Structural modeling and forecasting using a cluster of dynamic factor models

Christian Glocker* Serguei Kaniovski†‡

July 2020

Abstract

We propose a modeling approach involving a series of small-scale dynamic factor models. They are connected to each other within a cluster, whose linkages are derived from Granger-causality tests. This approach merges the benefits of large-scale macroeconomic and small-scale factor models, rendering our Cluster of Dynamic Factor Models (CDFM) useful for model-consistent nowcasting and forecasting on a larger scale. While the CDFM has a simple structure and is easy to replicate, its forecasts are more precise than those of a wide range of competing models and those of professional forecasters. Moreover, the CDFM allows forecasters to introduce their own judgment and hence produce conditional forecasts.

JEL-Codes: C22, C53, C55, E37

Key Words: Forecasting, Dynamic factor model, Granger causality, Structural modeling

*Austrian Institute of Economic Research, Arsenal Object 20, 1030 Vienna, Austria. Email: christian.glocker@wifo.ac.at.

†Austrian Institute of Economic Research, Arsenal Object 20, 1030 Vienna, Austria. Email: serguei.kaniovski@wifo.ac.at.

‡**Acknowledgment**: We would like to thank Astrid Czaloun for excellent research assistance.

1 Introduction

The aim of this paper is to provide a coherent methodology for large-scale macroeconomic nowcasting and forecasting using a rich set of economic indicators available on a monthly basis. To this end, we propose a *Cluster of Dynamic Factor Models* (*CDFM*).

The analysis of the economy's short-term situation, and the projection of its future course are fundamental tasks of central banks and national and international institutions. In this context, two different kind of models are commonly used. On the one hand, practitioners utilize large-scale structural macroeconomic models and, on the other hand, small-scale time series models are becoming increasingly common in the tool-kit of economic forecasters.

Large-scale macroeconomic models can be defined as a set of stochastic equations with definitional and institutional relationships denoting the behavior of economic agents. The strong reliance on economic theory makes them useful in terms of interpreting the forecast, but at the same time, their inability to allow for the use of the information of soft-indicators and other high frequency data makes them less valuable in producing unconditional forecasts.

Small-scale time series models usually come in the form of dynamic factor models or versions thereof. They attempt to exploit the reduced-form correlations in observed macroeconomic time series, with little reliance on economic theory. These models are most commonly used for producing unconditional forecasts in a variety of environments, ranging from firm-level business forecasting to economy-wide macroeconomic forecasting.

The popularity of large-scale models stems from the need to forecast a wide range of macroeconomic variables. The set-up allows for producing conditional forecasts, rendering these models useful for scenario analysis. While these models allow to establish a model-consistent forecast of many macroeconomic variables jointly, they, however, face various drawbacks. These involver the rigidity of these models as regards (quick) extensions and modifications, the difficulty of integrating information from soft indicators, the rigid corset imposed on the equations due to their expected compliance with economic theory, and the lack of the possibility to process data of different frequencies and time series with missing observations. It is especially the last point that weighs heavily on the suitability of these models for producing unconditional forecasts, primarily in the context of short-term forecasting.

The drawbacks of the large-scale models, however, characterize the strength of small-scale models. The small dimension of the latter shapes their high degree of flexibility that proves advantageous in the context of quick extensions in the form of the inclusion of further variables, etc. The biggest advantage, though, is – in particular when estimated with the Kalman filter – the possibility to allow for missing observations and mixed frequencies. This makes these models especially useful for nowcasting and short-term forecasting. The key drawback of small-scale models is, however, the fact that they involve only a limited number of variables which impedes the creation of model-consistent forecasts of a large number of macroeconomic variables. This, in turn, impairs the usefulness of such models for scenario analysis, which would require conditional forecasts for a wide range of variables.

We propose a forecasting framework that merges the advantages of both large-scale and small-scale models, and that thus simultaneously excludes their individual drawbacks. We do so by means of a cluster of a series of small-scale factor models. In particular, we create a variety of dynamic factor models. Each of these is individually specified, estimated and evaluated in terms of its out-of-sample predictive accuracy. The individual models are linked together in a cluster that establishes the interfaces between the models and thus between the variables. The outline of the cluster, and hence also the specific interfaces, relies on economic theory giving rise to a *structural cluster*. Although the individual interfaces are motivated on the basis of economic theory, we use a statistical approach for the econometric confirmation thereof. Putting all individual dynamic factor models into the structural cluster results in a large-scale macroeconomic forecasting model. Its composition in the form of a multitude of small-scale models allows for a high degree of flexibility. The interfaces of the structural cluster, on the other hand, offer the possibility of creating a model-consistent forecast for a large number of variables. Furthermore, this approach allows the determination of conditional forecasts, rendering this set-up useful for scenario analysis.

We create this set-up by means of a two-step approach. In the first step we identify the structural cluster of key macroeconomic variables. To this purpose we utilize variables of the production (supply), expenditure (demand) and income account of the System of National Accounts. We rely on the concept of Granger-causality tests to identify the interfaces of the variables within the cluster. The cluster in a sense represents a network in which each node is represented by means of a particular economic variable from the System of National Accounts. The cluster identifies the link between the economic variables based on both economic theory and econometric tests. In a second step, we specify small-scale factor models for each variable in the cluster. The cluster identifies linkages between economic variables. Once having introduced a small-scale model fort each variable, the cluster then interlinks a series of small-scale dynamic factor models.

The CDFM features some eighty variables, including variables of the National Accounts, leading indicators, financial market variables, labor market indicators, price and wage specific variables and a series of variables of key trading partners. These variables are contained in twenty-five individual dynamic factor models. These in turn are inter-linked by having so-called *link-variables* in each individual dynamic factor model. The specification of the individual models involves a decision on (i) which variables to include and in case a variable is included, (ii) along which temporal displacement (contemporaneous, lagged, etc.). We do so by relying on a combinatorial algorithm (Glocker and Wegmüller, 2020) for the selection procedure where we consider the out-of-sample forecasting accuracy as the sole target. We use the Kalman filter for the estimation which in turn implies that the individual models, and hence also the CDFM, allow for missing observations and mixed frequencies in the selected time series. The model is thus especially useful for nowcasting and short-term forecasting, as it can deal with ragged ends easily.

Our proposed forecasting set-up allows for (i) establishing model-consistent predictions for a wide range of variables, (ii) performing scenario analysis by relying on the concept of conditional forecasts, (iii) using data of mixed frequency and series with missing observations, (iv) quick extensions for particular variables and models etc. Moreover, the estimation of all the macroeconomic aggregates allows us to not only to forecast real-time GDP but also to incorporate information on the components that explain the forecast, providing an insight into the causes of GDP forecast revisions. Not least, our approach allows to incorporate the subjective judgment of a practitioner and hence establish conditional forecasts.[1]

We find that the nowcasting and forecasting performance of the CDFM is not just more accurate in comparison to naive models; indeed, the forecasting performance of the CDFM

[1]Fildes and Stekler (2002) stressed the importance of interventions in the form of the introduction of judgment in producing the forecasts.

turns out to be superior to that of peers which allow for richer dynamics. For instance, the CDFM's forecasts for GDP are noticeably more accurate than those of classical small-scale dynamic factor models directly applied to GDP as well as various other competing models (e.g., ARIMA). The results prove to be robust for different model specifications. For instance, adding more indicators to our model does not necessarily improve the forecasting performance. Not least, the forecasts of the CDFM model encompass the forecasts from the professional forecasters.

Our contribution is related to Banbura and Rünstler (2011), Marcellino and Schumacher (2010), Angelini et al. (2011), Schumacher and Breitung (2008) and Barhoumi et al. (2008), who use the approximate dynamic factor model proposed by Giannone et al. (2008) to compute GDP forecasts which are continuously updated as well. As in these proposals, we diverge from the commonly used univariate bridge equations employed by Rünstler and Sedillot (2003) and Diron (2006) and from those which try to measure high-frequency objects (as real-time activity) on a daily or hourly basis, such as Aruoba et al. (2009). Our paper is also related to studies assessing the gains from considering a disaggregated approach for forecasting GDP. In this context, Heinisch and Scheufele (2018), among others, compare the forecasting accuracy of models forecasting aggregate GDP directly, as opposed to aggregating forecasts of GDP components. Similarly to our case, they consider the expenditure side and production side as a means to disaggregate GDP in its components. Their results favor the direct approach in forecasting GDP rather than a disaggregated approach which stands in contrast to our conclusion.

Our paper is closely related to Higgins (2014), Loscos et al. (2020) and Giovannelli et al. (2020). Loscos et al. (2020) specify dynamic factor models for GDP and several expenditure components for the purpose of forecasting. In line with our results, they also find an improvement of the GDP forecast when applying a disaggregated approach. Giovannelli et al. (2020) propose a method for nowcasting and forecasting sixteen main components of GDP along the production and expenditure side at a monthly frequency, using a high-dimensional set of monthly economic indicators. Their methodology relies on estimating all possible mixed frequency bivariate models of the quarterly GDP components and each monthly indicator. They, too, find that the indirect approach yields more accurate forecasts for GDP, confirming the results of Loscos et al. (2020) and ours too. While Loscos et al. (2020) and Giovannelli et al. (2020) rely on a pure statistical approach, our methodology explicitly takes economic theory into account and is extended by means of individual time series models. The possibly biggest advantage of our approach compared to Loscos et al. (2020) and Giovannelli et al. (2020) is its simplicity and transparency, which makes the proposed method attractive for applied work.

The outline of the paper is as follows. Section 2 introduces the cluster which defines the basic environment of the CDFM. It aligns with economic theory which gives rise to a structural element in the CDFM. Section 3 describes the individual models, their structure and the basic workings of the overall model. We pay particular attention to consistency. Section 4 provides an extensive model assessment. This concerns both the in-sample fit of each individual model and the CDFM's out-of sample predictive accuracy. Finally, Section 5 concludes.

2 The structure of the cluster

The starting point of our analysis is the System of National Accounts. It offers a coherent and consistent set of macroeconomic aggregates for the analysis of the economic structure. These are compiled quarterly according to the methods and definitions outlined in Eurostat (2013).

We use data for Austria, a small open economy, to illustrate the concept of the CDFM. Our set of data comprises the time series of GDP and various components taken from the production, income and expenditure side of quarterly National Accounts compiled by the Austrian National Statistical Agency. The series range from 1996:q1–2019:q4 at chained volumes. The series for the income account are available in nominal terms, i.e. at current prices, only.

Table 1 gives an overview on the variables of the three accounts used. In a first step we identify a structure within this set of variables. We use an agnostic approach, however, with a view to the central application of the model. To this purpose, we take the mutual forecasting contribution of the variables into account. And in this context, the concept of Granger-causality tests proves useful.

The Granger-causality test is a statistical hypothesis test for determining whether one time series is useful in forecasting another. We run bivariate Granger-causality tests among the year-over-year (y-o-y) growth rates of all variables listed in Table 1. A directional link from variable x_t to variable y_t is established if x_t Granger-causes y_t, but y_t does not Granger-cause x_t. The resulting link is thus unidirectional. It allows us to construct an ordered sequence, or a hierarchy of individual models. Nominal variables such as for instance aggregate labor and capital income from the income account are deflated by using the GDP deflator. We rely on the Bayesian information criterion (BIC) to determine the optimal lag lengths of the individual bivariate models involved in the hypothesis testing.

Table 1: National Accounts: Variable coverage

GDP (Production)	GDP (Expenditure)	GDP (Income)
Manufacturing VA (NACE B-E)	Private consumption	Labor income
Construction VA (NACE F)	Investment	Manufacturing
Services VA (NACE G-N)	Construction	Construction
	Equipment	Services
	Intangibles	Capital Income
	Exports	
	Goods	
	Services	
	Imports	
Residual	Residual	Residual

The Cluster of Dynamic Factor Models covers the production, expenditure and income side of the quarterly National Accounts, as well as employment and other monthly indicators.

In an attempt to identify a plausible structural cluster among the variables in Table 1, we supplement the Granger-causality tests with consideration motivated by economic theory.[2] For this we proceed in two steps: The first involves the identification of Granger-causal dependencies. Here we assign a level of statistical significance of ten percent within the Granger-causality tests. If a dependency between two variables, as identified by means of the Granger-causality, is at odds with the theory, we discard this linkage. While this intervention was not necessary in our case for the dependencies identified at the five

[2]The structure of the production side depicted in Table 1 has a very basic form. However, it could easily be extended to take into account country specific production characteristics. In case of a typical raw-material goods producing country, one could decompose the value added in manufacturing into further sub-categories (for instance crude-oil extraction in case of Russia and Saudi Arabia, copper production in case of Chile, etc.); the same applies to the value added in the service sector (financial sector activity in case of Luxembourg, tourism industry in case of Mediterranean or Caribbean countries, etc.). The flexible structure of the CDFM allows for extensions along various dimensions in this context.

percent level or higher, it becomes important at the ten percent level.[3] The results of this exercise are depicted as arrows in Figure 1. We show the results in the form of a cartographic representation to facilitate the analysis. The arrows replicate the bi-variate Granger-causality test results and hence identify the inter-linkages that are important to improving a forecast from a statistical point of view.

The results conform with common sense for a typical small open economy. Exports (of goods and services) are the key variable in driving the dynamics. Shocks therein immediately affect the manufacturing and service sectors, and this in turn provide incentives for entrepreneurs to expand on investment. Expansions in the value added of manufacturing raise the need for an expansion in the productive capacities, giving rise to higher equipment and construction investment. The increasing importance of digital elements in equipment investment motivates equipment investment as a means of stimulating intangible investment (intangible investment primarily involves spending on computer software, etc.). These inter-linkages identify a tight dependency between the variables of the production and expenditure accounts and shape the form and the dynamics of the cycles of these variables and hence those of the economy as a whole.

With a view on the variables of the income account, we observe that labor income is solely affected by variables of the production account (value added in manufacturing, construction and services). Capital income is in turn only significantly affected by exports, once more highlighting the superior role of exports in shaping the overall income path. Changes in both labor and capital income transmit to private consumption. While labor income only affects consumption, capital income also affects equipment investment. The linkage arises because, from an entrepreneur's perspective, capital income contributes to building up the equity base. This in turn reduces the level of indebtedness, which facilitates and favors borrowing, which ultimately stimulates investment. This link is statistically different from zero at the one percent level, highlighting the importance of capital income in shaping investment dynamics.

While we interpreted the arrows depicted in Figure 1 from the point of view of causation in the previous paragraphs, this is, however, only done for the purpose of explaining them within an economic context. The Granger-causality tests do not necessarily allow

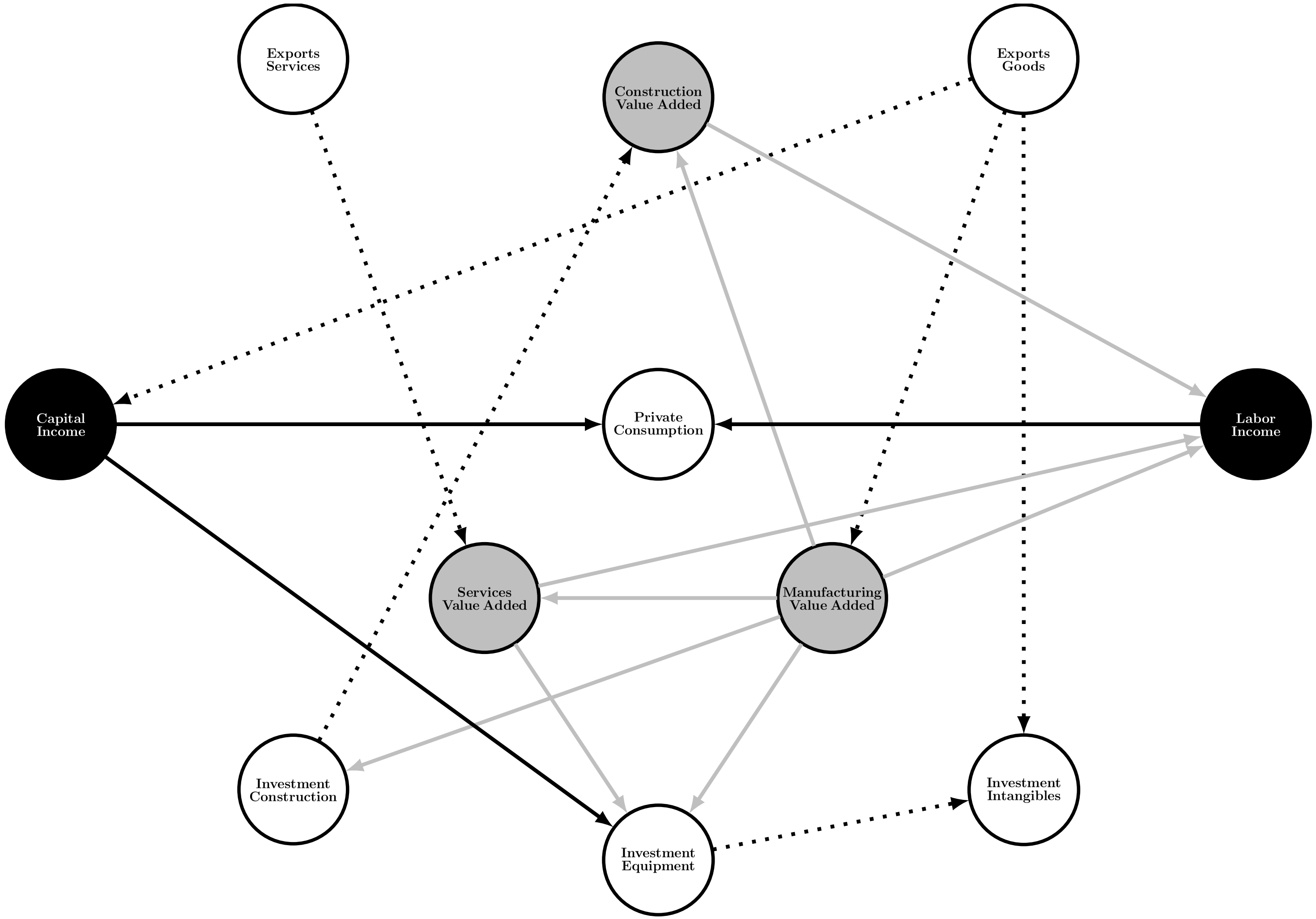
[3] We checked our final results for these omitted linakges, and we find that neither of these omissions has the potential to improve the forecast accuracy of the CDFM.

for that; instead, they only assess the variables' informational content for the purpose of forecasting (see Hamilton, 1994, for a more detailed discussion).

The structure of the production account proposed here is commonly considered by macroeconomic forecasters. While it comprises a reasonable approach from the point of view of a practitioner, this structure nevertheless relies only on the classification of the production account within the System of National Accounts, but it is not an approach proposed by economic theory. Against this background, we also assess the CDFM when using a different approach in modeling the production account. We do so by relying on a distinction between tradable and nontradable goods and services (TNT). Further details are provided in Section A in the Appendix.

In what follows, we incorporate the Granger-causal linkages within a series of small-scale dynamic factor models. We specify individual models for each variable depicted in Figure 1 and Table 1. The linkages constitute the structural element of the cluster, and the small-scale dynamic factor models comprise a reduced form element. The result thereof is a *structural* cluster of dynamic factor models. While the linkages depicted in Figure 1 allow only for a limited degree of dependency between the variables, this nevertheless proves sufficient for establishing model-consistent forecasts for a wide range of variables.

Figure 1: Granger-causal links



Note: Variables in white circles are from the expenditure account, the ones in black circles from the income account and the ones in grey circles from the production account.

3 The cluster of dynamic factor models (CDFM)

The basic element of the cluster is a dynamic factor model (DFM) that provides a parsimonious representation of macroeconomic data, because a small number of dynamic factors is sufficient to explain the majority of co-movements among macroeconomic data series (see Camacho and Pérez-Quirós, 2010, among others). The cluster of dynamic factor models (CDFM) comprises a series of dynamic factor models (DFMs), each estimated using the Kalman filter. Link variables across the DFMs connect the individual models to each other.

3.1 Some preliminaries

Let $\boldsymbol{x}_t = [x_{1,t}, ...x_{n,t}]^\intercal$, $t = 1, ..., T$ denote a set of standardized stationary monthly variables. Specifically, $\boldsymbol{x}_t$ will be a collection of monthly data taken either in levels or monthly year-on-year (y-o-y) growth rates. To incorporate quarterly y-o-y data (mostly from the National Accounts, e.g. GDP, investment, etc.) in this setting we construct a partially observable monthly y-o-y series and link it to the monthly variables by applying a modification of the approximation in Mariano and Murasawa (2003). In what follows, we adopt the convention that time indices for the quarterly variables refer to the third month of each quarter. Following Modugno et al. (2016), we consider quarterly level data for a given quarter to be the sum of monthly unobserved contributions. In particular, let X_t^q be a quarterly variable (in log-level), X_t^m its monthly (unobserved) counterpart, and let $x_t^{m,y}$ denote its unobserved monthly y-o-y growth rate. The monthly unobserved y-o-y growth rate can then be linked to a partially observed (at every third month of the quarter) quarterly y-o-y growth rate $x_t^{q,y}$ using the following:

$$\begin{aligned} x_t^{q,y} &= X_t^q - X_{t-12}^q \\ &= (1 - L^{12})X_t^q \\ &\approx (1 - L^{12})(1 - L - L^2)X_t^m \\ &= (1 - L - L^2)x_t^m \\ &= x_t^{m,y} + x_{t-1}^{m,y} + x_{t-2}^{m,y} \end{aligned} \tag{1}$$

where L is the lag operator. Viewing equation (1) as a factor model implies that quarterly variables should load equally on the current and lagged values of the unobserved monthly growth rate. We apply this set-up for all DFMs. More important, however, is the decomposition of vector $\boldsymbol{x}_t$ in terms of its role for the cluster. The cluster contains a series of DFMs of which each single DFM (i) addresses a specific target variable and (ii) establishes linkages with other DFMs.

The specification of a DFM for some target variable $x_t^{(j)}$ features an n dimensional vector $\boldsymbol{x}_t^{(j)}$ of observed monthly or quarterly time series which is partitioned as follows:

$$\boldsymbol{x}_t^{(j)} = \begin{bmatrix} x_t^{(j)} \\ \boldsymbol{x}_t^l \\ \boldsymbol{x}_t \end{bmatrix} \begin{array}{l} \text{– target variable} \\ \text{– link variables} \\ \text{– other variables} \end{array} \tag{2}$$

where vector $\boldsymbol{x}_t^l$ defines the link variables and $\boldsymbol{x}_t$ is a vector of other variables useful for the purpose of forecasting. Important for our application are the link variables $\boldsymbol{x}_t^l$, which the DFM for target variable $x_t^{(j)}$ receives from preceding DFMs. Inversely, this DFM also passes on variables to subsequent DFMs. We define the vector $\left\{\boldsymbol{x}_t^{*,(j)} = \left[x_t^{(j)}, (\tilde{\boldsymbol{x}}_t)^\intercal\right]^\intercal | t = 1, \dots, T;\ \tilde{\boldsymbol{x}}_t \subset \boldsymbol{x}_t\right\}$, of dimension p where $p \leq n$ to this purpose. It is important to note that $\boldsymbol{x}_t^*$ and $\boldsymbol{x}_t^l$ do not have any elements in common.

Our disaggregated modeling approach requires the aggregation of a series of variables at several points in order to account for identities of the System of National Accounts. To this purpose we distinguish between *behavioral* models and *aggregator* models. In what follows we address each of the two in more detail.

3.2 A behavioral DFM

We identify *behavioral* models to explain the dynamics of particular variables of interest. Consider again $x_t^{(j)}$ as a target variable. We specify a small-scale dynamic factor model using a q-dimensional, $0 < q < n$, vector of factors $\mathbf{f}_t$:

$$\boldsymbol{x}_t^{(j)} = \boldsymbol{\Lambda}(L)\mathbf{f}_t + \boldsymbol{D}(L)\boldsymbol{\epsilon}_t \tag{3}$$

$$(\boldsymbol{I} - \boldsymbol{\Phi}(L))\mathbf{f}_t = \boldsymbol{e}_t \tag{4}$$

where $\boldsymbol{\Lambda}(L)$ are $n \times q$ loading matrices which take into account equation (1). The common component $\boldsymbol{\Lambda}(L)\mathbf{f}_t$ and the idiosyncratic term $\boldsymbol{D}(L)\boldsymbol{\epsilon}_t$ are assumed to be uncorrelated, and, moreover, $\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma})$, such that $\boldsymbol{\Sigma}$ is an $n \times n$ diagonal covariance matrix. The matrix $\boldsymbol{\Phi}(L)$ is a lag-polynomial governing the dynamics of the latent factors in $\mathbf{f}_t$. The error term $\boldsymbol{e}_t$ of the dynamic equation satisfies: $\boldsymbol{e}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$. For identification reasons we impose that $\boldsymbol{\Sigma}_e$ is equal to the identity matrix.

We cast equations (3) and (4) into a state space format and estimate its parameters by running the Kalman filter. We set up the Kalman filter to deal with missing observations as discussed in Durbin and Koopman (2001).

A sequence of such models is estimated individually. We standardize each element in the vector $\boldsymbol{x}_t^{(j)}$. This is advisable because it reduces the number of estimated parameters (for instance no constant terms need to be estimated) and homogenizes variances of the idiosyncratic components which, in turn, allows for a significant acceleration of the estimation process of the models. Given the large number of behavioral models, this comprises an important aspect.

The specification of the behavioral models involves a decision on (i) which variables to include and in case a particular variable is included, (ii) along which temporal displacement (contemporaneous, lagged, etc.). We do so by relying on a combinatorial algorithm (Glocker and Wegmüller, 2020) for the selection procedure, where we consider the improvement of the out-of-sample forecasting accuracy of the target variable $x_t^{(j)}$ as the sole objective[4]. In the course of this selection process we find that for most target variables comparatively small models already have a forecast precision that cannot be further improved when including additional variables.[5] This results in a multitude of fairly small models. The big advantage here is that the estimation process of small models is fast in the context of a non-linear optimization routine, which is the case within the Kalman filter, and the variables can be adapted optimally with respect to their temporal displacement relative to the target variable.

The first column in Table 6 gives an overview of all the variables for which we specify a behavioral DFM. They are referred to as target variables and captured by $x_t^{(j)}$ in the vector in equations (2) and (3). The variables involve those mentioned in Figure 1, but for

[4]In order to accelerate the process of variable selection, we incorporate the findings from Heinisch and Scheufele (2018), Lehmann (2020), among others.

[5]As described in Boivin and Ng (2006) and Banbura and Rünstler (2011), the inclusion of additional variables, despite possible high correlation with the target variable, does not necessarily improve the forecast. When an additional variable is correlated with a subset of variables already in the model, the factors have a bias towards this subset of variables. As a consequence, the resulting factors explain a large fraction of the variation in each variable of this subgroup, but less of the variance in the target variable, rendering worse the overall model fit for the target variable and hence also its forecast.

several variables we consider a more disaggregated approach. This applies, for instance, to labor income, where we distinguish between labor income arising from the manufacturing, construction and service sector. The ordering of the variables follows a specific pattern: from mostly exogenous (i.e. foreign variables) to steadily more endogenous variables (pure domestic variables). This sequence replicates the empirically observed dependence structure and thus constitutes a key feature of the CDFM.

The second column provides a list for the link variables ($\boldsymbol{x}_t^l$) for each behavioral DFM. The link variables replicate the arrows depicted in Figure 1 and hence connect the individual DFMs to each other. We try to avoid circular dependency structures. In principle, our approach could take this into account, but this would increase computational complexity significantly, and this is beyond our scope. Hence, the link variables of any behavioral DFM in the list of Table 6 only involve variables of previously listed DFMs, though not of subsequent ones. This implies that the order of the DFMs matter in the CDFM.

Finally, the third column in Table 6 lists a series of additional variables. These are used to improve on the forecast of the target variable of the individual DFMs. Their selection as well as their temporal displacement is based on their contribution to improving the out-of-sample forecast, for which we rely upon the root-mean-squared-error (RMSE).

3.3 An aggregator DFM

While the behavioral DFMs identify appropriate reduced form models for the target variables for the purpose of forecasting, they, however, exclude truly aggregate variables. This arises from our modeling approach, which is highly disaggregated. In order to be able to establish forecasts for aggregate variables, we therefore construct *aggregator models*. Aggregator models essentially rely on identities. The idea is the following: Consider the computation of GDP along the production account. The summation of all sectors thereof yields the total value added and, once taking into account taxes and subsidies on products, we end up at GDP by definition. This is an identity for which we consider only a selected number of subcomponents within our approach. From this, two problems arise: first, modeling levels or growth rates, and second, how to deal with omitted components? We proceed in two steps. In the first, we consider a log-linearization of this identity. This gives us the GDP growth rate as a weighted average of the growth rates of each subcomponent. The weights are given by the share of each sector in GDP. This allows us to continue working with growth rates instead of levels. In the second step, we add a residual term. This term captures (i) changes in the shares of the sectors in GDP over time and (ii) collects all those subcomponents which are not specifically addressed by means of a behavioral DFM.

More formally, an aggregator model for y_t involves a weighted sum of the constituent component series. Let the series y_t comprise r components $x_t^{(i)}$ for $i = 1, \ldots, r$. An aggregator model for y_t reads:

$$y_t = \sum_{i=1}^{r} \omega_i x_t^{(i)} + \theta(L)\eta_t \tag{5}$$

$$(1 - \varphi(L))(\eta_t - \mu) = \epsilon_t \tag{6}$$

where $x_t^{(i)}$ $\forall i = 1, ..., r$ and y_t are non-standardized growth rates, $\theta(L)$ and $\varphi(L)$ are lag-polynomials, μ is a constant term and the error term ϵ_t satisfies $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. The weights ω_i $\forall i = 1, ..., r$ are fixed at their most recently observed values. We proceed as in Section 3.2 and cast equations (5) and (6) into a state space format and estimate its parameters by running the Kalman filter.

We proceed in this fashion to establish aggregate growth rates for GDP along the production, expenditure and income accounts. Further aggregator models are established for labor income, employment, investment, exports and imports. Table 7 provides an overview for each aggregator model and its subcomponents.

3.4 How the CDFM works

The key element for the working of the model is the concept of conditional forecasts. Conditional forecasting concerns forecasts of endogenous variables conditional on a certain path and length of path for some other endogenous variables. Specifically, it is assumed in our context that the conditioning information satisfies *hard conditions* (a particular path) rather than *soft conditions* (a range for the path). This stands in contrast to unconditional forecasts, where no knowledge of the future path of any variable is assumed.

The conditional forecasts of any DFM are passed on to another DFM within the cluster. To see how this works in practice, consider the following. For a small open economy, the most relevant shocks usually emerge from foreign demand. Therefore let us start with the DFM for goods exports (DFM (3)). This model can be used to produce either conditional or unconditional forecasts. In this context, a reasonable conditional forecast for goods exports could arise from assuming a certain future path for the GDP of the foreign economies. In case no future path for GDP for the foreign economies is assumed, the DFM for goods exports will establish an unconditional forecast for all the variables in the vector $\boldsymbol{x}_t^{(j)}$ of the DFM for goods exports. From this vector, the unconditional forecast of several variables is in turn used in yet other DFMs. The passing on is established by means of the link-variables ($\boldsymbol{x}_t^l$). This concerns, among others, the use of foreign PMIs for the model for service exports (DFM (4)). This linkages essentially guarantee consistency concerning the external environment surrounding the DFMs for the exports of goods and services. Furthermore, the unconditional forecast of goods exports is used in the DFM for the value added in the manufacturing sector (DFM (5)), capital income (DFM (12)) and intangible investment (DFM (16)). Once the forecast of goods exports is used in these models, conditional forecasts are established for the value added in the manufacturing sector and capital income. These conditional forecasts are in turn used in yet other DFMs.

This shows the role of conditional forecasts as a means of operationalizing the linkages within the CDFM. While the previous example started from an unconditional forecast for goods exports, the exercise would essentially be unchanged once assuming a certain future path of GDP for foreign economies, or these countries' PMIs, in order to establish a conditional forecast for goods exports and hence use this forecast in the subsequent DFMs. The prior knowledge, albeit imperfect, of the future evolution of some economic variables may carry information for the outlook of other variables. We assess this in more detail in Section 4.1.

In principle, the joint reliance on unconditional and conditional forecasts renders an assessment of the adequacy of our model difficult. For instance, one can immediately see why the failure of unconditional forecasts tells us very little about how good a model is at conditional forecasting. A macroeconomic model may be reasonably good at saying how a change in, for instance, oil prices will influence output, but it can still be pretty poor at predicting what output growth will be next year because it is bad at predicting oil prices in the first place. In this context, the CDFM offers the possibility of substituting imprecise unconditional forecasts of the CDFM for a specific variable with forecasts from outside. The CDFM then establishes forecasts for all variables in the model, conditional on the specific path assumed for this variable. The CDFM hence offers a flexible environment where inadequate forecasts can quickly be adjusted. This in turn also allows for scenario analysis in order to assess the sensitivity of forecasts to changes in the variables of some preceding DFMs.

3.5 Consistency

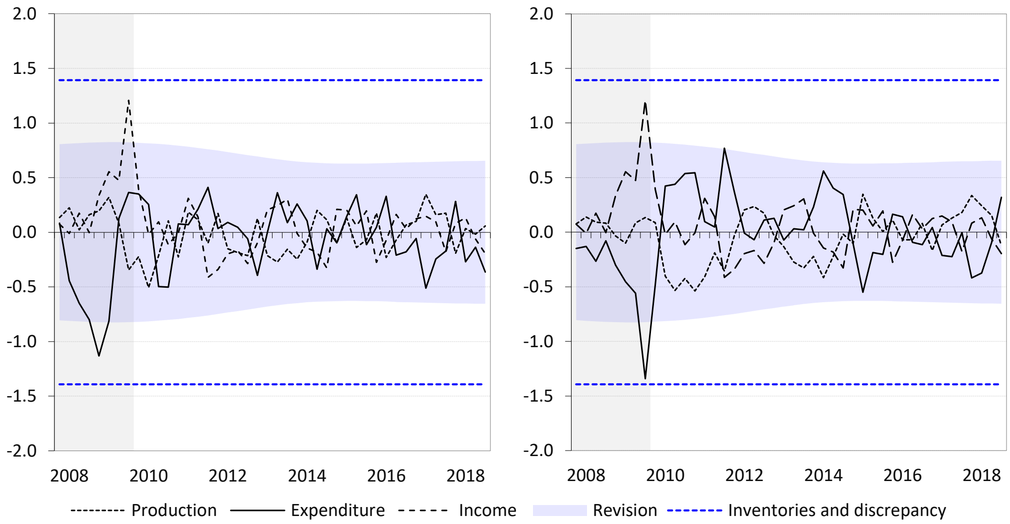
The CDFM computes forecasts for GDP along three dimensions. This seemingly independent three-fold approach might cast doubts on the predicted values once they diverge. Against this background, we assess the extent to which the GDP predictions arising from the production, expenditure and income accounts are consistent.

The CDFM interlinks variables from the three accounts to each other. Consider, for instance, investment as one particular expenditure variable. Its subcomponents (equipment investment, construction investment and intangible investment) are Granger-caused by variables from both the production account (manufacturing value added) and the income account (capital income). Variables from the expenditure account in turn also Granger-cause variables of the production and income accounts. A multitude of inherent linkages of this kind is likely to contribute to the overall model consistency as shocks in a particular variable are transmitted across all three accounts via these linkages. Since discrepancies between the aggregates can still arise, it is important to compare the discrepancies from the CDFM with the empirically observed counterparts.

We consider two empirically observed counterparts: (i) data revisions and (ii) inventory investment (we consider inventory investment[6] jointly with the official statistical discrepancy). As regards the first, the data of the National Accounts are continuously revised. This emerges from reconciliation of the three accounts. Revisions usually result in large changes in past growth rates. As regards the second, while for most countries the primary approach in computing aggregate GDP involves considering the production account, the expenditure and income accounts are adjusted accordingly. In principle, a GDP figure can be established from each of the three accounts. Since discrepancies are likely to arise, statistical authorities in general use the subcomponents to correct for them. These subcomponents are inventory investment for the expenditure account and capital income for the income account. In this respect, the CDFM closely mimics the approach and the difficulties statistical agencies face when establishing GDP data within

[6]This refers specifically to changes in inventory investment and acquisitions less disposals of valuables.

Figure 2: Internal consistency



Note: Internal consistency at 3m(1q) (left) and 12m(4q) (right). The values are in percentage points.

the System of National Accounts.

We compare the discrepancies of the CDFM for GDP from the three accounts (production, expenditure and income) to empirically observed revisions in the official data and differences in the GDP growth figures arising from inconsistencies across the three accounts. Figure 2 provides the results. The figure shows a band for the average revisions[7] and a measure for the empirically observed discrepancies in the growth rates across the production and expenditure account[8] depicted by the blue dashed line. In addition to this we show the one-quarter and four-quarter ahead forecast errors of the CDFM for GDP growth from all three accounts. As can be seen, the GDP growth rates from the CDFM as of the production, expenditure and income account differ, though the discrepancy is small compared to the average of the empirically observed discrepancy (around

[7]Average revisions are computed as the standard deviation across each quarter; this is possible since for each quarter, various estimates are available (flash estimate, first, second, ..., official figures; in our case for some quarters we have up to fifteen distinct figures for the quarterly growth rates). The blue band shows the (smoothed) average revisions and the change thereof over time. The extent of the data revisions depends, among other things, on the level of the growth rates. Although revisions are on average zero, their standard deviation is large compared to the mean of the GDP growth rate.

[8]We compute the GDP growth contribution of inventory investment and the statistical discrepancy using the final data vintage. We then calculate the standard deviation of the growth contributions. The value thereof is depicted by the blue dashed line in Figure 2. The standard deviation is shown in positive and negative territory in order to establish an interval.

1.5 percentage points over the horizon 2008–2018). This applies to both the one-quarter and four-quarters ahead forecasts. More importantly, however, is the observation that the discrepancies as implied by the CDFM are noticeably smaller than the discrepancies arising from the revision of official figures. The extent of official data revisions surpasses the discrepancies of the CDFM to a large extent, with only two exceptions in case of the one-quarter ahead forecast and three exceptions within the four-quarter horizon of the CDFM forecast.

When looking at the second measure of comparison – the standard deviation of the difference in actual growth rates – we find an even larger divergence in favor of the CDFM. The average gap between GDP growth as of the production and expenditure accounts from the official figures surpasses the discrepancies of the CDFM by a factor of up to four. We conclude that the GDP forecasts from the CDFM arising from the production, income and expenditure accounts differ to some extent. This difference, however, is small when compared to official figures from the National Accounts. We interpret this result in favor of a model inherent consistency.

In principle, our approach could easily be extended so as to have only one GDP measure across the three accounts. One possibility for this is by considering the average across the three accounts, which is something we will do later on. Another possibility is given by the error terms (η_t in equation (5)). Since all aggregator models for GDP include an error term, discrepancies could always be cast into these error terms in order to make sure that only one overall GDP measure emerges from the CDFM. This, however, requires the application of further statistical procedures as, for instance, described in van der Ploeg (1982, 1985). In this context, Loscos et al. (2020) establish overall consistency by imposing a balancing procedure which allows consistent forecasting of macroeconomic aggregates through an equilibrium model. Their approach is still distinct to ours as they consider a DFM directly applied for GDP and further DFMs for each demand component. Discrepancies therein arise because of a lack of consistency along the individual DFMs as they are not connected to each other. Within their approach, a balancing procedure is therefore essential.

4 Results

We start the assessment of the CDFM with an in-sample analysis. To this purpose we consider the behavioral DFMs only and use the R^2 as a measure of fit in this context. It is computed by regressing the factors of the jth behavioral DFM on the target variable $x_t^{(j)}$. The results thereof are depicted in the second column in Table 2.

Table 2: Factor correlation and NRMSE by core DFM (2007-2018)

		in-sample (R^2)	out-of-sample ($NRMSE$)			
DFM	Variable		3m(1q)	6m(2q)	9m(3q)	12m(4q)
(1)	Import deflator	0.77	0.44	0.86	1.14	1.29
(2)	Private consumption deflator	0.60	0.38	0.61	0.81	0.97
(3)	Export of goods	0.78	0.41	0.74	1.04	1.26
(4)	Export of services	0.57	0.57	0.59	0.56	0.63
(5)	Manufacturing VA	0.93	0.54	0.75	1.01	1.15
(6)	Investment construction	0.52	0.46	0.63	0.76	0.85
(7)	Construction VA	0.73	0.53	0.79	1.04	1.20
(8)	Services VA	0.52	0.44	0.64	0.79	0.94
(9)	Labor income manufacturing	0.83	0.38	0.59	0.82	1.04
(10)	Export deflator	0.97	0.43	0.85	1.15	1.31
(11)	Capital income	0.92	0.65	0.87	1.05	1.16
(12)	Labor income construction	0.75	0.55	0.62	0.69	0.75
(13)	Labor income services	0.87	0.38	0.58	0.79	0.96
(14)	Private consumption	0.56	0.75	0.88	0.95	0.98
(15)	Investment equipment	0.73	0.66	0.76	0.86	0.91
(16)	Investment intangibles	0.98	0.51	1.12	1.79	2.27

NRMSE refers to the normalized (standard deviation) root-mean-squared-error and hence ensures the comparability across all variables. The largest forecast errors are observed for consumption, capital income and equipment investment.

The in-sample fit ranges from a low of 0.52 in the case of the models for the value added of services and construction, up to 0.98 in case of intangible investment. Across all models the average value of the R^2 is around 0.80. This value conforms well with those of other studies. Glocker and Wegmüller (2020), for instance, identify a small-scale DFM for Switzerland; their preferred model has an R^2 of 0.74. In the specification of Camacho and Pérez-Quirós (2011) the factor explains a share of 0.8 of the variance of GDP growth. While these numbers seem sufficiently high for the models to produce good forecasts, this does not necessarily have to be the case. One reason for this is that a model with a good in-sample fit does not necessarily produce a precise out-of sample forecast (Clark, 2004; Granger and Jeon, 2004). The opposite also holds owing to the dynamic structure of the error term ($\boldsymbol{D}(L)\boldsymbol{\epsilon}_t$) that improves the forecast beyond the contribution of the factors $\boldsymbol{f}_t$. All this can been seen when considering the out of sample forecasts to which we turn next.

4.1 Out-of-sample analysis

The starting point of our out-of-sample exercise is the construction of a real-time data set, as originally proposed by Stark and Croushore (2002). This allows to assess the models' forecasting performance in real-time. Since we only use final data and not data vintages available to a researcher at a particular point in time, we consider this approach to be a pseudo-real-time analysis.[9]

We construct our real-time data set on monthly vintages. For each month within the 2006-2019 period we collect the whole set of time series available. We end up with 156 different vintages for the period 01/2006 to 12/2019. This data set allows us to closely mimic the forecasting procedure a practitioner would have performed at any point in time during the last few years when computing forecasts.

4.2 Predictive accuracy: a general view

We compute quarterly out-of-sample forecasts and compute root-mean-squared-error (RMSE) statistics. In order to establish a statistical measure that allows to compare the values across a wide range of variables, we normalize the RMSE by the standard deviation of the same variable. The resulting measure is referred to as a normalized-root-mean-squared-error (NRMSE). We provide the values thereof for the target variables specified within behavioral DFMs in the third to fifth columns in Table 2. Across all variables we find that the NRMSE increases with the horizon, which implies that the out-of-sample forecasting precision declines with the horizon. As regards the one-quarter ahead horizon, we find that the DFM for the consumption deflator (DFM (2)) and labor income in the manu-

[9] Although we have real-time data for GDP and various sub-components, the time series for this data are comparatively short, which limits the scope of a truly real-time analysis especially with regard to the cluster structure.

facturing sector (DFM (9)) and in the service sector (DFM (13)) yield the most precise forecasts. For the four-quarter ahead horizon, the DFM for service exports (DFM (4)) has the highest forecast accuracy. While this provides some information on the relative adequacy of each individual DFM, the comparison with the in-sample fit is, however, more interesting. In this context, the DFMs for service exports and intangible investment stand out. While the DFM for service exports has a rather low R^2, it at the same time has the highest forecast precision at the four-quarter horizon. We observe the opposite in the case of the DFM for intangible investment: with an R^2 of 0.98, this model has the highest in-sample fit; however, its four-quarter ahead forecast accuracy is the worst across all models. This once more highlights the commonly found observation that a model with a good in-sample fit does not necessarily produce accurate out-of-sample forecasts.

We provide the values of the NRSME for the variables which are captured by means of aggregator models in Table 3. As can be seen exports and labor income have a comparably high forecast accuracy at the one-quarter horizon. At the four-quarter horizon, the high precision of the model for labor income stands out.

Table 3 also provides NRMSE values for the GDP measures. Since the CDFM allows to establish GDP figures along the production, expenditure and income account jointly, we can hence assess the forecast accuracy for GDP from each of these three accounts. We add a fourth measure for GDP, which is simply the (unweighted) average across the three former measures. We find that the GDP forecast accuracy is highest with respect to the expenditure side, followed by the production side. The GDP forecast of along the production side could be noticeably improved when relying on the concept of tradable and nontradable goods and services (see Section A in the Appendix). We find that the GDP forecast from the income account has the lowest precision. This, however, only applies to the one-quarter horizon. Considering the four-quarter horizon, we find that the highest forecast precision for GDP now emerges from the income account. The average GDP measure performs reasonably well: worse than the expenditure approach, but better than the other two at a short horizon. For the four-quarter horizon, the average GDP forecast now outperforms the expenditure account.

We compare the forecast accuracy of the CDFM with a series of competing models. These are (i) a random-walk model, (ii) an AR(1) model, (ii) an optimal ARIMA model, and finally, (iv) a small-scale dynamic factor model directly applied to GDP. The optimal

ARIMA model is specified by relying on the Bayesian information criterion (BIC) for the lag-lengths of the autoregressive and moving-average lag-polynomials. Details on the specification of the small-scale dynamic factor model for GDP can be found in Section B in the Appendix.

Table 3: NRMSE by aggregator DFM (2007-2018)

DFM	Variable	3m(1q)	6m(2q)	9m(3q)	12m(4q)
(17)	Exports	0.36	0.65	0.91	1.14
(18)	Imports	0.52	0.79	1.04	1.26
(19)	Investment	0.85	0.90	1.02	1.15
(20)	Labor income	0.36	0.55	0.76	0.95
(21)	Employment	0.39	0.67	0.90	1.09
(22)	GDP deflator	0.57	0.73	0.76	0.76
(23)	GDP production	0.43	0.62	0.86	1.04
(24)	GDP expenditure	0.39	0.62	0.89	1.09
(25)	GDP income	0.50	0.70	0.90	1.03
	GDP average	0.41	0.61	0.86	1.05
			Competing models		
	GDP random walk	0.60	0.96	1.28	1.58
	GDP AR(1)	0.58	0.87	1.09	1.27
	GDP ARMA(2,1)	0.56	0.82	1.02	1.18
	GDP Small DFM	0.50	0.76	0.96	1.14

NRMSE refers to the normalized (standard deviation) root-mean-squared-error and hence ensures the comparability across all variables. The GDP average denotes the mean forecast of the three GDP National Account concepts. The competing models produce larger forecast errors than the CDFM throughout.

The out-of-sample forecasting performance of the competing models is also shown in in Table 3 where we again use the normalized RMSE (NRSME). As can be seen, the competing models have noticeably higher values for the NRMSE for both short and long forecasting horizons compared to the CDFM. The small DFM applied directly to GDP notably performs better than the rest of the competing models. Moreover, for a short horizon it performs as well as the GDP forecast of the CDFM arising from the income account.[10]

[10]Motivated by the findings in Döpke et al. (2019), we analyzed the forecast accuracy of the CDFM for the periods before and after the Great Recession. We find only small differences in the quantitative accuracy measures between the two periods. However, due to the short length of the respective time

4.3 Predictive accuracy: the importance of the linkages

Since the linkages between the individual DFMs are a fundamental feature of the CDFM, we will now briefly demonstrate their relevance in terms of the gain in the out-of-sample predictive accuracy for the GDP.

Tables 2 and 3 show the values of the NRMSE for different horizons based on the CDFM involving the linkages shown in Figure 1. We repeat this forecasting exercise based on a version of the CDFM, in which all of these linkages have been removed. This means that each individual DFM now produces forecasts independently to those of other DFMs. Therefore, for all those variables that are included in several DFMs, multiple forecasts are produced, which are all different to each other.

We compare the predictive accuracy of this alternative model with the CDFM. Table 4 relates the values of the NRMSE of the alternative model (no linkages) to those of Table 3. Values greater than unity imply that the CDFM's predictions are superior to those of the alternative model (no linkages). As can be seen, the values of the relative NRMSE are always larger than unity. The differences are considerable, implying up to 51 percent improvement in the forecast precision of the CDFM. We conclude that the CDFM's forecasts are superior to the alternative model where the linkages are ignored. This underscores the effect of the linkages on the predictive accuracy of the CDFM. The second purpose of including the linkages is ensuring the internal consistency of a model solution, which is particularly important for forecasting as well as simulations.

Table 4: Forecast error inflation without linkages (2007-2018)

Variable	3m(1q)	6m(2q)	9m(3q)	12m(4q)
GDP production	1.09	1.11	1.03	1.01
GDP expenditure	1.51	1.26	1.10	1.05
GDP income	1.08	1.14	1.11	1.07
GDP average	1.15	1.16	1.08	1.04

series for the forecast evaluation, these results should be interpreted with caution. Still, we also interpret this result in favor of our assumption for the time invariant variances in equations (4) and (6). Liu (2019) argued in this context that ignoring changes in macroeconomic volatility can lead to a biased estimation of macroeconomic tail fatness. The short sample size available for forecast evaluations, however, does not allow us to study this aspect in more detail.

4.4 Predictive accuracy: a closer look

While the results in Table 3 provide clear hints towards the predictive accuracy of distinct models, the question concerning whether some models are systematically better than others is, however, yet left unanswered. To this purpose we use the Diebold-Mariano test (Diebold and Mariano, 1995) for predictive accuracy. In particular, we use the modified Diebold-Mariano test of equal forecast accuracy according to Harvey et al. (1997).

With a view to the CDFM only, several results emerge: (i) the forecast of the CDFM along the production, expenditure and income accounts produce equal forecasts. For any horizon there is no evidence of a difference at any level of statistical significance. This once more undermines the consistency of the CDFM as already highlighted in Section 3.5, though with a different approach. (ii) The tradable-nontradable goods approach for modeling the production account as motivated in Section A of the Appendix also produces forecasts as good as the basic version of the CDFM. Again, there is no evidence of a difference for any level of statistical significance for any forecasting horizon. While this result shows that different approaches to modeling the production side yield fairly similar levels of forecast accuracy, it once more undermines the consistency inherent to the CDFM despite having three (four) measures for GDP.

When comparing the CDFM with the competing models, several further interesting results emerge: (iii) the CDFM forecasts are more accurate than those of the random walk and the AR(1) model for any horizon. This applies to the CDFM model in general and to its prediction for GDP along the production and expenditure accounts in particular. (iv) The predictions of the ARIMA model tend to be worse than those of the CDFM for any horizon, and the difference thereof is significantly different from zero only for short horizons (one-quarter and two-quarter ahead horizon). This observation undermines the usefulnes of the information from the monthly indicators, rendering the CDFM's forecast more precise at short horizons. (v) The CDFM's forecast from the production and the expenditure accounts tends to be superior to the forecast of the small-scale DFM applied to GDP directly. The difference in the precision of the forecast is statistically significant at the five-percent level for the one-quarter and two-quarter horizons. For higher horizons, the GDP forecasts of the CDFM are on average more precise, but the difference is not statistically different from zero. (vi) The small-scale DFM's forecasts are more accurate than those of the random-walk (all horizons) and the AR(1) and ARIMA models (one-quarter and two-quarter ahead horizon). This result replicates the findings in Camacho and Pérez-Quirós (2010), Camacho and Pérez-Quirós (2011) and Camacho and García-Serrador (2014), to mention a few.

4.5 Predictive accuracy: Comparison with expert forecasts

Having compared the forecast precision of the CDFM to various competing models, we extend the comparison to expert forecasts. The previous section relied on the Diebold-Mariano test for assessing the predictive accuracy of two distinct forecasts. While this test is commonly used, it faces significant drawbacks, especially when the length of the forecast series is short. A related test is the forecast-encompassing test. This test is used to determine whether one of the forecasts encompasses all the relevant information from the other. The resulting test statistic provides guidance as to whether to combine distinct forecasts or drop a particular forecast that contains no additional information.

To provide statistical evidence in terms of the predictive accuracy of the CDFM relative to expert forecasts, Table 9 presents the p-values of the forecast-encompassing test based on testing the significance of the parameter α_1 in the following OLS regression:

$$y_t - \hat{y}_t^{CDFM} = \alpha_0 + \alpha_1 \cdot \hat{y}_t^{PF} + \epsilon_t \tag{7}$$

where y_t is the realized GDP growth rate, and $\hat{y}_t^{CDFM}$ and $\hat{y}_t^{PF}$ are the real-time forecasts of the CDFM and of professional forecasters, respectively. To address the potential limitations of our short sample, we use robust standard errors as proposed by Bell and McCaffrey (2002). The encompassing test investigates whether there is still information contained in the CDFM's forecast error that can be explained by the professional forecasts.

The p-values indicate that the forecasts of the CDFM model encompass the forecasts from the professional forecasters. This holds across the CDFM's GDP forecast along the expenditure and income side, and the average GDP measure. As regards the GDP forecast from the production side, we find some evidence for improvement arising from the professional forecasts along the two-quarter horizon only. Despite this, we conclude that the CDFM's forecasts are hard to beat by any professional forecasts.

4.6 Forecast trajectory

In addition to investigating the CDFM's forecasting performance over a long sample, we also look into the model performance during specific historic episodes. In this context, we can assess the role of new, updated incoming information for the forecast. Our (pseudo) real-time data-set allows for such an assessment. We focus our analysis on the global financial and economic crisis of 2008/2009.

We compute real-time forecasts for the four quarters of 2009 to establish the annual growth rate of this year. We carry out this exercise with the information available at different points in time starting in January 2008. The path of the forecast trajectory is displayed in subplot (a) in Figure 3.

This figure helps us to address a question that has been the source of intensive debates in many countries: When did the authorities realize that the downturn had started? It is worth recalling that forecasting this turning point was a rather difficult task. The financial turmoil had increased the forecast uncertainty to unprecedented levels. In addition, at the beginning of the recession period, the financial variables and soft indicators were giving signals of a recession that were not associated with clear signals from real activity. Finally, for many countries it turned out to be the first negative annual growth after a long period of sustained growth. Figure 3 shows that signals of a business cycle turning point started to become clear around the summer of 2008.

For comparison purposes, forecasts from the competing models and the expert forecasts (IMF, OECD, EC, etc.) are shown in subplot (b) and (c). Each subplot is extended for the actual GDP growth values (blue solid line) jointly with a measure of average revisions of GDP growth (blue band; similarly to Figure 2). This figure displays several noticeable features which illustrate the advantages of real-time forecasting with the CDFM against alternative approaches. All forecasts display a declining path as the impact of the global downturn increasingly affected the economy. However, the CDFM's forecasts display the quickest downward adjustment. It anticipated negative growth rates for 2009 from August/September 2008 onwards, while most of the competing models' forecasts did so only from the end of 2008. In terms of detecting the point in time when growth for 2009 turned out to be negative, the small-scale DFM for GDP is more or less as good as the CDFM.[11]

[11] There are several alternative approaches for the analysis of turning points. Camacho et al. (2018)

With regard to the expert forecasts, it should be noted that the hesitant adjustment of the annual forecast in the form of excessive restraints as regards the publication of negative growth rates weighs heavily on the forecast error of the expert forecasts. This appears to be particularly pronounced at the end of the year: The CDFM forecast was already at -3.6 percent (average across the three accounts) and thus already very close to the realized value (-3.5 percent); the expert forecasts, on the other hand, were only slightly negative (-0.5 percent) and in some cases even positive.

Although a reluctance to publish negative growth rates may also result from the ambition not to spread excessive pessimism, the credibility of the expert forecasts is likely to suffer increasingly if such forecast errors occur repeatedly. This systematic bias does not apply to a model forecast though. The forecast is unbiased, objective and can be illustrated easily and transparently by means of the indicators used.[12]

5 Summary

In this paper, we propose a methodology for large-scale macroeconomic nowcasting and forecasting using a rich set of economic indicators available on a monthly basis. To this end, we use the concept of unidirectional Granger-causality to link a series of small-scale dynamic factors models in a cluster. Since the individual models are estimated using the Kalman filter, the resulting *Cluster of Dynamic Factor Models* can handle data at different frequencies, as well as data that feature gaps and ragged edges resulting from

use a two-state Markov-switching approach. They find that their recessionary state probability provides early guidance for the economy's current position in the course of the business cycle. Schreiber and Soldatenkova (2016) in turn use a subset of vector-autoregressive models (VARs) with automated zero restrictions to derive the probability of turning points in real time. Peláez (2015) specifies a model that forecasts all business cycle peaks and troughs for the US economy, out-of-sample, from 1970 to 2015. The predictive power is remarkably stable over time and yields a 100-percent proportion of correct recursive forecasts over this period. Finally, Hwang (2019) presents a recession forecast model and finds that time-varying logit models lead to large improvements in forecast performance, beating the individual best predictors as well as other popular alternative methods.

[12] An appealing extension in this context is motivated by the results put forward in Mathy and Stekler (2017). They show that business analysts were highly able to comprehend the situation along the quarters surrounding the global financial crisis: They use qualitative statements published in the financial press and convert them into quantitative information. Such quantitative information could be used in the CDFM to improve upon the forecast. We leave this venue open for future work.

asynchronous publication.

We find that the nowcasting and forecasting performance of the CDFM is superior to that of naive forecasting models and purely technical time series models that allow for rich dynamics. Moreover, the CDFM's forecasts for GDP are noticeably more accurate than those of classical small-scale dynamic factor models directly applied to GDP. Finally, the forecasts produced by the CDFM compete well with professional forecasts.

While the CDFM provides a transparent and highly flexible structure, it confers several further advantages. First, each individual DFM from the cluster can be used independently. This allows for producing both unconditional forecasts when used individually and conditional forecasts when used within the CDFM. Second, the CDFM allows forecasters to introduce their own judgment, which is integrated into a consistent conditional forecast. Third, the CDFM allows for scenario analysis in the context of forecasting on a large-scale, which renders our approach particularly useful for assessing the sensitivity of a macroeconomic forecast with respect to the underlying assumptions. Fourth, the CDFM not only allows to forecast real-time GDP, but also to incorporate information on the components of the expenditure, production and income accounts that determine the GDP forecast, providing broader insight into the causes of revisions in GDP forecasts. Finally, the CDFM has a simple structure that is easy to replicate. This makes the proposed modeling approach particularly attractive for applied work.

To summarize, we think that the *Cluster of Dynamic Factor Models* presented in this paper is a practical tool for nowcasting and short-term forecasting on a large-scale. It has a good forecasting record, is automatically updated when new information becomes available, provides a way of measuring the effects of new developments in GDP indicators and their subcomponents, and allows extensions to be implemented quickly due to its transparent and simple framework.

References

Aastveit, Knut, and Tørres Trovik (2012) 'Nowcasting Norwegian GDP: The role of asset prices in a small open economy.' *Empirical Economics* 42(1), 95–119

Angelini, Elena, Gonzalo Camba-Mendez, Domenico Giannone, Lucrezia Reichlin, and Gerhard Rünstler (2011) 'Short-term forecasts of euro area GDP growth.' *Econometrics Journal* 14(1), 25–44

Arnoštová, Kateřina, David Havrlant, Luboš Růžička, and Peter Tóth (2011) 'Short-term forecasting of Czech quarterly GDP using monthly indicators.' *Czech Journal of Economics and Finance* 61(6), 566–583

Aruoba, S. Boragan, Francis Diebold, and Chiara Scotti (2009) 'Real-time measurement of business conditions.' *Journal of Business & Economic Statistics* 27(4), 417–427

Bai, Jushan, and Serena Ng (2002) 'Determining the number of factors in approximate factor models.' *Econometrica* 70(1), 191–221

Banbura, Marta, and Gerhard Rünstler (2011) 'A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP.' *International Journal of Forecasting* 27(2), 333–346

Barhoumi, Karim, Gerhard Rünstler, Riccardo Cristadoro, Ard Den Reijer, Audrone Jakaitiene, Piotr Jelonek, António Rua, Karsten Ruth, Szilard Benk, and Christophe Van Nieuwenhuyze (2008) 'Short-term forecasting of GDP using large monthly datasets: A pseudo real-time forecast evaluation exercise.' Working papers 215, Banque de France, July

Bell, Robert M., and Daniel McCaffrey (2002) 'Bias reduction in standard errors for linear regression with multi-stage samples.' *Survey Methodology* 28, 169–181

Boivin, Jean, and Serena Ng (2006) 'Are more data always better for factor analysis?' *Journal of Econometrics* 132(1), 169–194

Camacho, Maximo, and Agustín García-Serrador (2014) 'The Euro-Sting revisited: The usefulness of financial indicators to obtain Euro Area GDP forecasts.' *Journal of Forecasting* 33(3), 186–197

Camacho, Maximo, and Gabriel Pérez-Quirós (2010) 'Introducing the euro-sting: Short-term indicator of euro area growth.' *Journal of Applied Econometrics* 25(4), 663–694

— (2011) 'Spain-sting: Spain short-term indicator of growth.' *Manchester School* 79(s1), 594–616

Camacho, Maximo, Gabriel Pérez-Quirós, and Pilar Poncela (2018) 'Markov-switching dynamic factor models in real time.' *International Journal of Forecasting* 34(4), 598–611

Clark, Todd (2004) 'Can out-of-sample forecast comparisons help prevent overfitting?' *Journal of Forecasting* 23(2), 115–139

Diebold, Francis X., and Roberto S. Mariano (1995) 'Comparing predictive accuracy.' *Journal of Business & Economic Statistics* 13(3), 253–263

Diron, Marie (2006) 'Short-term forecasts of euro area real GDP growth: An assessment of real-time performance based on vintage data.' Working Paper Series 622, European Central Bank, May

Döpke, Jörg, Ulrich Fritsche, and Karsten Müller (2019) 'Has macroeconomic forecasting changed after the Great Recession? Panel-based evidence on forecast accuracy and forecaster behavior from Germany.' *Journal of Macroeconomics* 62, 1–26

Durbin, James, and Siem Jan Koopman (2001) *Time Series Analysis by State Space Methods* (Oxford University Press)

Eurostat (2013) *Handbook on quarterly national accounts* (Luxembourg: Publications Office of the European Union)

Fildes, Robert, and Herman Stekler (2002) 'The state of macroeconomic forecasting.' *Journal of Macroeconomics* 24(4), 435–468

Friesenbichler, Klaus S., and Christian Glocker (2019) 'Tradability and productivity growth differentials across EU Member States.' *Structural Change and Economic Dynamics* 50, 1–13

Giannone, Domenico, Lucrezia Reichlin, and David Small (2008) 'Nowcasting: The real-time informational content of macroeconomic data.' *Journal of Monetary Economics* 55(4), 665–676

Giovannelli, Alessandro, Tommaso Proietti, Ambra Citton, Ottavio Ricchi, Cristian Tegami, and Cristina Tinti (2020) 'Nowcasting GDP and its components in a data-rich environment: The merits of the indirect approach.' CEIS Research Paper 489, Tor Vergata University, CEIS, May

Glocker, Christian, and Philipp Wegmüller (2020) 'Business cycle dating and forecasting with real-time Swiss GDP data.' *Empirical Economics* 58(1), 73–105

Glocker, Christian, and Serguei Kaniovski (2014) 'A financial market stress indicator for Austria.' *Empirica* 41(3), 481–504

Glocker, Christian, and Werner Hölzl (2019) 'Assessing the Economic Content of Direct and Indirect Business Uncertainty Measures.' WIFO Working Papers 576, Austrian Institute of Economic Research (WIFO), January

Granger, Clive, and Yongil Jeon (2004) 'Forecasting performance of information criteria with many macro series.' *Journal of Applied Statistics* 31(10), 1227–1240

Hamilton, James D. (1994) *Time Series Analysis* (Princeton University Press)

Harvey, David, Stephen Leybourne, and Paul Newbold (1997) 'Testing the equality of prediction mean squared errors.' *International Journal of Forecasting* 13(2), 281–291

Heinisch, Katja, and Rolf Scheufele (2018) 'Bottom-up or direct? Forecasting German GDP in a data-rich environment.' *Empirical Economics* 54(2), 705–745

Higgins, Patrick C. (2014) 'GDPNow: A model for GDP nowcasting.' FRB Atlanta Working Paper 2014-7, Federal Reserve Bank of Atlanta, July

Hwang, Youngjin (2019) 'Forecasting recessions with time-varying models.' *Journal of Macroeconomics* 62, 1–26

Lehmann, Robert (2020) 'Forecasting exports across Europe: What are the superior survey indicators?' *Empirical Economics, forthcoming*

Liu, Xiaochun (2019) 'On tail fatness of macroeconomic dynamics.' *Journal of Macroeconomics* 62, 103–154

Loscos, Ana Gómez, Mercedes de Luis López, Ana Arencibia Pareja, and Gabriel Pérez-Quirós (2020) 'A short-term forecasting model for the Spanish economy: GDP and its demand components.' *Economía* 43(85), 1–30

Marcellino, Massimiliano, and Christian Schumacher (2010) 'Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP.' *Oxford Bulletin of Economics and Statistics* 72(4), 518–550

Mariano, Roberto S., and Yasutomo Murasawa (2003) 'A new coincident index of business cycles based on monthly and quarterly series.' *Journal of Applied Econometrics* 18(4), 427–443

Mathy, Gabriel, and Herman Stekler (2017) 'Expectations and forecasting during the Great Depression: Real-time evidence from the business press.' *Journal of Macroeconomics* 53, 1–15

Modugno, Michele, Barış Soybilgen, and Ege Yazgan (2016) 'Nowcasting Turkish GDP and news decomposition.' *International Journal of Forecasting* 32(4), 1369–1384

Peláez, Rolando F. (2015) 'A biannual recession-forecasting model.' *Journal of Macroeconomics* 45, 384–393

Rünstler, Gerhard, and Franck Sedillot (2003) 'Short-term estimates of euro area real GDP by means of monthly data.' Working Paper Series 276, European Central Bank, September

Schreiber, Sven, and Natalia Soldatenkova (2016) 'Anticipating business-cycle turning points in real time using density forecasts from a VAR.' *Journal of Macroeconomics* 47, 166–187

Schumacher, Christian, and Jörg Breitung (2008) 'Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data.' *International Journal of Forecasting* 24(3), 386–398

Stark, Tom, and Dean Croushore (2002) 'Forecasting with a real-time data set for macroeconomists.' *Journal of Macroeconomics* 24(4), 507–531

van der Ploeg, Frederick (1982) 'Reliability and the adjustment of sequences of large economic accounting matrices.' *Journal of the Royal Statistical Society* 145(2), 169–194

— (1985) 'Econometrics and inconsistencies in the national accounts.' *Economic Modelling* 2(1), 8–16

A An alternative view on the production side: tradable and nontradable goods (TNT)

The approach proposed in Section 2 for modeling production is frequently used by international organizations, central banks and economic research institutes. It follows the sectoral composition of the National Accounts, but does not reflect the macroeconomic theory, which usually groups the sectors in producers of tradable goods (T) and nontradable goods (NT).[13] While the TNT classification is clear for some sectors, it can be ambiguous for others. Moreover, structural change might turn a previously tradable sector into a nontradable one, and vice versa. We follow the approach to the sectoral classification in Friesenbichler and Glocker (2019), determine the nominal value added for tradable goods and nontradable goods and calculate the corresponding deflators, which allows us to determine the real value added of these two categories.

The next step involves specification of separate behavioral models for tradable goods and nontradable goods, and an aggregator model for the GDP. The aggregator model features an error term, as the tradable and nontradable goods do not sum up to GDP, the difference being product taxes and subsidies. Another reason for including an error term is that we again consider a log-linearized representation of a weighted sum, in which the weights of the components can change over time. These changes in the weights are addressed by an autoregressive error term, as shown in Equation 5.

Table 5: Tradables (T) and nontradables (NT) (2007-2018)

Variable	3m(1q)	6m(2q)	9m(3q)	12m(4q)
Tradables VA	0.44	0.65	0.92	1.08
Nontradables VA	0.47	0.69	0.86	0.99
GDP average	0.41	0.61	0.86	1.05
GDP (TNT)	0.44	0.71	0.98	1.12
GDP Small DFM	0.50	0.76	0.96	1.14
Tradables deflator	0.64	0.89	1.05	1.10
Nontradables deflator	0.38	0.55	0.68	0.78
GDP deflator (TNT)	0.64	0.83	0.91	0.94
GDP deflator	0.57	0.73	0.76	0.76

Table 3 provides the normalized values of the RMSE for a GDP forecast based on the TNT approach. The forecasting accuracy of the TNT approach is similar to that of the standard production-side GDP approach considered in Section 3. Although the normalized RMSE of GDP is slightly smaller than in the conventional three-sector approach

[13] We use the term *goods* for both goods and services.

(manufacturing, construction, services), this difference is not statistically significant. The forecasts of the value added of tradable goods and nontradable goods seem to be comparatively precise. The values for the normalized RMSE for these two components are smaller than those for the sectors in the conventional approach for all forecasting horizons.

Subplot (d) in Figure 3 shows the annual GDP-forecast for the year 2009 obtained using the TNT approach over a period of 2008 and 2009. The TNT-based forecasts indicate negative annual growth relatively early, somewhat overestimating the extent of the recession at the end of 2008, but nonetheless approaching the realized value quickly.

Finally, the deflators for tradable and nontradable goods allow for an alternative approach to modeling and forecasting the GDP deflator. We specify a behavioral model for each of the two deflators and link them to the GDP deflator in an aggregator model. The forecast evaluation in Table 5 shows the normalized RMSE of the deflators for tradable and nontradable goods to be comparatively small. The TNT-based forecast of the GDP deflator are thus fairly precise and comparable to the baseline approach.

B The small-scale dynamic factor model as competing model

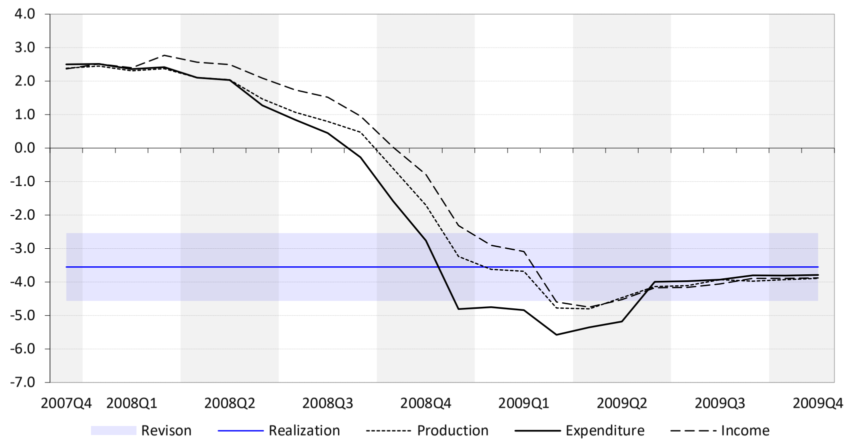
We consider a small-dynamic factor model as popularized by Mariano and Murasawa (2003); Camacho and Pérez-Quirós (2010, 2011); Arnoštová et al. (2011); Aastveit and Trovik (2012) as a competing model. This approach comprises a small-scale and hence simple factor model applied directly to GDP growth. Following Mariano and Murasawa (2003), we combine monthly and quarterly data, expressing the quarterly data as a function of monthly data. In particular, if the sample mean of the three monthly observations in a given quarter can be approximated by the geometric mean, then the quarterly growth rates can be decomposed as weighted averages of monthly growth rates. We follow the outline put forward in Section 3 and utilize the approach motivated by Glocker and Wegmüller (2020) to select an appropriate set of variables. This approach explicitly takes into account the fact that additional variables do not necessarily improve the model's forecast. The available set of variables contains around sixty variables.

The principal criterion for variable selection is out-of-sample forecasting ability, producing a set of variables geared towards economic expectations. We have already seen in Section 4.6 that the resulting model performs well in forecasting the 2009 economic downturn. The final specification of the small-scale factor model (SDFM) for GDP includes: (1) expectations in the construction sector, (2) expectations in the manufacturing sector, (3) expectations in the service sector, (4) Purchasing Managers Index (PMI), (6) order backlog (manufacturing sector), (7) employment (all sectors), (8) vacancies (all sectors), (9) retail sales (total) and (10) truck mileage. We add further variables only if they improve the out-of-sample forecasting performance of the model. We find that some additional variables could be included, however, they do not improve the forecast (e.g, Economic Sentiment Index (ESI) from the European Commission, ATX/Austrian Traded Index volatility, the financial market stress indicator as considered in Glocker and Kaniovski (2014), term-structure – i.e. the difference between 10-year and 2-year government bond yield, industrial production – excluding the construction sector, and retail sales). Other variables worsened the out-of-sample forecasts and were subsequently discarded from the model. The final selection proved robust to enlargements of the model in various directions. We tested our model using disaggregated versions of the variables already included in the model. For instance, we used retail sales without oil-related products instead of total retail sales. We also checked for the employment of different sectors (manufacturing sector, construction sector) instead of the aggregate measure, failing to improve the model in all cases.

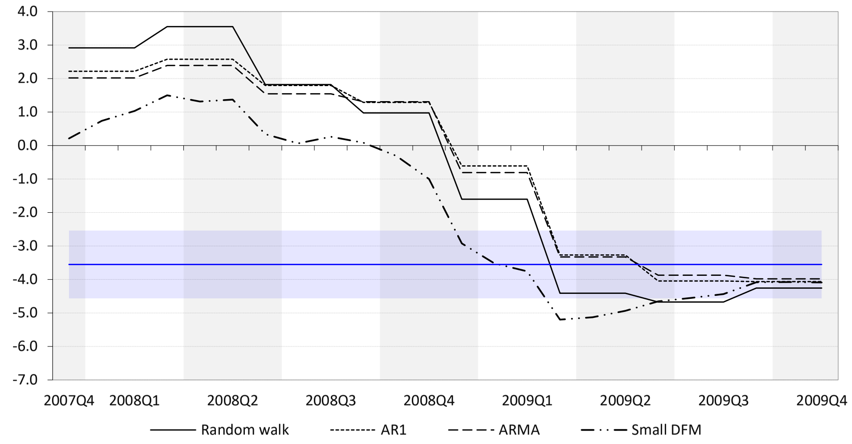
Since the variables considered address both the outlook and the current situation, we allow for a temporal displacement between GDP as the target variable and the additional variables, for which we follow Camacho and García-Serrador (2014). This set-up follows Camacho and Pérez-Quirós (2010) with a dynamic factor structure involving one factor with two lags (we omit the elements concerning data revisions from the model). The number of factors are selected by using Bai and Ng (2002) information criteria (BG) modified to take into account that the parameters are estimated using maximum likelihood. The number of lags in the factor equation and for the error terms was chosen by relying on the Bayesian Information Criterion (BIC).

C Additional figures and tables

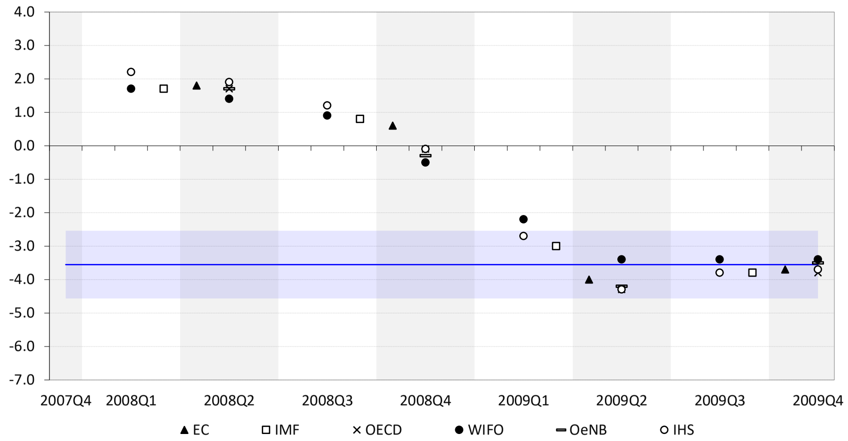
Figure 3: Forecasting performance 2009



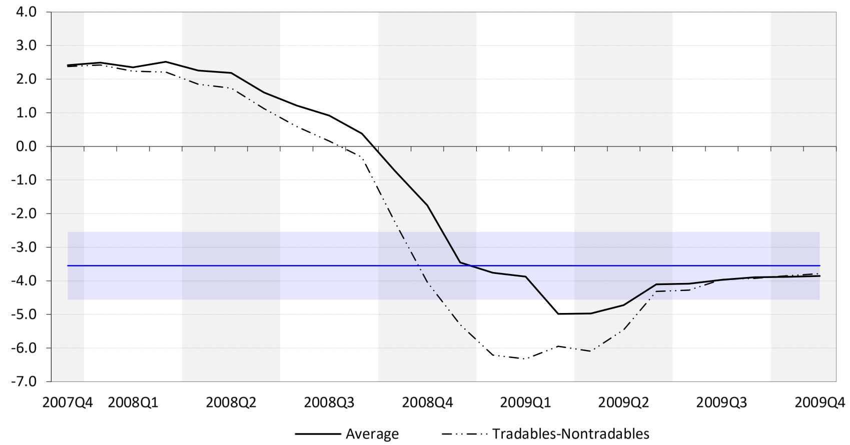
(a) Cluster of Dynamic Factor Models (CDFM)



(b) Competing models



(c) Expert forecasts



(d) CDFM with TNT extension

Note: The figure plots real-time forecasts of the CDFM (in sub-plot (a) and (d)), the competing models (sub-plot(b)) and expert forecasts (sub-plot(c)) for the annual GDP growth rate for the year 2009 established during 2008 and 2009. The values are in percent. The acronyms for the expert forecasts refer to European Commission (EC), International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), Austrian Institute of Economic Research (WIFO), Institute for Advanced Studies (IHS) and Austrian National Bank (OeNB).

Table 6: Ordering of behavioral models in the CDFM

DFM	Target variable $x_t^{(i)}$	Link variables $\boldsymbol{x}_t^l$	Other variables $\boldsymbol{x}_t$
(1)	Import deflator		Oil price in Euro (Brent) • EU PPI manufacturing
(2)	Consumption deflator	Import deflator (1)	CWI
(3)	Exports of goods		EU PMI • US GDP • EU GDP • Truck mileage
(4)	Exports of services	EU PMI (3) • EU GDP (3) • US GDP (3)	Air passenger volume
(5)	Manufacturing VA	Exports of goods (3) • Truck mileage (3)	DE manufacturing confidence • Manufacturing orders • Industrial production • Manufacturing employment • Manufacturing vacancies
(6)	Investment construction	Truck mileage (3) • Manufacturing VA (5)	Construction expectations
(7)	Construction VA	Truck mileage (3) • Manufacturing VA (5) • Investment construction (6)	Construction employment
(8)	Services VA	Truck mileage (3) • Exports of services (4) • Manufacturing VA (5)	Services employment • Services situation • Services expectations • Services vacancies
(9)	Labor income manufacturing	Manufacturing employment (5)	Manufacturing CWI • Manufacturing foreign orders
(10)	Export deflator	Import deflator (1) • EU GDP (3) • Manufacturing CWI (9) • Manufacturing foreign orders (9)	
(11)	Capital income	Exports of goods (3) • Exports of services (4)	Yield curve (2-10y) • Manufacturing uncertainty
(12)	Labor income construction	Construction expectations (6) • Construction employment (7) • Manufacturing CWI (9)	
(13)	Labor income services	Services employment (8) • Services situation (8) • Manufacturing CWI (9)	
(14)	Private consumption	Labor income manufacturing (9) • Capital income (11) • Labor income services (13)	
(15)	Investment equipment	Manufacturing VA (5) • Manufacturing orders (5) • Manufacturing vacancies (5) • Services VA (8) • Capital income (11)	Manufacturing situation
(16)	Investment intangibles	Exports of goods (3) • Investment equipment (15)	

The behavioral models are estimated sequentially by running the Kalman filter. The link variables are indicated by the number of the DFM model which their forecast is sourced from. *CWI* refers to the collective wage index that results from the wage bargaining process. *Manufacturing uncertainty* refers to a direct measure of uncertainty from the manufacturing sector explained in detail in Glocker and Hölzl (2019). Additional behavioral DFMs are estimated for (i) tradables VA, (ii) nontradables (VA), (iii) tradables deflator and (iv) nontradables deflator.

Table 7: Aggregator models in the CDFM

DFM	Variable y_t	Weighted target variables $x_t^{(i)}$
(17)	Exports	Exports of goods • Exports of services
(18)	Imports	Consumption • Investment construction • Investment equipment • Investment intangibles • Exports
(19)	Investment	Investment construction • Investment equipment • Investment intangibles
(20)	Labor income	Labor income manufacturing • Labor income construction • Labor income services
(21)	Employment	Employment manufacturing • Employment construction • Employment services
(22)	GDP deflator	Consumption deflator • Export deflator • Import deflator • Public sector CWI
(23)	GDP production	Manufacturing VA • Construction VA • Services VA
(24)	GDP expenditure	Consumption • Investment • Exports • Imports
(25)	GDP income	(GDP deflator) • Labor income • Capital income
	GDP TNT	Tradable goods VA • Nontradable goods VA
	GDP deflator (TNT)	Tradables deflator • Nontradables deflator

The GDP deflator is used to express nominal variables (labor and capital income) in real terms. *Public sector CWI* refers to the collective wage index that results from the wage bargaining process in the public sector. Details for the tradable-nontradable goods approach in modeling the production side are provided in Section A of the Appendix. The aggregator model for imports is based on the import content of the consumption investment and exports. The import content is calculated using the input-output tables.

Table 8: Diebold-Mariano test for GDP forecasts (2007-2018)

	Production	Expenditure	Income	TNT	Random walk	AR(1)	ARMA(2,1)
				3m(1q)			
Expenditure	<						
Income	>	>					
TNT	>	>	<				
Random walk	>**	>**	>*	>**			
AR(1)	>**	>**	>*	>*	<		
ARMA(2,1)	>**	>**	>	>	<	<	
Small DFM	>**	>*	<	>	<**	<	<
				6m(2q)			
Expenditure	<						
Income	>	>					
TNT	>	>	>				
Random walk	>**	>**	>**	>**			
AR(1)	>**	>*	>**	>	<		
ARMA(2,1)	>**	>	>**	>	<*	<**	
Small DFM	>**	>	>	>	<**	<	<
				9m(3q)			
Expenditure	>						
Income	>	>					
TNT	>	>	>				
Random walk	>**	>**	>**	>*			
AR(1)	>**	>	>**	>	<**		
ARMA(2,1)	>*	>	>	>	<**	<**	
Small DFM	>	>	>	<	<**	<	<
				12m(4q)			
Expenditure	>						
Income	>	<					
TNT	>	>	>				
Random walk	>**	>**	>**	>**			
AR(1)	>**	>	>**	>	<**		
ARMA(2,1)	>**	>	>*	>	<**	<**	
Small DFM	>	>	>	>	<**	<	<

The inequality sign compares a row against a column variable.
The notation *** 1 percent; ** 5 percent; * 10 percent level of significance of a two-sided Diebold-Mariano test.

Table 9: Forecast encompassing test (p values)

Forecast	3m(1q)	6m(2q)	9m(3q)	12m(4q)
		GDP production		
EC	0.374	0.059	0.188	0.700
IMF	0.442	0.024	0.160	0.500
OECD	0.287	0.022	0.072	0.754
WIFO	0.341	0.031	0.144	0.862
OeNB	0.413	0.009	0.110	0.886
		GDP expenditure		
EC	0.489	0.489	0.621	0.517
IMF	0.360	0.360	0.570	0.703
OECD	0.405	0.405	0.415	0.301
WIFO	0.552	0.552	0.578	0.369
OeNB	0.542	0.542	0.447	0.472
		GDP income		
EC	0.325	0.295	0.325	0.828
IMF	0.508	0.353	0.508	0.899
OECD	0.182	0.301	0.182	0.782
WIFO	0.321	0.400	0.321	0.583
OeNB	0.223	0.392	0.223	0.450
		GDP average		
EC	0.674	0.160	0.353	0.674
IMF	0.657	0.143	0.397	0.657
OECD	0.516	0.105	0.191	0.516
WIFO	0.739	0.267	0.375	0.739
OeNB	0.889	0.224	0.243	0.889

The p-values refer to the results of a t-test on the parameter α_1 of equation (7). The acronyms refer to European Commission (EC), International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), Austrian Institute of Economic Research (WIFO), and Austrian National Bank (OeNB).