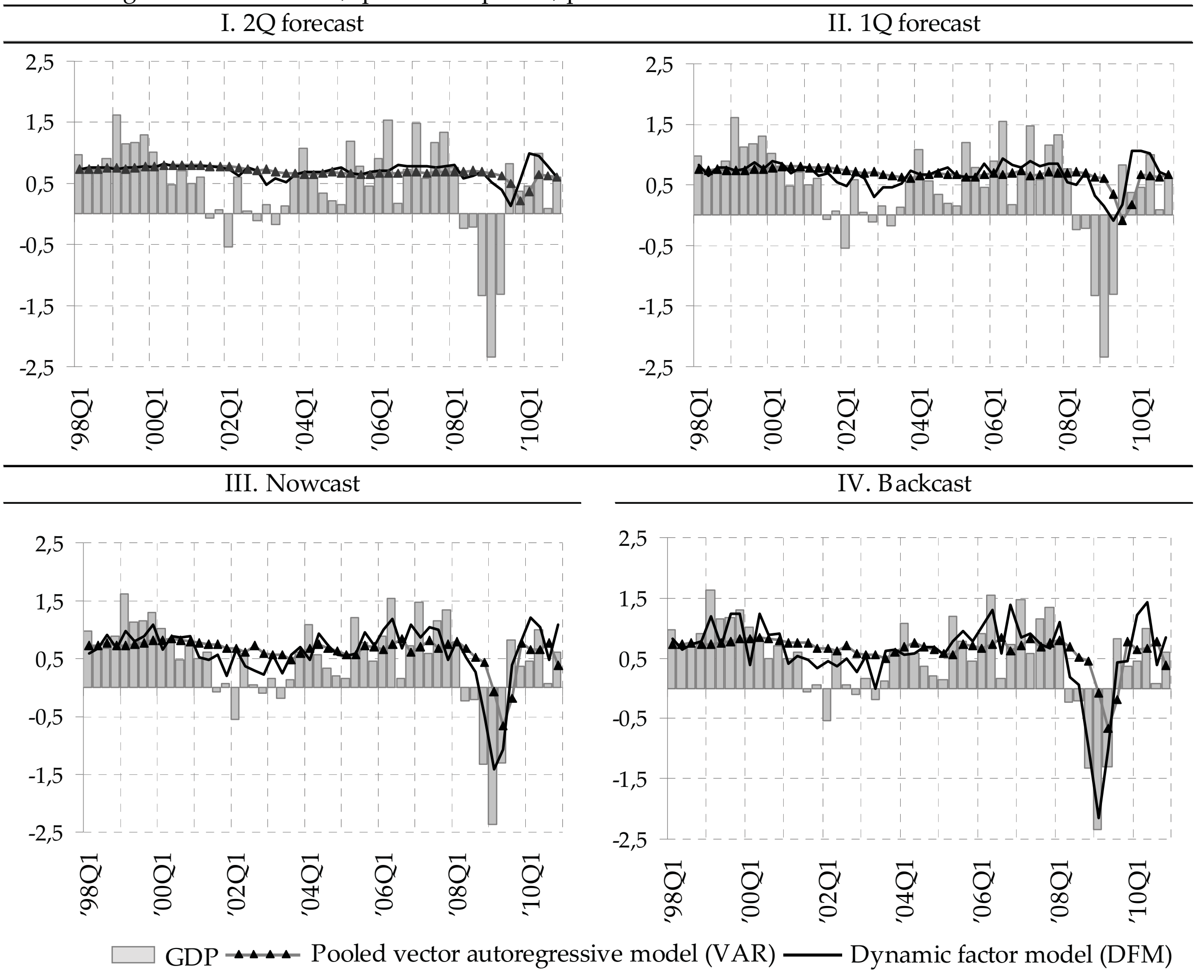
Figure A.2. GDP forecasts dynamic factor model and autoregressive model
Real GDP growth Netherlands, quarter on quarter, percent



Notes: month 3 for nowcasts, 1Q forecast and 2Q forecast and month 2 for backcast.

Source: own calculations.

Figure A.3 GDP forecasts dynamic factor model and pooled vector autoregressive models
Real GDP growth Netherlands, quarter on quarter, percent



Notes: month 3 for nowcasts, 1Q forecast and 2Q forecast and month 2 for backcast.
Source: own calculations

Figure A.4. Largest negative contributors to Backcast 2009Q1
May 2009

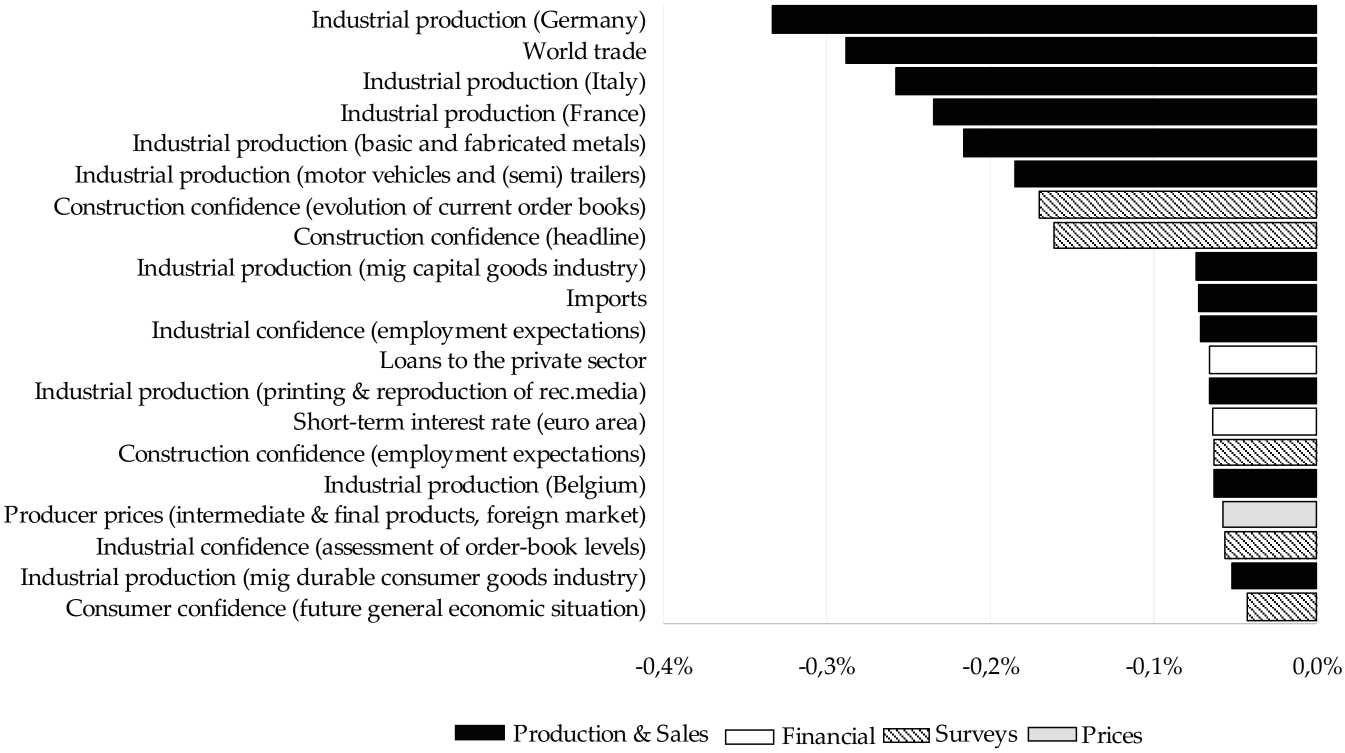


Figure A.5. Largest negative contributors to Nowcast 2009Q1
March 2009

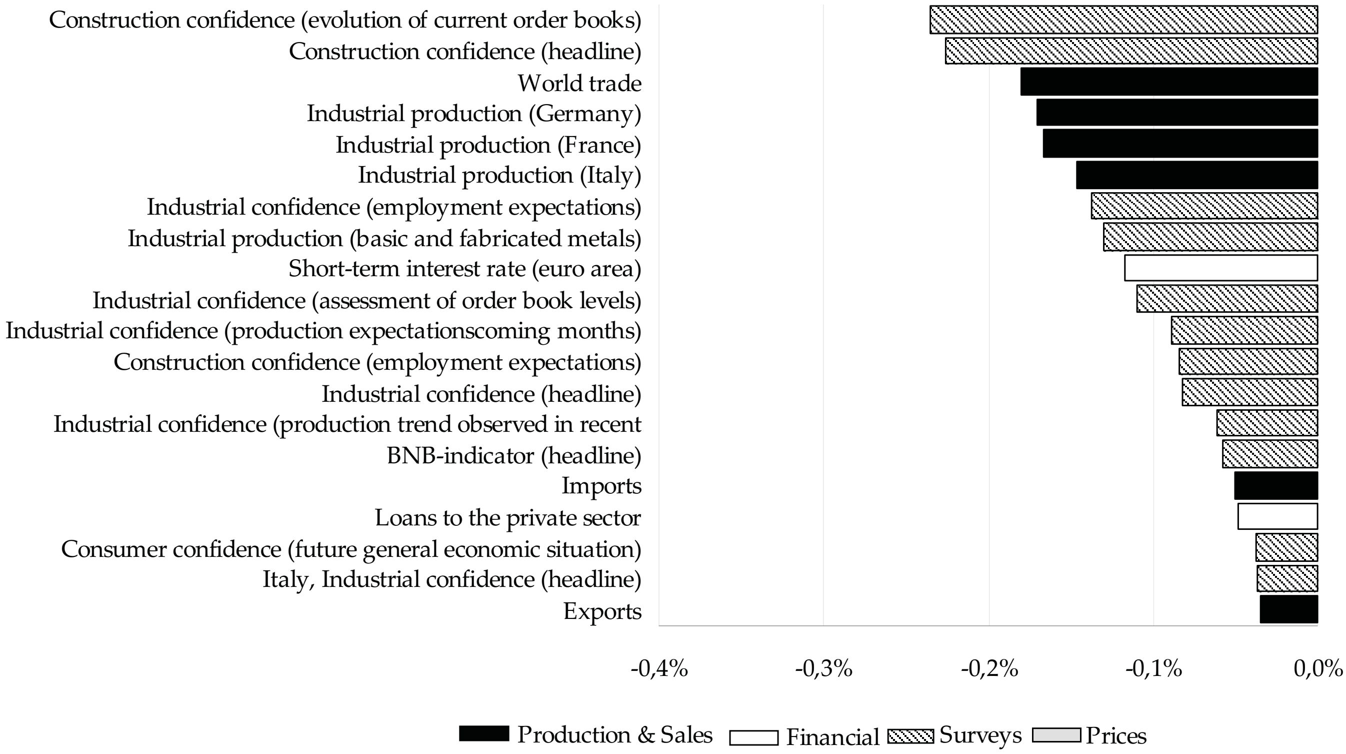


Figure A.6. Largest negative contributors to one quarter ahead Forecast 2009Q1
December 2008

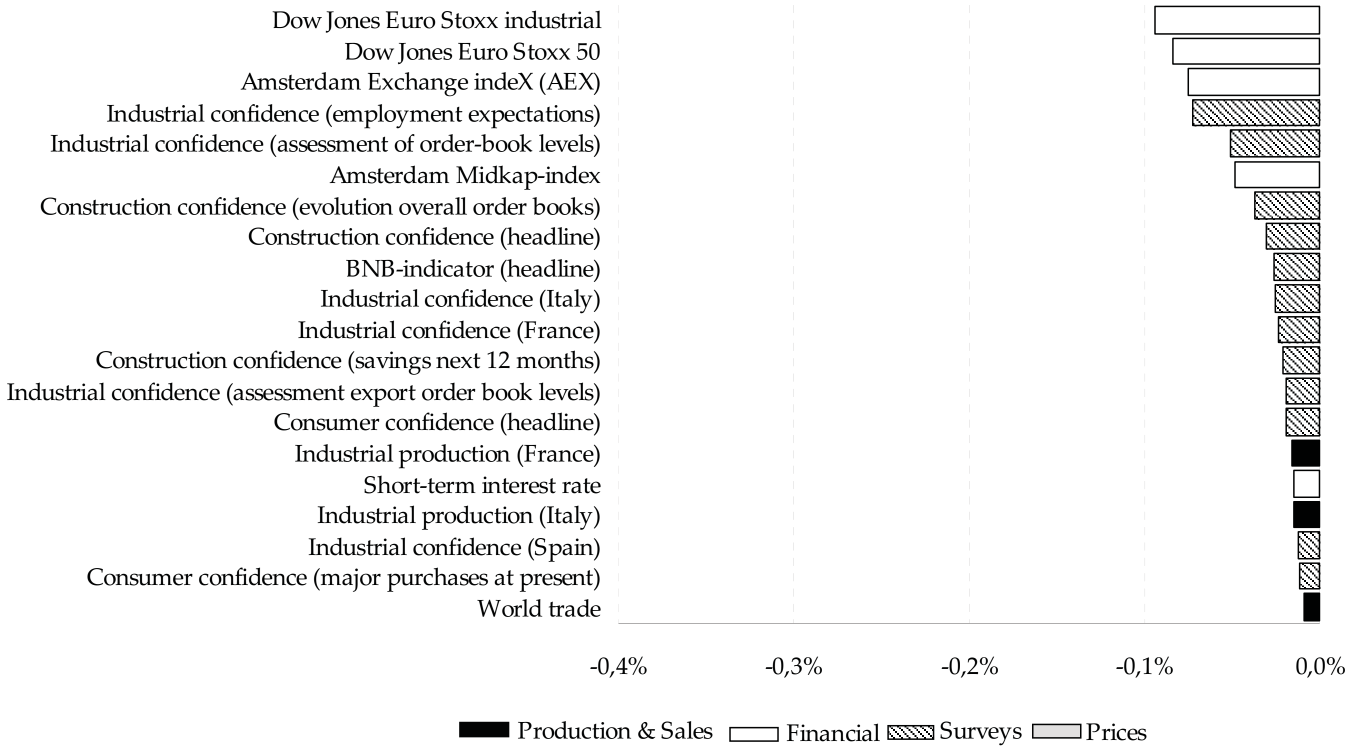
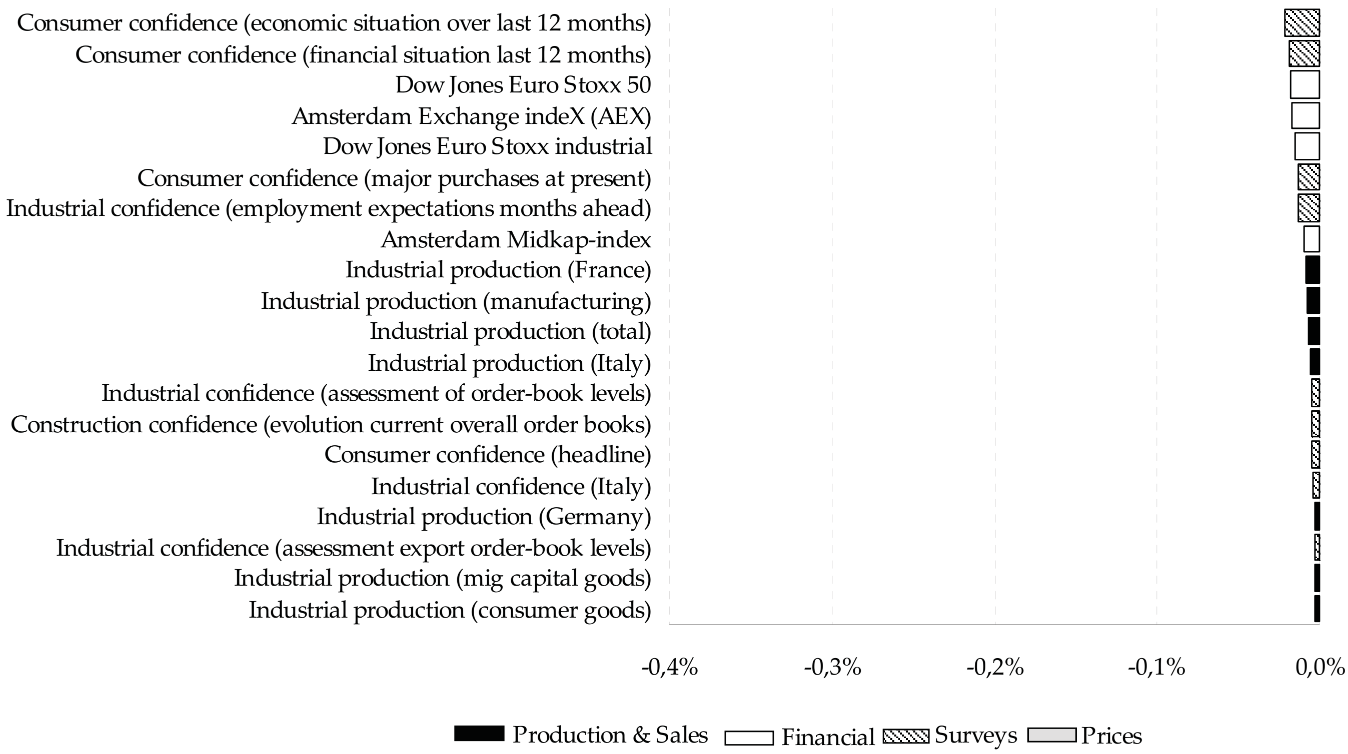


Figure A.7. Largest negative contributors to two quarter ahead Forecast 2009Q1
September 2008



Previous DNB Working Papers in 2011

No. 276 **Ronald Heijmans, Richard Heuver and Daniëlle Walraven**, Monitoring the unsecured interbank money market using TARGET2 data
No. 277 **Jakob Bosma**, Communicating Bailout Policy and Risk Taking in the Banking Industry
No. 278 **Jakob de Haan and Fabian Amtenbrink**, Credit Rating Agencies
No. 279 **Ralph de Haas and Neeltje van Horen**, Running for the Exit: International Banks and Crisis Transmission
No. 280 **I Kadek Dian Sutrisna Artha and Jakob de Haan**, Labor Market Flexibility and the Impact of the Financial Crisis
No. 281 **Maarten van Oordt and Chen Zhou**, Systematic risk under extremely adverse market conditions
No. 282 **Jakob de Haan and Tigran Poghosyan**, Bank Size, Market Concentration, and Bank Earnings Volatility in the US
No. 283 **Gabriele Galati, Peter Heemeijer and Richhild Moessner**, How do inflation expectations form? New insights from a high-frequency survey
No. 284 **Jan Willem van den End**, Statistical evidence on the mean reversion of interest rates
No. 285 **Marco Hoeberichts and Ad Stokman**, Price dispersion in Europe: Does the business cycle matter?
No. 286 **Cai Cai Du, Joan Muysken and Olaf Sleijpen**, Economy wide risk diversification in a three-pillar pension system
No. 287 **Theoharry Grammatikos and Robert Vermeulen**, Transmission of the Financial and Sovereign Debt Crises to the EMU: Stock Prices, CDS Spreads and Exchange Rates
No. 288 **Gabriele Galati, Federica Teppa and Rob Alessi,** Macro and micro drivers of house Price dynamics: An application to Dutch data
No. 289 **Rob Alessie, Maarten van Rooij and Annamaria Lusardi**, Financial Literacy, Retirement Preparation and Pension Expectations in the Netherlands.
No. 290 **Maria Demertzis**, Public versus Private Information
No. 291 **Enrico C. Perotti and Javier Suarez**, A Pigovian Approach to Liquidity Regulation
No. 292 **Jan Willem Slingenberg and Jakob de Haan**, Forecasting Financial Stress
No. 293 **Leo de Haan and Jan Willem van den End**, Banks' responses to funding liquidity shocks: lending adjustment, liquidity hoarding and fire sales
No. 294 **Janko Gorter and Jacob A. Bikker**, Investment risk taking by institutional investors
No. 295 **Eric J. Bartelsman, Pieter A. Gautier, and Joris de Wind**, Employment Protection, Technology Choice, and Worker Allocation
No. 296 **Maarten R.C. van Oordt and Chen Zhou**, The simple econometrics of tail dependence
No. 297 **Franka Liedorp, Robert Mosch, Carin van der Cruijsen and Jakob de Haan,** Transparency of banking supervisors
No. 298 **Tiziana Assenza, Peter Heemeijer, Cars Hommes and Domenica Massaro**, Individual Expectations and Aggregate Macro Behavior
No. 299 **Steven Poelhekke**, Home Bank Intermediation of Foreign Direct Investment
No. 300 **Nicole Jonker**, Card acceptance and surcharging: the role of costs and competition
No. 301 **Mark Mink**, Procyclical Bank Risk-Taking and the Lender of Last Resort
No. 302 **Federica Teppa**, Can the Longevity Risk Alleviate the Annuitization Puzzle? Empirical Evidence from Dutch Data
No. 303 **Itai Agur and Maria Demertzis**, "Leaning Against the Wind" and the Timing of Monetary Policy
No. 304 **Gabriele Galati, John Lewis, Steven Poelhekke and Chen Zhou**, Have market views on the sustainability of fiscal burdens influenced monetary authorities' credibility?
No. 305 **Itai Agur**, Bank Risk within and across Equilibria
No. 306 **Iman van Lelyveld and Marco Spaltro**, Coordinating Bank Failure Costs and Financial Stability
No. 307 **Enrico Perotti, Lev Ratnovski and Razvan Vlahu**, Capital Regulation and Tail Risk

Previous DNB Working Papers in 2011 (continued)

No. 308 **Ayako Saiki**, Exchange Rate Pass-Through and Monetary Integration in the Euro Area
No. 309 **Maria Demertzis and Nicola Viegi**, Interdependence of Fiscal Debts in EMU
No. 310 **Matej Marinč and Razvan Vlahu**, The Economic Perspective of Bank Bankruptcy Law
No. 311 **Jacob Bikker, Thijs Knaap and Ward Romp**, Real Pension Rights as a Control Mechanism for Pension Fund Solvency
No. 312 **Stefan Trautmann and Razvan Vlahu**, Strategic Loan Defaults and Coordination: An Experimental Analysis
No. 313 **Maarten van Rooij, Annamaria Lusardi and Rob Alessie**, Financial Literacy, Retirement Planning, and Household Wealth
No. 314 **Irina Stanga**, Sovereign and Bank Credit Risk during the Global Financial Crisis
No. 315 **Carin van der Cruijsen, Jakob de Haan, David-Jan Jansen and Robert Mosch**, Household savings behaviour in crisis times
No. 316 **Ronald Heijmans and Richard Heuver**, Is this bank ill? The Diagnosis of doctor TARGET2
No. 317 **Michel Beine, Elisabetta Lodigiani and Robert Vermeulen**, Remittances and Financial Openness
No. 318 **Joras Ferwerda, Mark Kattenberg, Han-Hsin Chang, Brigitte Unger, Loek Groot and Jacob Bikker**, Gravity Models of Trade-based Money Laundering
No. 319 **Jos Jansen and Ad Stokman, International Business Cycle Comovement**: Trade and Foreign Direct Investment