Models for short-term economic forecasts: an update

An accurate assessment of the current economic situation and how it will develop in the near term is crucial for monetary policymakers. The sooner changes in the economic situation and the resulting risks to price stability are identified, the sooner any need for monetary policy action can be determined. The Bundesbank regularly publishes its assessment of the economic outlook in Germany in its Monthly Report. The underlying short-term economic forecast also serves as a starting point for the semi-annual macroeconomic projections for Germany, which are incorporated into the macroeconomic projections for the euro area alongside the forecasts of other euro area central banks. The growth rate of gross domestic product (GDP) is the most important aggregate indicator in this regard.

Econometric forecast models serve as a key tool for short-term business cycle analysis. They can be used on an automated basis and draw on empirically observed relationships between a large number of leading economic indicators and the target variables to be forecast. Econometric models therefore provide a valuable basis for the ongoing assessment of economic activity.

Until now, the Bundesbank has been using three econometric models for short-term forecasts as part of its regular analysis of the German economy: a bridge equation model, a dynamic factor model and – owing to the particular importance of the manufacturing sector for the German economy – a model for industrial output, with each model consisting of different variants. Both the bridge equation model and the model for industrial output recently underwent a fundamental revision. The factor model, on the other hand, will remain in its present form for the time being as its structure is still fit for purpose. Furthermore, the existing set of forecasting instruments has been supplemented by a new vector autoregressive (VAR) model. An evaluation of the new and/or revised forecast models has shown that they deliver forecasts up to three quarters in advance, for which the information content is greater than that of a simple extrapolation using the historical average.

Short-term economic forecast models at the Bundesbank – scope of application and requirements

Short-term business cycle analysis crucial for monetary policy

The monetary policy decisions of the ECB Governing Council are based on a comprehensive assessment of macroeconomic and financial indicators, the aim being to identify risks to price stability and to identify any need for action. This analysis is divided into two pillars, the economic and the monetary analysis.[1] In the economic analysis, the assessment of the current economic situation and the outlook for the short and medium-term future play an important role because they can provide indications of increasing or decreasing price pressure. This is where GDP takes on an especially significant role as an aggregate indicator of economic activity. As for the single monetary policy, it is the outcome for the euro area as a whole that is decisive. However, developments in Germany are of considerable importance owing to the country's high weight. Data for German GDP are published on a quarterly basis. The first official flash estimate by the Federal Statistical Office for the past quarter is issued with a time lag of just over six weeks.[2] Thus, depending on the given point in time, an assessment of the developments in the past quarter, which has not yet been published by the Federal Statistical Office, or for the current quarter is needed as well. Although such an assessment does not refer to the future, it will also be referred to below as a forecast for the purpose of this article.[3]

Short-term economic forecast regularly communicated in qualitative form

The short-term economic forecast for Germany covers up to three quarters. It is incorporated into the Bundesbank's ongoing economic assessment, which is also regularly communicated to the general public. This is usually issued in qualitative form, such as in the Bundesbank's Monthly Report. In addition, the results of the short-term forecast form the starting point for the macroeconomic forecasts for Germany, which are prepared every six months using the Bundesbank's macroeconometric model and are incorporated into the Eurosystem staff macroeconomic projections for the euro area.

Combining model-based forecasts with expert knowledge

The Bundesbank's short-term economic forecasts are based on several automated, econometric forecast models.[4] These models use systematic relationships observed in the past between a large number of relevant economic indicators and the respective target variable. These purely model-based estimates serve as the starting point for the economic forecast and are supplemented with expert knowledge. The results of the individual models are weighted or corrected on the basis of their specific strengths and weaknesses in order to obtain as accurate a picture as possible of the economic situation and the short-term outlook. Furthermore, account is taken of additional information that is difficult to capture in the models. Such information includes, for example, one-off factors such as strikes, flu epidemics or other exceptional events.

Taking different forecast models into account is advantageous

The economic literature offers a wide range of forecast models, which differ, for example, in their basic approach, their degree of complexity, or in terms of their underlying indicators. Some features – such as a good interpretability of the model results – bring obvious advantages. Other features turn out to be an advantage in certain situations, but a drawback in others. This is true, for example, with regard to the speed and the extent to which the fore-

1 See European Central Bank, The outcome of the ECB's evaluation of its monetary policy strategy, ECB Monthly Bulletin, June 2003, pp. 79-92.
2 For the euro area, Eurostat publishes a preliminary flash estimate just over four weeks after the end of each quarter.
3 In the English-language academic literature, it has become common to use the term "backcasts" for forecasts that refer to periods in the past but for which no data have yet been published. By contrast, forecasts for the current quarter are referred to as "nowcasts" and those for future quarters as "forecasts". For a definition of the terms, see M. Bańbura, D. Giannone and L. Reichlin (2011), Nowcasting, in M. P. Clements and D. F. Hendry (eds.), The Oxford Handbook of Economic Forecasting, pp. 193-224.
4 The short-term economic forecast methods used in the Bundesbank's day-to-day work were described in detail in Deutsche Bundesbank, Forecasting models in short-term business cycle analysis – a workshop report, Monthly Report, September 2013, pp. 69-83.

casts are adapted to take account of new information. A speedy adjustment can be advantageous, for example, if an economic turning point should occur. On the other hand, it can be disadvantageous if, for example, current data are affected on a large scale by erratic disruptive factors. For risk diversification purposes, it has therefore proven helpful to take forecast models with differing features into consideration.[5]

Criteria for selecting the models

The Bundesbank selects the models to be used for its short-term economic forecasts based on several criteria. First of all, the most important aspect in this regard is the forecast performance, i.e. the ability of a model to produce accurate forecasts of the target variable. Moreover, it should be possible for an economically plausible explanation of the results to be derived from the model. In addition, the forecasts should demonstrate a certain degree of stability over time up to when the target variable is published. Although the degree of accuracy generally tends to increase along with the inflow of new information, the forecast results of some models fluctuate quite strongly in practice. Frequent and large forecast adjustments in different directions make it more difficult to interpret and communicate the results.

Further requirements: inclusion of various indicators, …

Besides these general criteria, a number of more specific requirements are placed on each forecast model. As a general rule, a large number of different economic indicators should be used in order to cover, as far as possible, all areas relevant to economic activity. Furthermore, certain special factors such as weather or calendar effects can also be taken into account using appropriately designed variables. This prevents special developments in individual areas from being overlooked. In addition to its economic relevance, the availability of sufficiently long time series is ultimately also crucial in deciding whether to include an indicator in a forecast model's dataset.

As in the case of GDP, economic indicators are also published with a time lag in some cases. Furthermore, these publication lags differ from indicator to indicator. At the end of September, for example, data on industrial output are available only up to and including the end of July, whereas the ifo business climate index is already available for September. The resulting gaps of varying lengths in their availability create the characteristic "ragged edge" which is inherent to macroeconomic datasets. The forecast models should make use of all available information at any given time and therefore fill or bridge such gaps in a suitable manner. The predictive power of an indicator can therefore result not only from a possible leading character (e.g. for new orders in industry or survey-based business expectations), but also from the fact that it is available earlier than the target variable. The publications of numerous important "hard" economic indicators in the official statistics, such as industrial output, new orders in industry or foreign trade figures conglomerate with a publication lag in the second week of each month, while many of the "soft" survey-based sentiment indicators are usually published in the fourth week of each month. This is why the Bundesbank updates its short-term forecasts twice per month.

… diverging publication lags …

… and mixed data frequencies

A further typical feature of macroeconomic datasets is that the indicators are published at different time intervals. While new observations for GDP and its components (but also for some other economic indicators, such as the ifo capacity utilisation in the manufacturing sector) are available only for quarterly periods, most hard and soft indicators are published on a monthly basis.[6] The forecast models should therefore be able to process such differences in the data frequencies.

5 See A. Timmermann (2006), Forecast combinations, in G. Elliot, C. Granger and A. Timmermann (eds.), Handbook of Economic Forecasting 1, pp. 135-196.
6 Although some indicators are also published weekly, daily or even every minute (e.g. oil prices, weather data or stock prices), a potential gain in information through the direct modelling of higher-frequency time series is usually offset by a more complex estimation procedure, which means that indicators aggregated to the monthly frequency are usually used.

Established model classes in academia and in practice: single equation models, …

In the academic literature, several model classes are used for forecasting purposes. A first model class, which has also become established in practice, is made up of single equation models. In these models, the influence of a small number of selected indicators on the target variable (e.g. GDP or one of its components) is estimated using single equations. The forecast values from several single equations are then often consolidated by means of simple or weighted averaging. Bridge equation models are one representative of this model class that are often used by central banks.[7]

… models that can process large volumes of data …

The characteristic feature of a second model class is the ability to process large volumes of data. On the one hand, this includes "condensing" models where the information from all indicators is summarised. Dynamic factor models are among those that belong to this model class. They consolidate the information of a potentially very large number of indicators, which is often similar over the course of the economic cycle, into just a few factors. A simultaneous or lagged relationship is established between these factors and the target variables.[8] On the other hand, it includes models in which the complexity is reduced by means of an implicit variable selection instead of aggregating information from a variety of indicators.[9]

… and VAR systems

Vector autoregressive models (VAR models) form another model class. In a system of multiple variables, each variable is dependent on its own past values and those of the other variables contained in the system.[10] Due to their strong interdependencies, VAR models have so far provided precise estimation results only for relatively small systems, and could therefore be used only to a limited extent for short-term forecasting. In the meantime, however, promising approaches for large VAR systems are emerging from research.

Bundesbank uses factor and bridge equation models, supplemented by a forecast model for industrial output

At the Bundesbank, two representatives of the above-mentioned model classes, a bridge equation model and a dynamic factor model, have been in use for quite some time for the short-term forecasting of GDP.[11] Both models are supplemented by a separate forecast model for industrial output. The manufacturing sector occupies a prominent position in terms of growth dynamics in the German economy. Not only does this sector account for a large share of total economic value added, at somewhat more than one-fifth, the industrial sector in Germany also has close ties with many other domestic economic sectors and is, not least owing to its strong focus on exports, firmly integrated into the global economy. This means that industry is an important impulse generator for the economy. Separate modelling makes it possible to cross-check the GDP forecasts with those of the other models.

7 In addition to bridge equation models, their counterparts for data with mixed frequencies, MI(xed) DA(ta) S(ampling) models, also belong to the single equation models, see C. Schumacher (2016), A comparison of MIDAS and bridge equations, International Journal of Forecasting 32, pp. 257-270. Error correction models, in which potential long-term relationships are explicitly recorded, also belong to this model class. Single equations with monthly indicators are used by Norges Bank, for example. See K.A. Aastveit, K. Gerdrup and A.S. Jore (2011), Short-term forecasting of GDP and inflation in real-time: Norges Bank's system for averaging models, Norges Bank, Staff Memo 9/2011. Another example is the Bank of England, which regularly produces forecasts based on bridge equations and MIDAS models, see N. Anesti, S. Hayes, A. Moreira and J. Tasker (2017), Peering into the present: the Bank's approach to GDP nowcasting, Bank of England Quarterly Bulletin Q2 2017.

8 Dynamic factor models are a widely used tool among central banks for short-term forecasts. One example is the approach of the Federal Reserve Bank of New York, see B. Bok, D. Caratelli, D. Giannone, A. Sbordone and A. Tambalotti, Macroeconomic nowcasting and forecasting with big data, Federal Reserve Bank of New York Staff Reports, No 830, November 2017.

9 These include the Lasso and Boosting approaches. The "Least absolute shrinkage and selection operator" (Lasso) approach is a regression procedure in which the coefficient of a variable is either unequal to zero (significant indicator) or is "shrunk" to zero (insignificant indicator). Thus, a variable selection takes place simultaneously in the estimation, see R. Tibshirani (1996), Regression analysis and selection via the Lasso, Journal of the Royal Statistical Society Series B 58, pp. 267-288. Boosting is an iterative procedure in which the indicator with the greatest explanatory content in relation to the variation of the target variable that is still to be explained is selected in each step, see Y. Freund (1995), Boosting: a weak learning algorithm by majority, Information and Computation 121 (2), pp. 256-285.

10 The use of VAR models for macroeconomic analyses and forecasts was originally recommended by Christopher Sims, see C.A. Sims (1980), Macroeconomics and reality, Econometrica 48 (1), pp. 1-48.

11 See Deutsche Bundesbank (2013), op. cit.

Different variants for each model

Forecasts of different variants are calculated for all three models, which differ, for example, with regard to the indicators considered and the various specifications. In order to reduce the resulting multiplicity of results, the outcome of the different variants is averaged for each model. By doing so, each variant contributes to the overall result of the respective model. Averaging across different variants enhances the temporal stability of the results. Furthermore, the dispersion of the results among the model variants provides an initial indication of the uncertainty of the model forecasts. By considering the results for each model independently, the various strengths and weaknesses of each model are taken into account in the overall assessment.

Revision and enhancement of the set of instruments for short-term economic forecasts

Revision of forecast models resulted in modifications and addition of VAR model

Both the bridge equation model and the factor and industry models have provided satisfactory results in recent years. Nevertheless, it is advisable to review the models used from time to time and, if necessary, revise or replace them. For example, their forecasting quality may change over time owing to new framework conditions. Possibilities of improvement can also arise from weaknesses in the given model or new findings in the academic literature. It is against this backdrop that both the bridge equation model and the industry model have been revised.[12] Furthermore, the set of instruments used for forecasts has been supplemented by a VAR model.

Basic features of the bridge equation model

The bridge equation model is an established cornerstone of the Bundesbank's model-based short-term business cycle analysis. It consists of a system of single equations, the structure of which is based on that of the national accounts. It can be used not only to forecast GDP directly, but also to forecast its components on the supply and demand side. In addition to direct GDP forecasts, two variants disaggregated to different depths are calculated for each side. The modelling of the sectoral driving forces and demand impulses behind GDP growth makes it easier to interpret and communicate the forecast results. Above and beyond that, the disaggregated approach plays a vital role in dovetailing the short-term forecast with the medium-term projection, which also focuses on the expenditure structure of the GDP projection. The core idea of the bridge equations is to establish a link between the quarterly variables to be forecast, i.e. the GDP growth rate or one of its components, and the monthly economic indicators: the various data frequencies are "bridged", as it were. To this end, the respective monthly economic indicators are themselves extrapolated as a prior step, with suitable leading indicators also being used where available. The resulting forecasts on the monthly frequency are then aggregated over time to the quarterly frequency and inserted into the previously estimated bridge equation with the national accounts variable.

Improvements thanks to deeper disaggregation of components, …

As part of the fundamental overhaul, the original version of the model was improved upon in a number of respects.[13] One of the key modifications is an enhanced degree of detail in the disaggregated approaches.[14] This means that calculations are now performed on the basis of five (instead of four) and 15 (instead of seven) components on the supply side, and on the basis of four (as previously) and 14 (instead of eight) components on the demand side of GDP (see the table on p. 20). A particular point to note is

12 The factor model satisfactorily fulfils a number of requirements – use of large data volumes, filling the ragged edge, taking account of different publication frequencies – even by current standards. With this in mind, the decision was taken not to revise the factor model for the time being.
13 A description of the original version may be found in Deutsche Bundesbank (2013), op. cit.
14 The revised model framework is documented in N. Pinkwart, Short-term forecasting economic activity in Germany: a supply and demand side system of bridge equations, Deutsche Bundesbank Discussion Paper No 36/2018. The system outlined therein provides the basic framework for the model presented in this article for day-to-day business cycle analysis at the Bundesbank.

System of bridge equations

Supply side	Demand side
Disaggregated, 29 components	
GVA[1] agriculture, forestry and fishing	Private consumption
GVA mining and quarrying	Public consumption
GVA manufacturing	Private investment in machinery and equipment
GVA energy and water supply, waste management, etc.	Public investment in machinery and equipment
GVA construction	Private residential investment
GVA wholesale and retail trade; repair of motor vehicles and motorcycles	Corporate construction investment
GVA transportation and storage	Public construction investment
GVA accommodation and food service activities	Private other investment
GVA information and communication	Public other investment
GVA financial and insurance activities	Changes in inventories
GVA real estate activities	Exports of goods
GVA business services	Exports of services
GVA public administration, education, human health	Imports of goods
GVA other service activities	Imports of services
Net taxes on products	
Disaggregated, 9 components	
GVA agriculture	Consumption
GVA production sector excluding construction	Gross investment
GVA construction	Exports
GVA services	Imports
Net taxes on products	
Directly aggregated GDP forecasts	
Gross domestic product	Gross domestic product

1 Gross value added.
Deutsche Bundesbank

that gross value added in the services sectors, on the supply side, and investment, on the demand side, have been disaggregrated more deeply.[15]

… a broader pool of indicators and a more flexible extrapolation method, …

Another new feature is the upstream forecasts of monthly indicators. The dataset from which the most accurate indicators are selected has been extended to cover around 130 time series. In another modification, some economic indicators are now no longer extrapolated, as they were before, with just a single leading survey indicator; instead, the information content from multiple leading indicators is tiered and analysed in multiple steps (for example, the ifo export expectations help in extrapolating foreign industrial orders and these, in turn, can be used to forecast exports of goods). Also, the revised model now estimates the effects of "bridge" days, school holidays or unseasonal weather conditions for some economic indicators.[16] Lagged regressors are added to capture the subsequent countermovements typically associated with such one-off effects.

… combined forecasts potentially based on multiple single indicators, …

Alongside the single equation models used hitherto, which each, as a rule, analyse just a single selected indicator, potential multi-indicator-based model selection and combination methods were also examined with a view to forecasting the components of GDP.[17] While a number of the established single indicators (e.g. industrial output for forecasting gross

15 Deeper disaggregation also allows the calculation of special aggregates, such as corporate investment (private gross fixed capital formation excluding residential construction), which are used in the context of the Bundesbank's macroeconomic projections.
16 Specifically, calendar regressors take account of "bridge" days between public holidays and weekends, and the summer school holidays. Exceptional weather conditions are captured using an ice day indicator; see Deutsche Bundesbank, The impact of weather conditions on gross domestic product in the latter part of 2013 and early part of 2014, Monthly Report, May 2014, p. 54-55. In line with the relevant European guidelines, no corrections are made for these effects in the official seasonal and calendar adjustment; see Deutsche Bundesbank, Calendar effects on economic activity, Monthly Report, December 2012, pp. 51-60; and Eurostat (2015), ESS guidelines on seasonal adjustment, ISSN 2315-0815.
17 Means of multiple single equations have hitherto been used only for direct GDP forecasts.

value added in manufacturing) were ultimately retained, in many cases it was found useful to prepare combined forecasts based on multiple single indicators (see the adjacent table). Both simple arithmetic averaging and mean values weighted according to their historical forecast performance are used here. For some components that are particularly difficult to forecast (e.g. the gross value added of financial and insurance service providers or public consumption expenditure), however, it was not possible to find any model specification that produced a forecast which outperformed a naive benchmark forecast based on the historically observed mean. In such cases, the historical mean or autoregressive extrapolation will be used as the forecast in future.

… and weighted averaging of supply- and demand-side GDP forecasts

The variants of the bridge equation model yield six different GDP forecasts overall. To detect tensions between the supply and demand sides of GDP, the three variants on both sides were each previously condensed by way of arithmetic averaging, but then evaluated separately. Given that the forecast errors of supply- and demand-side GDP forecasts are not fully correlated, above all for short forecast horizons, it is, however, possible to significantly reduce the mean error by combining the supply- and demand-side forecasts. It was found that a slight overweighting of the supply-side results delivers the best forecast performance.

VAR models can capture interaction between indicators

The new bridge equation model (and, to some extent, the factor model as well) analyses the impact of the indicators used on the target variable only from a single direction. Furthermore, it takes little or no account of the relationships among the variables. VAR models estimated using traditional methods, which permit dynamic interaction between all the variables, have been used by central banks rather infrequently for regular short-term forecasts. This is because each variable in VAR models depends on the lagged values of all the variables fed into the model. Hence there is a need to estimate many parameters, even in models with a small number of variables, which tends

Bridge equation specifications

Component	Specification[1]
GDP supply side	Combined forecast (60 indicators)
GVA[2] agriculture, forestry and fishing	Combined forecast (9 indicators)
GVA production sector excluding construction	Production in the production sector excluding construction
GVA mining and quarrying	Mining production
GVA manufacturing	Industrial production
GVA energy and water supply, waste management, etc.	Energy production
GVA construction	Production in the main construction sector
GVA services	Combined forecast (18 indicators)
GVA wholesale and retail trade; repair of motor vehicles and motorcycles	Combined forecast (15 indicators)
GVA transportation and storage	Naive mean forecast
GVA accommodation and food service activities	Real revenues from accommodation and food service activities
GVA information and communication	Simple average (10 indicators)
GVA financial and insurance activities	Naive mean forecast
GVA real estate activities	Combined forecast (19 indicators)
GVA business services	Combined forecast (8 indicators)
GVA public administration, education, human health	Naive mean forecast
GVA other service activities	Naive mean forecast
Net taxes on products	Combined forecast (9 indicators)
GDP demand side	Combined forecast (71 indicators)
Consumption	Combined forecast (17 indicators)
Private consumption	Combined forecast (47 indicators)
Public consumption	Naive mean forecast
Gross investment	Combined forecast (23 indicators)
Private investment in machinery and equipment	Combined forecast (32 indicators)
Public investment in machinery and equipment	Naive mean forecast
Private residential investment	Production in the main construction sector
Corporate construction investment	Combined forecast (14 indicators)
Public construction investment	Production in the main construction sector
Private other investment	Combined forecast (19 indicators)
Public other investment	AR forecast
Changes in inventories	Simple average (7 indicators)
Exports	Combined forecast (14 indicators)
Exports of goods	Combined forecast (11 indicators)
Exports of services	Combined forecast (5 indicators)
Imports	Combined forecast (44 indicators)
Imports of goods	Combined forecast (14 indicators)
Imports of services	Combined forecast (6 indicators)

1 Selected single indicators, autoregressive (AR) forecast, extrapolated from the sample mean (naive mean forecast), the simple (arithmetic) average of multiple forecasts, or the combined forecast with weights based on past forecast errors in the Q2 2006-Q1 2018 evaluation period (specifying the number of indicators included in the combination with a weight different from zero for at least one forecast horizon). **2** Gross value added.

Deutsche Bundesbank

to be linked to a high degree of forecasting uncertainty and severely restricts the number of variables that the system can analyse.[18] Diverging publication lags and data frequencies impede modelling further.

New VAR model complements models in use …

Recent developments in econometric methods as well as advances in the performance of modern computing systems mean that it is now possible to use flexible VAR models estimated employing Bayesian methods for short-term forecasts, which meet the requirements set out above.[19] It was for this very reason that the short-term forecasting models used hitherto at the Bundesbank were augmented by a VAR model of this kind.[20] This model is based on monthly data in order to maximise the amount of information it can analyse. Time series available only at quarterly intervals (such as GDP) therefore need to be transformed into monthly data. This, just like the filling of data gaps caused by the ragged edge problem, is done within the framework of the model.[21] Thus, (previously) observed variables are taken into account when filling data gaps, and interpolated monthly values (as in the case of GDP) always add up to the known quarterly figure.

… and provides density forecasts for all variables as well as monthly GDP series

One advantage of the Bayesian VAR approach is that it can also account for the uncertainty of a forecast in a consistent manner. Unlike in a point forecast (where just a single value is estimated), which is the focus of the models described above, Bayesian VARs produce what are known as density forecasts (i.e. forecasts for the entire probability distribution). This yields a more comprehensive picture of the possible path of indicators and GDP. The range and potentially asymmetric shape of the probability distributions can point to downside and upside forecast risks. One interesting by-product of this VAR model is that it also delivers monthly estimates for GDP, both for the past and for the forecast period.[22]

The model used to forecast monthly industrial output, which augments the forecast models used for GDP, has also been overhauled. The key objective of this was to enhance transparency and improve the readability of forecasts. The number of model variants was therefore substantially reduced.[23] Non-linear specifications have been added to some of the proven linear approaches, however.

Modified industry model based on few, albeit better specified, model variants

Industry model based on relationship between output, new orders and inventories or order volumes

In Germany's industrial sector, incoming orders which are processed over a period of time account for a very substantial share of economic activity. In addition, enterprises can be expected to adjust their production if their inventories and order volumes diverge from values they consider desirable from a commercial perspective. This is why the modelling approach used in the industry model is based on a close relationship between industrial output, new orders as well as inventories and order volumes.

The fundamental relationship between these variables can be modelled in a number of dif-

18 This issue was often circumvented in empirical short-term forecasting by combining multiple small VARs; see, for example, K. A. Aastveit, K. Gerdrup and A. S. Jore (2011), op. cit.

19 Bayesian estimation methods allow direct estimates to be made of relatively large VARs; see M. Bańbura, D. Giannone and L. Reichlin (2010), Large Bayesian vector autoregressions, Journal of Applied Econometrics 25 (1), pp. 71-92.

20 The VAR model is presented in T. B. Götz and K. Hauzenberger, Large mixed-frequency VARs with a parsimonious time-varying parameter structure, Deutsche Bundesbank Discussion Paper No 40/2018, which also explains additions made to the model to account for time-varying parameters and stochastic volatility. The model outlined therein provides the basic framework for day-to-day business cycle analysis at the Bundesbank, where 12 monthly indicators are used alongside GDP. This approach is based on F. Schorfheide and D. Song (2015), Real-time forecasting with a mixed-frequency VAR, Journal of Business and Economic Statistics 33 (3), pp. 366-380.

21 To this end, an iteration is made between two "blocks" of the model, with the last result of one block being used as the starting point for estimating the other block. Here, the first block interpolates the data gaps described above, while the second block is used to estimate the relationships between the variables.

22 The Office for National Statistics recently introduced a new publication model for GDP in the United Kingdom. In the new publication model, a rolling three-month estimate is calculated based on monthly estimates of GDP; see J. Scruton, M. O'Donnell and S. Dey-Chowdhury, Introducing a new publication model for GDP, Office for National Statistics article of 3 May 2018.

23 The previous version of the model calculated slightly more than 3,400 model variants. See Deutsche Bundesbank (2013), op. cit.

Error correction and multi-co-integration approaches used for modelling

ferent ways. In a first, simple group of variants of the industry model, there is assumed to be a relationship – one that may arise with a certain time lag – between the fluctuations in new orders and the change in production. This short-term effect is modelled using a two-dimensional VAR structure. Given that production and new orders are likely to be driven by the same trend over the long term, their relationship is captured in a second group of variants using an error correction model.[24] The third group of variants, based on a more detailed multi-co-integration approach, goes one step further by additionally including the relationship between production and inventories or order volumes as a further long-term relationship in the error correction model.[25] Since it is not yet possible at present to use the time series on order volumes provided by the Federal Statistical Office owing to their short data history, the order volumes data are either determined by measuring the accumulated deviations between new orders and output or they are proxied using suitable survey data. Moreover, as with the bridge equations, the effects of school holidays and "bridge" days are taken into account beyond the usual calendar and seasonal factors. Lastly, the results of the 12 variants in total are arithmetically averaged.

Model forecast performance

Analysis of forecast errors

The first step in assessing the opportunities and limitations presented by these models is to evaluate how each of them has performed in forecasting in the past.[26] Furthermore, case studies for the first half of 2018 are used to highlight how the models can be deployed in conjunction with the expert assessment in practical business cycle analysis (see the box on pp. 25-27). A commonly used measure of forecast performance – the mean absolute forecast error (MAFE) of the quarter-on-quarter rate of change of the target variable in question – is used here.[27] As a result of the steady inflow of information, the forecast error ought to diminish as the forecast horizon approaches. That is why a distinction is made between different forecast horizons – that is to say, in this specific case, the gap, measured in weeks, between the forecast date and the publication date.

Evaluation period: Q1 2010 to Q1 2018

The period from the first quarter of 2010 to the first quarter of 2018 has been chosen as the evaluation period.[28] While this means that the calculation of the mean forecast errors is generally based on a rather long period of time, an analysis of this kind should ideally cover an entire business cycle in order to gain an impression of how the models behave at every stage of the business cycle. This is possible only to a limited extent in the period selected because the German economy has been in an extended upswing since mid-2009, which faltered only briefly when the euro area crisis struck in 2012. As a result, a model which has proven to be quite accurate in such a protracted period of expansion but possibly generates large forecast errors during downturns might be overrated in

24 This specification consists of two parts: one relationship for the long-term equilibrium, and one relationship for short-term deviations of the two flows from this long-term relationship.

25 The multicointegration approach was already incorporated into an earlier version of the model; see Deutsche Bundesbank (2013), op. cit. Furthermore, the error correction models and the multicointegration approach now also used non-linear model variants which allow for an asymmetric adjustment of the variables to the respective equilibrium terms; see C. W. J. Granger and T.-H. Lee (1989), Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models, Journal of Applied Econometrics 4, pp. 145-159.

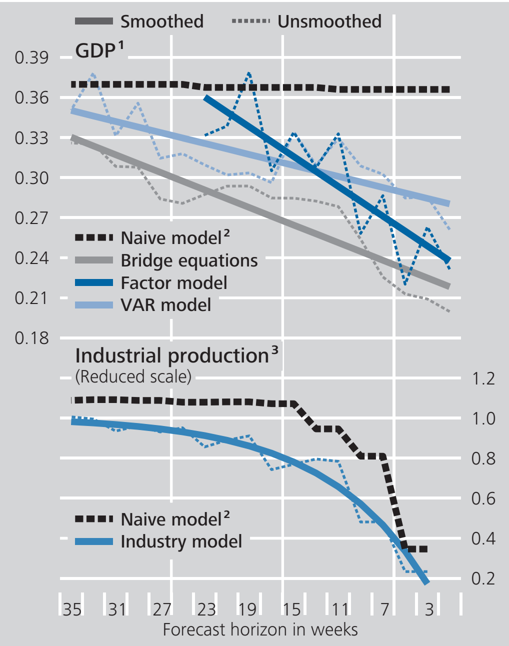
26 Put simply, simulations are used to check what forecasts and forecast errors the models would have produced if they had already been used in the past.

27 The MAFE for a given horizon is calculated by taking the absolute values of the differences between the forecasts and the actuals and averaging them arithmetically. The MAFE is just one possible measure of forecast performance; alternative statistics are available according to preference. For example, using the square root of the mean squared error would emphasise large forecast errors. Furthermore, rather than looking at growth rates, an alternative, which might be of particular interest to monetary policymakers, depending on their objective, would be to calculate the errors in levels.

28 The dataset used in the forecast evaluation is current as at 24 May 2018, i.e. following publication of the national accounts figures for the first quarter of the current year. Given that real-time data were not available for all the times series observed in this specific case, the evaluation was carried out in "pseudo" real time (i.e. on the basis of the final dataset without considering any historical data revisions).

Forecast errors*

Percentage points



***** Mean absolute forecast errors (from Q1 2010 to Q1 2018) for different forecast horizons, estimated from quarter-on-quarter changes. **1** Adjusted for price, seasonal and calendar effects. **2** The historical mean value of each target variable is extrapolated as the forecast. **3** Adjusted for seasonal and calendar effects.

Deutsche Bundesbank

terms of its forecast performance. Alternatively, it would be possible to choose a significantly longer period of time that includes the Great Recession of 2008-09. That would give analysts a better idea of whether (and possibly how quickly) models are capable of flagging up critical situations. Note, however, that this recession and the rapid recovery that followed it were exceptionally strong by historical standards. An atypical period ("outlier") of that kind could distort the evaluation results if it is fraught with particularly large forecast errors.

The chart above shows the forecast performance – in the form of the MAFEs – for each short-term forecast model. This performance is compared with a naive benchmark model in which the rate of change in the target variable is extrapolated by its historical mean.[29] Forecasts of economic activity often prove to be informative for up to three quarters ahead, in the sense that they each deliver a forecast performance for the quarter-on-quarter rate that is superior to that of the naive benchmark model.[30] It is for this reason that forecast horizons ranging from one week up to 35 weeks are considered.[31]

Models deliver accurate and complementary forecasts

In the period under review, the models were found to be more accurate than the naive benchmark model across almost all the horizons observed. Comparing the models with each other revealed that the bridge equations generate the smallest forecast errors for all the forecast horizons. It should, however, be noted that this error represents an average measure of the forecast performance over the entire evaluation period. For individual quarters, the factor model or the VAR model certainly predict the GDP increase with greater precision.[32] This makes looking at all three model classes a sensible course of action. The VAR model performs quite well, relative to the factor model, for forecast horizons of more than nine weeks, but it is less accurate for horizons of seven weeks or less. The industry model also clearly outperforms the naive benchmark model. In

29 Fairly small fluctuations in the forecast error of the naive benchmark model for GDP and – for long forecast horizons – for industrial output can be attributed to slight changes in the long-term average resulting from the publication of new figures. In the case of industrial output, publication of data from the first month, and particularly the second month of the quarter being forecast rapidly reduces the errors of the naive benchmark forecast. Publication of the latest monthly data for the prior quarter again significantly reduces the forecast error because the statistical overhang feeds into the forecast.
30 See J. Breitung and M. Knüppel (2018), How far can we forecast? Statistical tests of the predictive content, Deutsche Bundesbank Discussion Paper No 07/2018. A definition of "informative" forecasting may also be found in M. P. Clements and D. F. Hendry (1998), Forecasting economic time series, Cambridge University Press.
31 Regarding the definition of the forecast horizons, it is assumed here for the sake of simplicity that each month consists of exactly four weeks. For the factor model, the MAFEs are calculated for a forecast horizon of up to 23 weeks. Owing to its interpretation as an "average" forecast error that covers all the potential sources of error, the MAFE is also often used to measure the uncertainty of a point forecast. For this purpose, it is worth depicting the forecast error in a smoothed way, since the empirical mean forecast error often does not follow a consistently monotone path. See Deutsche Bundesbank, Uncertainty of macroeconomic forecasts, Monthly Report, June 2010, pp. 29-46.
32 For example, the factor model provided more accurate estimations of GDP growth for longer forecast horizons in the second quarter of 2018 than the bridge equations (see the box on pp. 25-27).

Business cycle analysis in practice – first half of 2018

This box describes how the results produced by the various econometric models which the Bundesbank uses to forecast short-term developments in gross domestic product (GDP) and industrial output are fed into practical business cycle analysis, using the forecasts for the first and second quarter of 2018 as an example. The forecasts are analysed in real time, i.e. as they actually developed under the constant influx of new information. In terms of day-to-day operations, the results of the various short-term forecasting models make up the foundation of the expert assessment, which is used as the basis for internal and external communication purposes. Besides empirical knowledge, the assessment by business cycle experts takes into account additional information which the models cannot process or are unable to process in an appropriate manner. In the first two quarters of 2018, these were, above all, the strikes in late January and early February in the metal-working and electrical engineering industries and the impact of the severe flu epidemic in February and March.[1]

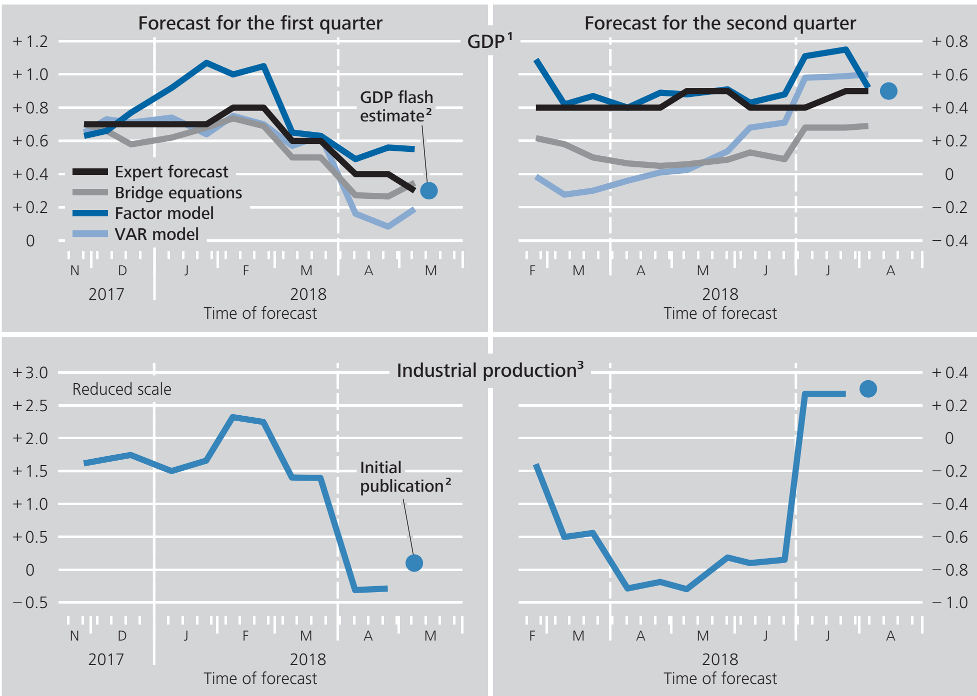
In the second half of November 2017, i.e. after publication of the GDP flash estimate for the third quarter of 2017, the time horizon of the short-term forecasting models was extended to include the first quarter of 2018. From then on until publication of the target value in mid-May, each model produced forecasts twice a month (see the chart on p. 26). New information from hard and soft indicators was fed into the models in the second and fourth week of each month, respectively. The forecasts generated by the four different models were then used by business cycle analysts as a basis for discussion in order to produce the expert forecast on the GDP growth rate. As from mid-February, the GDP growth rate for the fourth quarter of 2017, which had only just been published, was included in the calculations. The growth rate for the second quarter of 2018 was initially forecast at the same time, and the procedure outlined was repeated.

Following publication of a strong GDP growth rate of 0.8% for the third quarter of 2017 in mid-November 2017, the models quite unanimously indicated for some time that the brisk pace of economic growth would continue into the first quarter of 2018. The very good order volume and the excellent sentiment in the manufacturing sector suggested that industrial activity would stay strong and that this sector would continue to be the main engine driving the upswing.[2] This was reflected in the forecasts of the industry model as well. Correspondingly, the macroeconomic forecast from December 2017 predicted a sharp quarter-on-quarter increase in real GDP of 0.7% in both the last quarter of 2017 and the first quarter of 2018. Regarding the next few quarters, the expectation was that the rapid pace of expansion would normalise towards a growth rate that was slightly above its potential.[3] With the actual GDP growth rate for the fourth quarter of 2017 coming in at 0.6%, the industrial sector still in excellent shape, and export growth picking up sharply at year-end, a strong GDP growth rate in the first quarter of 2018 was

1 See Deutsche Bundesbank, Commentaries, Monthly Report, April 2018, pp. 5-12.
2 See Deutsche Bundesbank, Economic conditions in Germany, Monthly Report, November 2017, pp. 41-52.
3 See Deutsche Bundesbank, Outlook for the German economy – macroeconomic projections for 2018 and 2019 and an outlook for 2020, Monthly Report, December 2017, pp. 15-34.

Short-term forecasts for the first and second quarter of 2018

Change against previous period in %; end-of-period values



1 Adjusted for price, seasonal and calendar effects. 2 Source: Federal Statistical Office. 3 Adjusted for seasonal and calendar effects.
Deutsche Bundesbank

still conceivable at the end of February 2018.[4]

However, having touched record highs around the turn of the year, sentiment in the manufacturing sector became progressively more subdued in the following. In addition, the hard data for January and February published at the beginning of March and in early April, respectively, were far more downbeat than expectations based on leading indicators had suggested. This was reflected in significantly less favourable model results for the first quarter. Besides the quick pace of underlying economic activity beginning to return to normal sooner than expected,[5] it was assumed that, in particular for February, one-off factors that are difficult to quantify such as the strikes and the flu epidemic had played a role here, too. Given that the models are unable to identify these negative, temporary one-off factors, they implicitly extrapolate their effects into the future. It was for this reason that the expert forecast was deliberately placed at the upper end of the models' range until the end of April. Following large forecast errors initially, the growth rate of 0.3% published in mid-May for the first quarter was reached in the end.[6]

The fact that the forecast errors produced by the models for the first quarter initially grew over time in some cases (in particular for the factor model) was attributable to the exceptionally positive leading indicators

4 See Deutsche Bundesbank, Economic conditions in Germany, Monthly Report, February 2018, pp. 45-56.
5 In retrospect, the revision of the national accounts in August 2018 shows that the economy differs somewhat from the picture painted in the second quarter of 2018. While growth rates were even stronger at the beginning of 2017, they were lowered somewhat for the subsequent quarters. The Federal Statistical Office now reported rates of 0.6% for the third and 0.5% for the fourth quarter of 2017.
6 In August 2018, the growth rate was revised to 0.4%.

for industry at the beginning of the year. Yet, as the corresponding hard data were published for the reference period, the forecast errors gradually narrowed in almost every instance up to publication of the target value. The factor model almost consistently provided more optimistic forecasts than the VAR model or the bridge equation model. The latter two largely produced fairly similar results.

For the second quarter of 2018, the model forecasts painted a rather mixed picture for quite some time. This was a reflection of the elevated uncertainty surrounding the economic outlook. On the one hand, the factor model was signalling that the brisk pace of economic growth would continue. On the other, the bridge equations and the VAR model indicated that momentum would be very weak, significantly below the rate of expansion expected in the macroeconomic forecast of December 2017. In line with this, the industry model pointed to a distinct decline in industrial output for the second quarter.

Until the end of June, the expert forecast was at the upper end of the model results. Several factors played a role here. For one thing, it was assumed that the waning one-off effects, whose magnitude had been difficult to gauge and which had probably dampened growth in the first quarter, would lead to a countermovement in the second quarter, while the models implicitly extrapolated the effects of the one-off factors. That is why the business cycle experts even went as far as to revise their forecast slightly upwards when it became clearer at the beginning of May that GDP growth may even have been below potential growth in the first quarter. In addition, the order situation in industry, measured in terms of order volumes, was still very good despite a continuous decline in new orders. This was only partly taken into account by the models. Moreover, labour market developments remained favourable. This suggested that growth in the services sectors would stay robust. There are, however, only a few leading economic indicators available for these sectors, owing to which some model variants ascribe relatively little weight to them. In addition, the models tend, in the short term, to extrapolate the trajectory of the indicators in recent months. They therefore predicted further declines in sentiment indicators, ever decreasing new orders, and a steady drop in output figures in industry. By contrast, the experts assumed that the less favourable sentiment indicators were partly to be viewed as a return to normal following the very high levels reached in the second half of 2017 – a development they believed would be reflected in real economic data to a lesser degree than the models expected.

Published in June, weak hard data for industry in the reporting month of April suggested that the period of weakness in the industrial sector might even persist after the negative one-off factors had petered out. Good industrial data for May meant that those concerns faded into the background at the beginning of July. The VAR model and the bridge equations converged on the expert assessment as the subdued momentum expected by the models – in particular due to the extrapolated downward movement in industry – was overwritten by more favourable incoming data. The bridge equations, the VAR model and the industry model had already produced forecasts largely consistent with the once again weaker industrial data for June published at the beginning of August. The factor model was the only one to revise its forecast downwards. The expert forecast was also not adjusted any more, and it thus predicted the realised GDP growth rate of 0.5% published in mid-August quite well.

summary, then, it can be said that the short-term forecasting models already supply quite accurate forecasts for GDP and industrial output up to two to three quarters ahead and that the forecast performance improves even more strongly as the horizons shorten and the inflow of information increases.

Large forecast errors during 2008-09 crisis, though no qualitative impact on results

Two steps were taken to gauge the potential impact of the Great Recession. First, the evaluation period was extended to include the years of crisis. However, this did not produce any qualitative changes, apart from higher average errors, compared with the results for the shorter evaluation period. In the second step, the forecasts produced by the individual models for the 2008-10 period were compared with each other. As before, the bridge equation and factor models proved to be most accurate for short forecast horizons, whereas the VAR model performs well for medium- to long-term forecasts. All in all, the bridge equation model also turned out to be the most robust model on average for this period. Nevertheless, the VAR model would have pointed to a decline in GDP in the final quarter of 2008 as early as late August 2008, while the factor model and the bridge equations would not have done so until early and late October, respectively.[33]

Outlook

Modernising forecast models an ongoing process

The modifications described in this article do not mark the end of work on the forecast models. Indeed, the search for more suitable short-term forecasting methods is an ongoing process. Even though the factor model has not been revised for the time being, there should be an exploration of whether new insights from the academic literature could allow improvements to be made. Nonetheless, even if there are no immediate grounds for fine-tuning them, the existing models (as well as the expert forecast based on them) should be subject to evaluation from time to time.[34] This is the only way to measure their performance under changed economic and structural conditions. Furthermore, it is also possible to utilise entirely new models – such as, in the present case, the Bayesian VAR model – for short-term forecasting purposes.

Extend models to include density forecasts

Density forecasts are implicitly available for the VAR models presented in this article. These forecasts allow conclusions to be drawn about forecast uncertainty and statements to be made about the probability of certain events (for example, GDP growth above or below a certain threshold). Extending the other models discussed here to include density forecasts would offer additional insights into uncertainty and risk distribution.

Tapping new data sources

Furthermore, technological progress, particularly in the area of big data processing, allows new data sources to be tapped. Information obtained from online search queries or credit card transactions are just two examples of these. They could be of assistance in better assessing certain GDP components (e.g. private consumption) and thus also improve overall forecasting performance for GDP.[35] With a wide variety of data being surveyed and collected by private enterprises, research institutions and government agencies, it may be assumed that further promising data sources can be tested in the near future. Any such assessment needs to consider whether and to what extent such information can be useful for short-term business cycle analysis.

---

33 Even with close to a full body of information shortly before publication of the GDP data, none of the models would have indicated the severity of the economic downturn which set in at the end of 2008.
34 To simplify an evaluation process of this kind, all the forecast results – as well as the underlying datasets – are routinely archived, thus building up a real-time database that can be used for future analysis.
35 The extent to which data from online search queries are suited to forecasting German GDP has already been examined in a paper which uses a simplified supply-side version of bridge equations. However, the new data only have the potential to improve upon the existing body of survey data in isolated cases and following rigorous pre-selection. See T. B. Götz and T. A. Knetsch (2017), Google data in bridge equation models for German GDP, International Journal of Forecasting, forthcoming.