# Nowcasting Swedish GDP Growth*

Sebastian Ankargren† and Unn Lindholm‡

February, 2021

* We thank Erika Färnstrand Damsgaard and seminar participants at the National Institute of Economic Research for comments that helped improve the paper. The opinions expressed in this article are the sole responsibility of the author and do not necessarily reflect the views of the National Institute of Economic Research.

† National Institute of Economic Research, Fleminggatan 7, Box 12090, 102 23 Stockholm, Sweden. E-mail: sebastian.ankargren@konj.se Phone: +46 453 59 55.

‡ National Institute of Economic Research, Fleminggatan 7, Box 12090, 102 23 Stockholm, Sweden. E-mail: unn.lindholm@konj.se Phone: +46 453 59 23.

WORKING PAPER NO 154, FEBRUARY 2021

PUBLISHED BY THE NATIONAL INSTITUTE OF ECONOMIC RESEARCH (NIER)

ISSN 1100-7818

**NIER** prepares analyses and forecasts of the Swedish and international economy and conducts related research. **NIER** is accountable to the Ministry of Finance, and like other government agencies, has an independent status and is responsible for the assessments that it publishes.

The **Working Paper** series consists of publications of research reports and other detailed analyses. The reports may concern macroeconomic issues related to the forecasts of the institute, research in environmental economics, or problems of economic and statistical methods.

All publications can be downloaded from the NIER website.

Abstract

We nowcast Swedish GDP growth using several types of popular short-term forecasting models. Our results indicate that medium-sized MIDAS regressions and small-scale bridge equation models provided the most accurate nowcasts in 2010Q1-2019Q4. Among dynamic factor models, we find a larger set of variables to be more appropriate, but the nowcasts made by our large dynamic factor model have historically been inferior to those made by the bridge equation and MIDAS regression models. Nevertheless, equal-weighted pooling of forecasts is superior to any single method. In a closer examination of nowcasting during the Covid-19 pandemic, we find that the dynamic factor model reacted much more forcefully during 2020Q2 and 2020Q3, with nowcasts that to a large degree developed like professional forecasts. Our results reveal a clear divide between, on the one hand, historical forecasting performance in the period between the Great Recession and the Covid-19 pandemic, and, on the other hand, usefulness during the pandemic. Decomposing the revisions of the dynamic factor model's nowcasts into contributions, we find that updated parameters caused large revisions. In comparison with a model that is not re-estimated during the pandemic, however, the re-estimated model's nowcasts are more reasonable and accurate. Incorporating new data sources that measure economic activity at higher frequency does not improve forecasting accuracy historically, but amplifies the downturn signal during the peak of the pandemic.

**JEL classification code**: C11, C22, C32, C53, E17

Keywords: Bayesian, dynamic factor model, bridge equation, mixed-frequency, MIDAS, Covid-19

# Sammanfattning

Vi gör nulägesprognoser för svensk BNP-tillväxt med flera typer av populära kortfristiga prognosmodeller. Våra resultat tyder på att medelstora MIDAS-regressioner och småskaliga överbryggnadsekvationer gav de mest exakta nulägesprognoserna under 2010kv1-2019kv4. Bland dynamiska faktormodeller finner vi att en större uppsättning variabler är mer lämpligt, men de nulägesprognoser som gjorts av vår stora dynamiska faktormodell har historiskt sett varit sämre än de som gjordes av överbryggnadsekvationer och MIDAS-regressionsmodellerna. Icke desto mindre är en lika viktad sammanslagning av prognoser överlägsen en enskild metod. I en närmare granskning av nulägesprognoser under covid-19-pandemin finner vi att den dynamiska faktormodellen reagerade mycket kraftigare under 2020kv2 och 2020kv3, med nulägesprognoser som till stor del utvecklats som prognoser gjorda av professionella prognosmakare. Våra resultat visar en tydlig skillnad mellan å ena sidan historisk prognosutveckling under perioden mellan finanskrisen och covid-19-pandemin och å andra sidan användbarheten under pandemin. Genom att dekomponera revideringarna av den dynamiska faktormodellens nulägesprognos till bidrag finner vi att uppdaterade parametrar orsakade stora förändringar i prognoserna. Jämfört med en modell som inte skattas om under pandemin är dock den omskattade modellens nulägesprognoser mer rimliga och korrekta. Att införliva nya datakällor som mäter ekonomisk aktivitet på högre frekvens förbättrar inte prognosförmågan historiskt, men förstärker nedgången under pandemins topp.

# Contents

# 1 Introduction

An accurate assessment of the current situation is crucial for predicting the future. Forecasting the present is commonly referred to as nowcasting, and is particularly common for gauging economic output as measured by the gross domestic product (GDP). Many professional forecasters, both public and private, engage in nowcasting with well-known examples including the Federal Reserve Banks of New York (Bok et al., 2017) and Atlanta (Higgins, 2014), the European Central Bank (Bańbura and Saiz, 2020), and the OECD (Ollivaud et al., 2016). For a detailed discussion of the challenges involved in nowcasting, see Bańbura et al. (2013).

The real-time macroeconomic data flow is characterized by a staggered nature that yields so-called ragged edges—that is, unbalanced data, in particular at the end of the sample. Survey data is typically faster than other types of data, and published at the end of the reference month. Other important economic data, such as inflation and unemployment, are often made available around two to three weeks following the end of the reference month. Some data that are crucial for judging the state of the economy, like industrial production, retail sales, and household consumption, can often be associated with even greater publication lags of four to five weeks or more. The differences in publication structure means that models that include more data than only a handful of indicators can oftentimes be updated several times a week to account for new releases. Developing models that can efficiently incorporate the continuous flow of new data to improve the short-term forecasts is the core challenge that nowcasting tackles.

We study the Swedish economy and nowcasts of quarterly GDP growth in this work. Our contribution extends beyond the development of a suite of forecasting models for Sweden as we investigate the role of several key aspects when conducting model-based nowcasting. We compare the forecasting ability of three popular classes of nowcasting models: dynamic factor models, bridge equations, and MIDAS regressions. We further assess the role of model size by estimating the models on small (6 variables), medium (13 variables) and large (34 variables) sets of data. Since Sweden is a small and open economy, we also investigate the role of foreign variables in nowcasting domestic output growth. While our data comprises conventional sources of soft and hard data, the Covid-19 pandemic emphasized an urgent need for faster data, and ultimately led to several new indicators of economic activity being developed. We evaluate the use of some of these data in predicting the pandemic-induced downturn and subsequent recovery.

Our forecast evaluation shows that MIDAS regressions and bridge equation models achieved lower root mean squared errors than the dynamic factor model. The results also highlight one important aspect when comparing forecasting methodologies: the three methods we consider all favor different model sizes. The bridge equation model using our small set of variables does better than using the medium or large sets of variables, whereas we obtain the best performance among the MIDAS regressions when we use the medium-sized set of variables. In contrast, the dynamic factor model shows more satisfactory performance when employing the large set of variables.

To better understand how the results are affected by model specifications, we consider and evaluate some alternatives. We mostly obtain marginal differences in predictive ability. The dynamic factor model does not benefit from considering more or fewer international variables, nor does increases or decreases in the number of lags lead to more satisfactory forecasting performances. The bridge equation model does not improve from modeling monthly variables using a Bayesian VAR in lieu of independent AR(4) regressions, and neither bridge nor MIDAS show any major improvements when dropping lagged GDP from the equations. In our baseline specification, we use the unrestricted MIDAS approach suggested by Foroni et al. (2015) with four lags, but changing to an exponential Almon specification with 12 lags leads to negligible differences that is both better and worse than baseline depending on the horizon. Finally, we examine if more sophisticated forecast pooling methods can improve upon the simple unweighted average. We obtain no better nowcasts using weights that depend on historical accuracy.

We also investigate the performance of our suite of nowcasting models during the downturn and subsequent recovery in 2020Q2 and 2020Q3. We compute nowcasts every Friday from the beginning of March and onwards, and compare if new indicators for real economic activity in Sweden and the United States published at higher frequency made a difference for the nowcasts. Our results show that inclusion of the Weekly Economic Index for the US, released by the Federal Reserve Bank of New York, helped provide a stronger signal of severe economic effects than what was found by the benchmark dynamic factor model. Qualitatively, all versions of the dynamic factor model, with and without selected new indicators, essentially convey the same story during 2020Q2. Nowcasts were successively revised downwards in March and early April until reaching their minimum, and then being revised upwards as more signals of optimism were identified in the data. To offer some context, we contrast the model-based nowcasts with the professional forecasts published by the National Institute of Economic Research (NIER), one of Sweden's most prominent forecasting institutes. The evolution of the model-based nowcasts largely corresponds to that communicated by the NIER, whose view of 2020Q2 was considerably more negative by the end of April than in early April and June. For 2020Q3, the models' nowcasts were close to the view held by the National Institute of Economic Research in its August and September forecasts, and with only minor differences between versions of the dynamic factor model. The bridge equation and MIDAS regression models—showing the lowest RMSE in our forecast evaluation—both suffer from sluggishness, and do not at all react in the same way to new data in 2020Q2 and 2020Q3.

Forecasting results in the literature vary substantially depending on the scope of the forecasting exercise, particularly with respect to country and sample period. Andersson and den Reijer (2015) develop a dynamic factor model and bridge equation models for nowcasting Swedish GDP growth, and is thus closely related to our work. Similarly, den Reijer and Johansson (2019) evaluate various specifications and find that averaging predictions from MIDAS regressions with indicators included one at a time—the same approach we take—performs particularly well. Our result that simple bridge equation models do better or equally as well as larger and more complex dynamic factor models corroborates the results presented by Ollivaud et al. (2016). They argue that simple bridge equation models are often as good as large and more complex dynamic factor models, and demonstrate for six advanced economies that dynamic factor models offer no gains in forecast accuracy for short-term forecasts. We do, however, find a relatively more satisfactory performance by the more complex dynamic factor model in 2020Q2 and 2020Q3. This result can be viewed in light of work by, for example, Chauvet and Potter (2013), who show that improvements in accuracy achieved by more sophisticated models for the US are largely driven by crisis periods, and that during non-crisis periods improvements are instead at best modest in comparison with simple benchmarks. Siliverstovs (2020) obtain the same result for nowcasting models for the US, and our results, clearly favoring the dynamic factor model during the pandemic-affected second and third quarters of 2020 but simpler models otherwise, thus match previous findings in the literature.

Our examination of nowcasts during the Covid-19 pandemic also add to a growing body of literature on nowcasting during the pandemic; see, for example, Antolín-Díaz et al. (2020); Cimadomo et al. (2020); Foroni et al. (2020); Lenza and Primiceri (2020); Schorfheide and Song (2020). Antolín-Díaz et al. (2020) and Lenza and Primiceri (2020) both propose ways of adapting model specifications to accommodate the large economic shocks caused by the pandemic, Antolín-Díaz et al. (2020) suggesting a dynamic factor model that allows for outliers, and Lenza and Primiceri (2020) detailing how a Bayesian VAR can be estimated when the timing of a sudden increase in shock variances and covariances is known. Similarly, Schorfheide and Song (2020) show how a few months of pandemic-affected data increases estimated shock variances and covariances, instead favoring estimating their model on pre-pandemic data so that the distribution of shocks in the recovery period is the same as in the pre-pandemic era. Besides showing how our models' nowcasts developed during the pandemic, we also decompose the revisions to the nowcasts and find occasional large contributions from changes in parameters. Accounting for abnormal observations by either employing an outlier-adjusted model specification as used by Antolín-Díaz et al. (2020) or by letting the shock covariance matrix increase at the beginning of the pandemic as suggested by Lenza and Primiceri (2020) would effectively put less weight on data observed during the pandemic, and would lead to smaller changes in estimated parameters. However, when we compare nowcasts using a re-estimated model and a model estimated before the pandemic, an approach strongly resembling that taken by Schorfheide and Song (2020), we find that the re-estimated model yields much better and realistic nowcasts, in spite of some of its forecast revisions largely being attributed to changes in parameters.

The remainder of the paper is structured as follows. Section 2 gives brief descriptions of the forecasting methods that we use, Section 3 provides a description of the data, Section 4 presents the results of forecast evaluations, and Section 5 presents an analysis of nowcasts during the Covid-19 pandemic. Section 6 concludes.

# 2 Methods

In this section, we provide brief descriptions of the methods that we use.

## 2.1 Dynamic Factor Model

The (dynamic) factor model is one of the more popular methods for nowcasting, originally proposed by Geweke (1977). Further developments that have been crucial for its applicability in nowcasting include Mariano and Murasawa (2003); Giannone et al. (2008); Bańbura and Modugno (2014) with many additional studies providing empirical evidence of its usefulness. A key feature is that the model can decompose its forecast revisions into one of three sources:

parameter revisions, historical data revisions, and new data releases. The decomposition provides an opportunity to interpret changes to the model's forecasts in a meaningful way, and is useful for monitoring the economy in real time.

Basic specifications of the factor model let the variables load on the factors contemporaneously with idiosyncratic autocorrelation structures allowed by letting the error terms follow simple autoregressive processes. Such simple specifications are often effective and do well as evidenced by, for example, the factor model used by the Federal Reserve Bank of New York (Bok et al., 2017). However, lead-lag relationships between variables themselves and with the business cycle in general can be complex, with differences in the timing of downturns between, for instance, consumption, production and consumer confidence. To account for heterogeneous lead-lag relationships, we follow the work by Luciani and Ricci (2014); D'Agostino et al. (2016); Antolín-Díaz et al. (2020) and let the lag structure be more flexible than a contemporaneous impact of the factor on all variables. The model accommodates not only immediate effects of the business cycle factor on the variables, but also delayed effects through lags of the factor.

The model contains a single factor and is specified as

$$\begin{aligned} x_{i,t} &= \sum_{j=0}^{r} \lambda_{i,j} f_{t-j} + \sum_{k=1}^{p} \phi_{i,k} x_{i,t-k} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim \text{iid}\, \mathrm{N}(0, \sigma_i^2) \\ f_t &= \sum_{j=1}^{q} \varphi_j f_{t-j} + u_t, \quad u_t \sim \text{iid}\, \mathrm{N}(0,1). \end{aligned} \tag{1}$$

Lead-lag relationships that are different across equations are permitted because of the inclusion of $r$ lags of the factor. We let $x_{i,t}$ represent the underlying process that governs the system, and $y_{i,t}$ be the data that is observed. Whenever a monthly variable is observed, then $y_{i,t} = x_{i,t}$, but in cases where it is unobserved then $y_{i,t} = \varnothing$. Similarly, quarterly GDP is observed only four times a year. We follow Mariano and Murasawa (2003) and relate the quarterly observations to the monthly underlying series through

$$y_{i,t} = \begin{cases} \frac{1}{3}(x_{i,t} + 2x_{i,t-1} + 3x_{i,t-2} + 2x_{i,t-3} + x_{i,t-4}), & \text{if } y_{i,t} \text{ is observed} \\ \varnothing, & \text{otherwise.} \end{cases}$$

In our benchmark specification, we set $r = 3$ and $p = q = 1$ to permit for asynchronous relationships between factors and variables, but still keep the specification parsimonious. More details on the estimation of the model can be found in Appendix A.

## 2.2 Bridge Equations

Bridge equations have an extensive track record in nowcasting at major economic institutions, having been used by, e.g., the ECB (ECB, 2008), Deutsche Bundesbank (Bundesbank, 2013), the OECD (Ollivaud et al., 2016), and Sveriges Riksbank (Andersson and den Reijer, 2015). The approach uses simple models to forecast high-frequency series in a first step, and then aggregates the high-frequency series to the lower frequency at which the target variable is observed. The final step predicts the target using the aggregated high-frequency series, which now partially consist of observed data and forecasts, as predictors.

Let $y_t$ be quarterly GDP growth, and assume that it is observed for $t = 3, 6, 9, \dots, T_y$. Let $x_t$ be a monthly indicator observed for $t = 1, 2, \dots, T_x$, and let $x_t^Q$ represent the monthly variable aggregated to the quarterly frequency. Depending on the nature of the variable, $x_t^Q = \frac{1}{3}(x_t + x_{t-1} + x_{t-2})$ or $x_t^Q = x_t$ for $t = 3, 6, \dots$. The main forecasting equation in our bridge equation model is an autoregressive distributed lag model of the form

$$y_t = \phi_0 + \phi_1 y_{t-3} + \beta_1 x_t^Q + \beta_2 x_{t-3}^Q + \epsilon_t. \tag{2}$$

However, the caveat is that $x_t^Q$ must be observed to forecast $y_t$. If $x_t^Q$ is unobserved, we first employ an auxiliary AR(4) model that extends the monthly predictor into the future. That is, we estimate

$$x_t = \psi_0 + \sum_{i=1}^{4} \psi_i x_{t-i} + \nu_t, \quad t = 1, 2, \dots, T_x,$$

and compute the forecast $\hat{x}_t$ for $t = T_x + 1, \dots, T_y + H$, where $H$ refers to the maximal forecast horizon for GDP growth. We then proceed and compute the quarterly aggregates $x_t^Q$ using forecasts $\hat{x}_t$ when necessary, whereby (2) is feasible to use for recursively forecasting $y_{T_y+h}$ for $h = 3, 6, \dots, H$. We aggregate the predictions obtained from each indicator to a single number by taking the average of the predictions.

### 2.3 MIDAS Regressions

The final class of models that we consider is MIDAS (mixed data sampling) regressions (Ghysels et al., 2007). The model relates quarterly GDP growth at time $t$ to its value in the previous quarter, and to monthly lags of $x$ according to:

$$y_t = \phi_0 + \phi_1 y_{t-3} + \sum_{j=0}^{p} \beta_j x_{t-j} + \xi_t.$$

MIDAS regressions can suffer from parameter proliferation problems, in particular when the frequency mismatch is large. For this reason, the parameters $\beta_j$ are often approximated by a low-dimensional polynomial, whereby the number of parameters to estimate is reduced. However, we use only $p = 4$ lags and instead use the UMIDAS approach suggested by Foroni et al. (2015), who advocate estimating the parameters freely if the frequency discrepancy is low. We estimate one MIDAS regression per explanatory variable, and aggregate the individual forecasts by averaging the individual forecasts. Pooling single-indicator models as a way of handling large sets of indicators is a strategy with demonstrated success in the literature, see, for example, Kuzin et al. (2013) for multi-country evidence and den Reijer and Johansson (2019) for a successful application to Swedish data.

## 3 Data

Over time, common practice in nowcasting has shifted from including a large number of disaggregate series to using a more modest amount of aggregate variables. Table 1 contains

Table 1: Summary of Data Used in Factor Models

| Country | Reference | Data |
|---|---|---|
| Euro Area | Angelini et al. (2008) | 85 variables |
| Germany | Bundesbank (2013) | >100 variables |
| Sweden | Andersson and den Reijer (2015) | 126 variables |
| UK | Anesti et al. (2017) | 60 variables |
| USA | Bok et al. (2017) | 32 domestic |
| Canada | Chernis and Sekkel (2017) | 17 domestic, 6 foreign |
| Switzerland | Galli et al. (2019) | 536 domestic, 84 foreign |
| Euro Area | Bańbura and Saiz (2020) | 30 variables |

a non-exhaustive list of selected nowcasting models that are publicly described by major institutions. Around the beginning of the 2010s, nowcasting models used by the ECB, Deutsche Bundesbank and Sveriges Riksbank included a large number of disaggregate time series (ECB, 2008; Bundesbank, 2013; Andersson and den Reijer, 2015). More recently developed models at the Federal Reserve Bank of New York (Bok et al., 2017), Bank of Canada (Chernis and Sekkel, 2017) and the ECB (Bańbura and Saiz, 2020) instead leverage a much smaller amount of data. Galli et al. (2019) use a very large dynamic factor model for nowcasting Swiss GDP growth, saying that their going against the current tendency of medium-sized models can be explained by a lack of hard data that are timely available and that have well-established predictive power.

We take note of the evolvement of practice to primarily focus on aggregates and use a total of 34 time series. The variables we select reflect several aspects of the economy, from labor market developments to consumption and production. We include variables that are routinely monitored by Swedish forecasters, and variables that often make the news because they carry important information on the state of the economy. Because our interest lies in exploiting historical empirical relationships to nowcast GDP, we are agnostic with respect to theory and include variables that could potentially correlate with GDP even if theoretically guided causality is absent. The variables are listed in Table 2.

The table displays the variables split into three blocks: a small, core set of variables, a medium-sized augmented set, and a large set that comprises all of the listed variables. The division into smaller sets of core variables is based on balancing timeliness of data releases with experience of what indicators are useful predictors for GDP growth. The medium and large blocks include the variables listed in the smaller blocks, and so the total number of variables for each block size is 6, 13, and 34 variables. All time series, except GDP, enter our models as monthly variables.

# 4 Forecast Evaluation

We conduct a pseudo-real-time forecast evaluation to assess the models' forecasting ability. The evaluation period is 2010Q1–2019Q4, and we make out-of-sample forecasts on the 4th and 18th day of each month. At each forecast origin, we replicate the release pattern of data

Table 2: Summary of Data. The medium block includes variables listed in the small block, and the large block includes variables listed in both the medium and small blocks.

| | Mnemonic | Description | Category | Trans-formation | Publication lag (days) |
|---|---|---|---|---|---|
| Small block | afvtots | Redundancy notices | Labor market | - | 15 |
| | bhuscon100s | Consumer confidence | Survey | - | -6 |
| | sesurv0242 | PMI, total services | Survey | - | 3 |
| | pviindufs | Production value index, industry | Production | pc | 39 |
| | ussurv1055 | PMI, total manufacturing, USA | International | - | 1 |
| | nbnpmpfs | Gross domestic product | Production | pcq | 60 |
| Medium block | arali1574s | Unemployment rate | Labor market | diff | 18 |
| | btvi201s | Prod. volume (expectations) | Survey | - | -6 |
| | sesurv0177 | PMI, total manufacturing | Survey | - | 1 |
| | setrad1516 | Household consumption indicator | Retail/consumption | pc | 41 |
| | ESI_euroarea19 | Economic sentiment, Euro area | International | - | -2 |
| | pvitje5fs | Production value index, services | Production | pc | 39 |
| | uextotsfs | Foreign trade, export | Production | pc | 29 |
| Large block | agpers | Aggregate gross pay | Labor market | pc | 43 |
| | asy1574s | Employment | Labor market | pc | 18 |
| | ati1574s | Hours worked | Labor market | pc | 18 |
| | btotcons | Total industry indicator | Survey | - | -6 |
| | btviordps | New orders (outcome), manuf. | Survey | - | -6 |
| | btvi101s | Prod. volume (outcome), manuf. | Survey | - | -6 |
| | kifi | Economic sentiment | Survey | - | -6 |
| | sesurv0188 | PMI, planned production | Survey | - | 1 |
| | sesurv0249 | PMI, planned business volume | Survey | - | 3 |
| | bt_regs | New registrations, passenger cars | Retail/consumption | log | 4 |
| | omsfs | Retail trade, sales volume | Retail/consumption | pc | 26 |
| | hox | HOX real estate price index | Housing | pc | 19 |
| | omsbr | Real estate transactions | Housing | pc | 10 |
| | seb_bo | SEB housing survey | Housing | pc | 8 |
| | kpifs | Underlying inflation | Prices | pc | 11 |
| | pppb_e | Producer price index | Prices | pc | 26 |
| | usprod1022 | Industrial production index, USA | International | pc | 16 |
| | euprod0002 | Industrial production index, EU | International | pc | 43 |
| | qorbctifs | Orders in industry | Production | pc | 39 |
| | seprod2406 | Electricity consumption | Production | pc | 41 |
| | uimtotsfs | Foreign trade, import | Production | pc | 29 |

*Note:* Transformation keys: diff = first difference, log = natural logarithm, pc = month-on-month percentage change, and pcq = quarter-on-quarter percentage change. The final column shows after how many days following the end of the reference period data are published. A negative number means that data are published the given number of days before the end of the reference period.

using the stylized publication calendar presented in Table 2. We use data as of September 2020 in recreating the real-time data available at each forecast origin since many variables are not revised and historical vintages are unavailable. The sole exception is GDP, for which we use historical vintages so that the data available at each forecast origin more truthfully mimics what was known at that time. We use a rolling window of 15 years of data for estimation, and evaluate forecasts against the first published outcome.

## 4.1 Comparison of Baseline Models

We begin by comparing the root mean squared error (RMSE) within each class of models with respect to the size of the model. Figure 1 displays the RMSE for each model and model size as a function of the number of days until the first release.[1]

Figure 1a shows the forecasting performance of the dynamic factor model. The medium-scale model has historically produced forecasts with higher precision during the nowcasting quarter. For backcasts, the large model has instead been better. Of all three model sizes, the model based on the large set of variables is the only version that shows a decreasing RMSE as more information is obtained throughout the quarter.

Among the bridge equation models displayed in Figure 1b, the small-scale model has a small edge over the medium-sized version. The model based on the large set of variables is on par with the two smaller versions in the beginning of the nowcasting quarter, but its nowcasts do subsequently not improve at the same rate.

The RMSEs for the MIDAS regressions (Figure 1c) show that the medium-sized model has been more accurate at every horizon evaluated. The model based on the large set of indicators performs comparably better than the small-scale model, which stands in contrast to the results for the bridge equations.

Figure 2 shows a comparison of selected versions from each model class. The figure also includes a pooled forecast. The pooled forecast is computed by averaging over the three models' forecasts.

The figure shows that our MIDAS and bridge equation models give clear improvements compared with the dynamic factor model, and that all methods improve at the same rate with more information. The medium-sized MIDAS model shows a small advantage over the small-scale bridge equation model at most horizons, although the difference is occasionally negligible. Furthermore, the figure shows that pooling forecasts by a simple unweighted average leads to better forecasts. The difference between the RMSE of the pooled forecast and that of MIDAS is tied at two of the considered horizons, but otherwise in favor of pooling.

To get a better understanding of the forecasts, Figure 3 shows the nowcasts made on the 18th day of the final month of each quarter. The three models' nowcasts have, for the most part, been relatively similar, in particular the nowcasts by the MIDAS and bridge equation

[1] In additional results (not reported here), we have also evaluated mixed-frequency VARs. However, these models, which are larger and more heavily parametrized than the alternatives we report here, occasionally exhibit unwieldy behavior and are generally unstable due to the fact that many of the time series we include are short and plagued by missing values for long periods of time in the beginning of the estimation samples. Other studies—when time series are not as short as ours—have found large mixed-frequency BVARs to be useful for forecasting, see for example Schorfheide and Song (2015); Cimadomo et al. (2020).

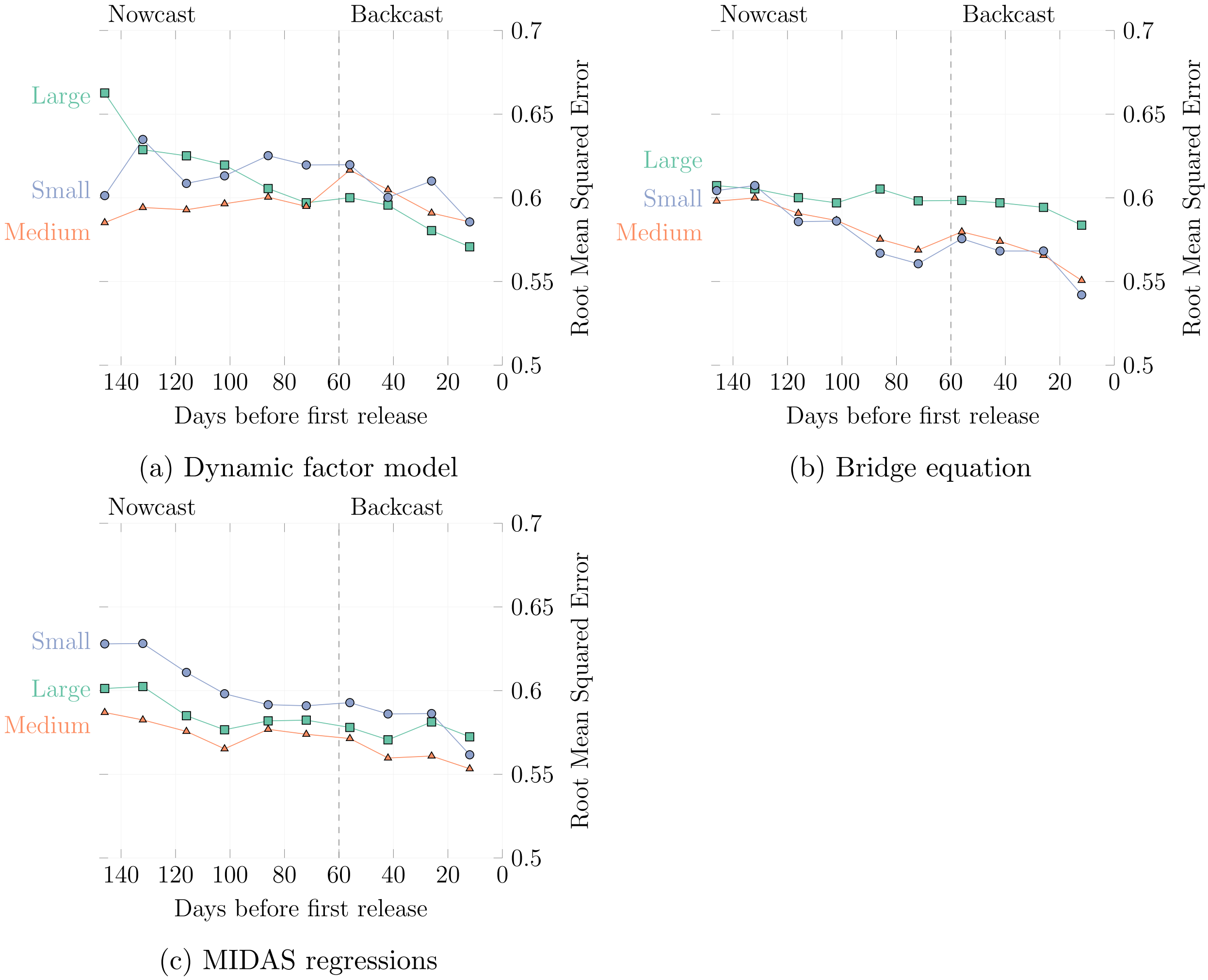



Figure 1: RMSE for forecast errors, 2010Q1–2019Q4

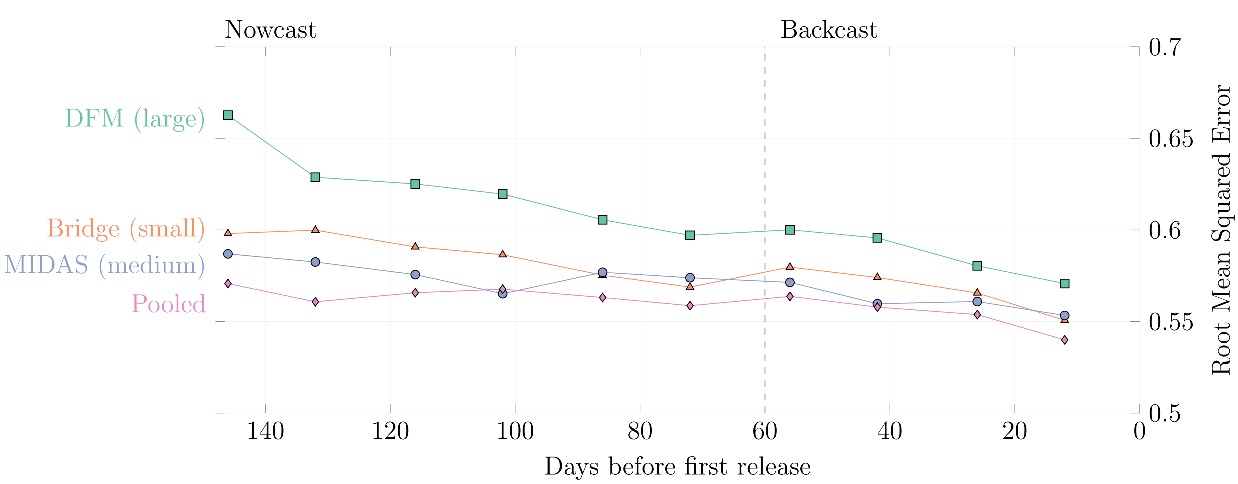



Figure 2: RMSE for forecast errors, 2010Q1–2019Q4

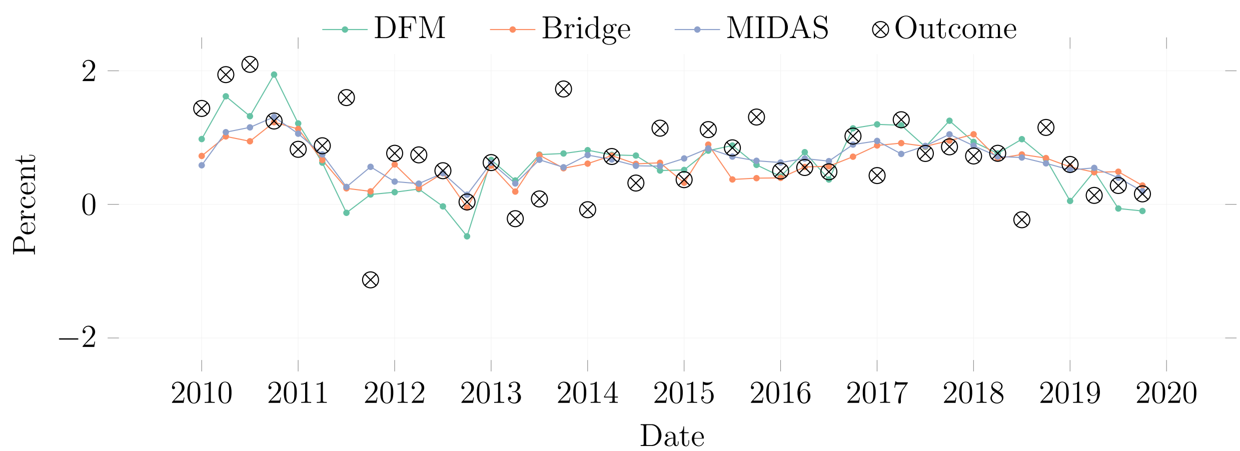



Figure 3: Nowcasts of GDP Growth. The nowcasts are made on the 18th day of the final month of each quarter.

models. One clear characteristic that is evident in the figure is that forecasts made by these two models are less volatile and vary less over time than those made by the dynamic factor model. The former two models thus exhibit a much larger degree of inertia. The dynamic factor model forecast, on the other hand, is subject to more movement, and the downside is that it occasionally makes larger forecasting errors, as for example in 2010Q4, 2011Q3, 2012Q4, and 2017Q1. The sluggishness likely explains the success of MIDAS and bridge equation models in terms of RMSE since they avoid some of the larger errors made by the dynamic factor model. Nevertheless, the dynamic factor model's tendency to make forecasts that deviate more from the local trend means that it is closer to some of the more atypical outcomes, as in the beginning of 2010.

## 4.2 Alternative Specifications

We next investigate the role of certain specification choices. To increase the relevance of this exercise for current nowcasting conditions, we shorten the evaluation period to cover 2016Q1–2019Q4, and for space considerations we limit the scope to a subset of horizons considered in the previous section. The alternative specifications we consider vary between the respective models. For the dynamic factor model, we assess the role of international variables by comparing our benchmark model with a model without any international variables, and one with additional international variables. The model with an enriched set of international variables contains 8 foreign variables in total, as compared to the benchmark model's 4 foreign variables.[2] To see whether a different number of lags would change the performance of the model, we let the number of lags of the factor that is included in each equation—$r$ in (1)—be either 1 or 6, respectively, instead of 3 as in the baseline version.

For the bridge equation model, the benchmark specification employs independent AR(4) models for the monthly auxiliary regressions, and the quarterly regressions are based on

[2] The four additional international variables that we include are Euro area manufacturing PMI, Euro area services PMI activity index, the Aruoba-Diebold-Scotti business conditions index for the US, and US imports.

regressing GDP growth on its lag as well as the contemporaneous and first lag of the aggregated monthly variable. To see how the specification of the monthly regressions affect the forecasting performance, we abandon the AR(4) regressions and instead use a monthly Bayesian VAR with four lags and an independent normal-Wishart prior (Karlsson, 2013). We evaluate the role of the specification of the quarterly regressions by removing the autoregressive term, thereby only regressing GDP growth on the contemporaneous and first lag of the aggregated monthly variable.

The baseline MIDAS specification that we employ is based on the UMIDAS approach taken by Foroni et al. (2015). To assess whether an alternative specification could improve the MIDAS model's forecasting performance, we switch to 12 lags and an exponential Almon lag parametrization of the model. Such a formulation of the model allows for large number of lags without increasing the number of parameters to be estimated, see Ghysels et al. (2007) for more details. Similar to the bridge equation model, we also evaluate our baseline model with the AR(1) component removed.

Finally, we investigate whether data-driven pooling strategies can improve upon the simple equal-weighted average that we considered in the previous section. The methods that we examine are stacking of means (Wolpert, 1992; Breiman, 1996) and inverse MSE weights (Bates and Granger, 1969; Timmermann, 2006). For the former, we begin by regressing the outcome in 2010Q1–2015Q4 on the two models' nowcasts for the same period, and compute the pooled nowcast for 2016Q1 by feeding the model-specific nowcasts through the estimated regression model. We then add 2016Q1 to the estimation sample, and continue in this manner throughout the evaluation period. The inverse MSE weights are based on the inverse of the MSE observed thus far when each nowcast is made. Table 3 shows the results of the evaluation for the alternative specifications.

The results show that none of the alternative specifications of the dynamic factor model leads to improved forecasts across all horizons, and that differences in performance are generally small. Adding additional foreign variables does not improve the model's predictive ability except for a minor increase in the second month of the nowcasting quarter, while removing all international variables leads to small improvements for the final nowcast as well as the backcast. Increasing the model's complexity by adding further lags of the factor in the observational equations of the model gives a substantial worsening of the first nowcast, and a small improvement to the second nowcast. Reducing the number of factor lags brings close to a 10 % reduction in RMSE compared with the benchmark specification for the second and third nowcasts, but comes at the cost of a notable increase in the RMSE of the backcast.

Turning to the bridge equation model, we find no improvement from dropping the autoregressive term from the quarterly regressions. The results also show that a joint forecasting approach to the monthly variables, utilizing a standard BVAR, is inferior to the simpler AR(4) approach that the benchmark uses with RMSE values that tend to increase with more information.

MIDAS regressions improve some at the final nowcast and backcast by eliminating the AR term from the models, but differences are small and its performance is inferior at the first two nowcasting horizons. Changing to the exponential Almon specification with 12 lags yields no major changes, with only minor differences at the first nowcast horizon and the backcast.

Lastly, we see no evidence that data-driven pooling is superior to the equal-weighted aver-

Table 3: RMSE, 2016Q1–2019Q4. Nowcasts and backcasts are made on the 18th day of each respective month. Backcast here refers to the forecast made in the first month following the end of the reference quarter.

| | Nowcast origin | | | |
|---|---|---|---|---|
| Model | Month 1 | Month 2 | Month 3 | Backcast |
| *Dynamic factor model* | | | | |
| Benchmark | 0.53 | 0.46 | 0.45 | 0.46 |
| No international | 0.56 | 0.46 | 0.42 | 0.45 |
| Additional international | 0.54 | 0.45 | 0.46 | 0.48 |
| $r = 1$ factor lags in eqs. | 0.54 | 0.41 | 0.40 | 0.49 |
| $r = 6$ factor lags in eqs. | 0.66 | 0.42 | 0.45 | 0.48 |
| *Bridge equation model* | | | | |
| Benchmark | 0.36 | 0.35 | 0.35 | 0.40 |
| No AR term | 0.36 | 0.35 | 0.36 | 0.41 |
| Monthly BVAR | 0.35 | 0.36 | 0.37 | 0.39 |
| *MIDAS regressions* | | | | |
| Benchmark | 0.36 | 0.35 | 0.36 | 0.34 |
| No AR term | 0.38 | 0.36 | 0.34 | 0.33 |
| 12 lags, exp. Almon | 0.35 | 0.35 | 0.36 | 0.36 |
| *Pooled forecast* | | | | |
| Benchmark | 0.34 | 0.35 | 0.36 | 0.38 |
| Inverse MSE | 0.34 | 0.35 | 0.36 | 0.40 |
| Stacking | 0.47 | 0.44 | 0.43 | 0.43 |

age that we use as our benchmark pooling strategy. Stacking leads to inferior RMSEs, while using inverse MSE weights yields the same level of forecasting precision as the benchmark except for at the backcast. These results are thus in line with many empirical studies that have identified a so-called forecast combination puzzle (Smith and Wallis, 2009; Claeskens et al., 2016).

In summary, we find no convincing evidence that any of the alternative model specifications is superior to our benchmark specifications. We therefore choose to keep our benchmark specifications intact.

## 4.3 New Data Sources

The rapid deterioration of the economic outlook following the spread of the Covid-19 virus highlighted the need for data sampled more frequently, and data that are released more rapidly. In this section, we evaluate if some of these new data help improve our dynamic factor model's nowcasts. We proceed as in the previous section, and evaluate selected nowcasts and backcasts in 2016Q1–2019Q4. Three new data series that were first released during the beginning of the pandemic are added to the benchmark dynamic factor model, one at a time and all together. The three series are: the Weekly Economic Index (WEI) published by the Federal Reserve Bank of New York, the monthly Activity Indicator (AI) published by Statistics Sweden, and the Retail Indicator (RI) published by Statistics Sweden.

WEI is a weekly index that closely tracks real economic activity in the US (Lewis et al., 2020). Because our model operates at the monthly frequency, we take a pragmatic approach and at a given forecast origin use the observed mean of the available weekly observations as a proxy for the monthly observation of WEI. The monthly observation is thus revised several times a month as new weekly releases are available. The index is published on the first Tuesday following the end of the reference week.

The Activity Indicator is a new indicator that "is intended to highlight the collected activities in the Swedish economy in a way similar to the description of growth in GDP" (Statistics Sweden, 2020b). As a complement, Statistics Sweden has also initiated the publication of a monthly indicator for retail sales that is based on preliminary statistics from the VAT register (Statistics Sweden, 2020a). The two indicators are published approximately 20–25 (RI) and 35–40 (AI) days after the end of the reference period.

The results of the forecast evaluations using these new data sources are presented in Table 4. Contrary to what one would expect, adding the two new Swedish indicators to the benchmark model leads to no better nowcasts according to the RMSE values. The differences are, however, minor and should not be taken to mean that these new indicators are not useful for nowcasting. A complication in evaluating the new data series in a forecasting application is that their history is relatively short—the indicators published by Statistics Sweden start in 2011, and so at the first nowcast in the evaluation only five years of data are made available to the model. Thus, the models that include the new indicators may be adversely affected by a lack of data and estimation uncertainty. Inclusion of the Weekly Economic Index for the US in the model reduces the RMSEs of the nowcasts for all three forecast origins, with no difference for the backcast. Nevertheless, despite the additional improvements obtained by adding the Weekly Economic Index to the benchmark dynamic factor model, it is still behind in terms of performance compared with our MIDAS regressions and our bridge equation

Table 4: RMSE, 2016Q1–2019Q4. Nowcasts and backcasts are made on the 18th day of each respective month. Backcast here refers to the forecast made in the first month following the end of the reference quarter. The National Institute of Economic Research's (NIER) forecasts are usually published by the end of the last month of a quarter, but the exact day varies.

| | Nowcast origin | | | |
|---|---|---|---|---|
| Model | Month 1 | Month 2 | Month 3 | Backcast |
| *Dynamic factor model* | | | | |
| Benchmark | 0.53 | 0.46 | 0.45 | 0.46 |
| Weekly Economic Index (WEI) | 0.49 | 0.45 | 0.44 | 0.46 |
| Activity Indicator (AI) | 0.55 | 0.47 | 0.45 | 0.45 |
| Retail Indicator (RI) | 0.55 | 0.46 | 0.45 | 0.46 |
| WEI+AI+RI | 0.53 | 0.48 | 0.45 | 0.47 |
| *Other models* | | | | |
| Bridge | 0.36 | 0.35 | 0.35 | 0.40 |
| MIDAS | 0.36 | 0.35 | 0.36 | 0.34 |
| Pooled | 0.34 | 0.34 | 0.36 | 0.38 |
| *Benchmark forecast* | | | | |
| NIER | | | 0.33 | |

model. Evaluated using approximately the same amount of information, the latter show RMSEs that are only marginally inferior to the forecasts made by the National Institute of Economic Research.

To interpret the magnitude of the differences in RMSE, Figure 4 shows the nowcasts from selected models. In general, the model-based nowcasts are typically similar, whereas the NIER's nowcasts stand apart to a greater degree. Nevertheless, the model-based nowcasts agree well with NIER's on several occasions, but were consistently higher between 2017Q3 and 2019Q2, and sometimes for the better.

# 5 Nowcasting During the Covid-19 Pandemic

The between-crises period 2010Q1–2019Q4 was characterized by relatively stable growth. However, economic conditions changed rapidly as the Covid-19 pandemic started to spread around the world in early 2020. In this section, we study this episode in some more detail from the perspective of nowcasting. We do so by first considering if the new data sources discussed in the previous section help produce a more timely signal of an impending severe downturn. We then examine the impact that the pandemic had on nowcasts through parameter instability by comparing nowcasts from models with and without re-estimation. Throughout the remainder of this section, we estimate models and produce nowcasts on Friday every week.

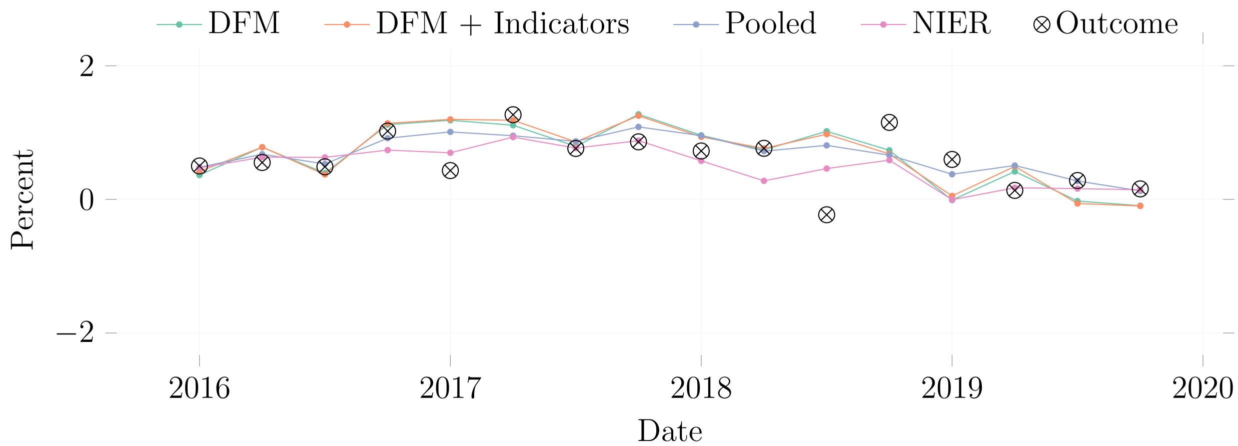


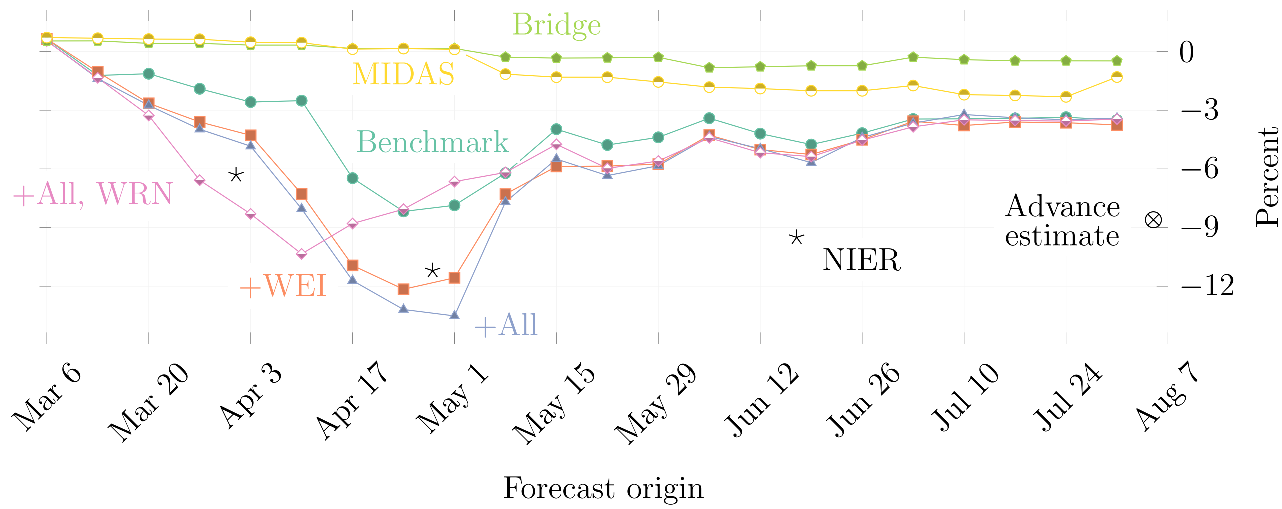

Figure 4: Nowcasts, 2016Q1–2019Q4. Model-based nowcasts are produced on the 18th of the final month of each quarter. The NIER's nowcasts are usually published towards the end of the final month of each quarter.

## 5.1 Using New Data Sources to Nowcast 2020Q2 and 2020Q3

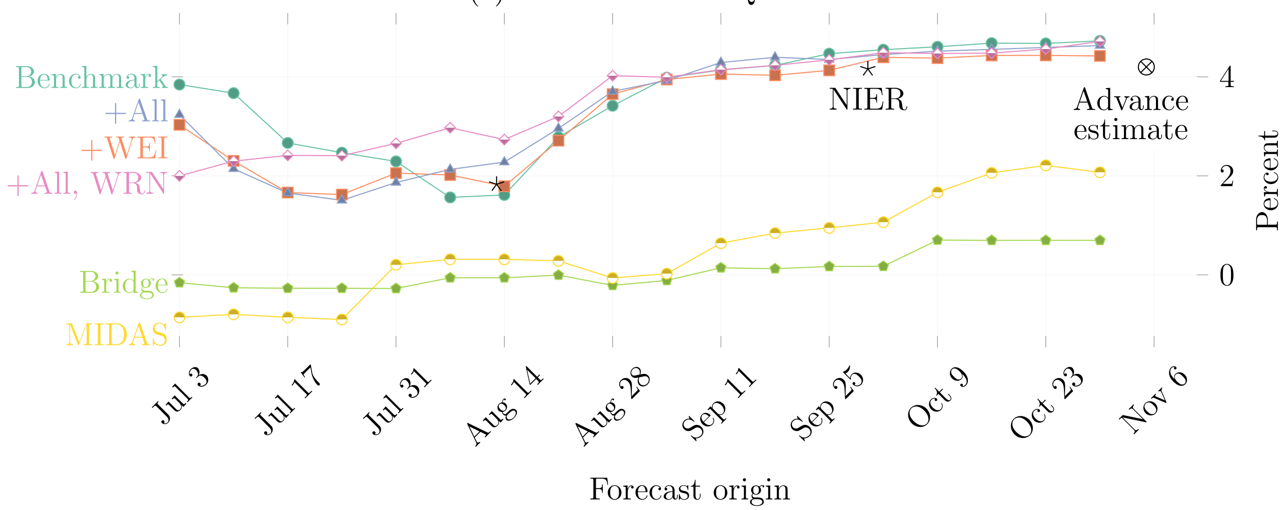
In Section 4.3, we evaluated three new indicators of economic activity and whether our factor model's nowcasts improved by adding these to the model. We next evaluate the same indicators in predicting the severe economic downturn that hit Sweden with full force in the second quarter of 2020. For ease of readability, we present in Figure 5 the benchmark model's nowcast and nowcasts obtained after adding the Weekly Economic Index, and all three indicators (Weekly Economic Index, Activity Indicator, Retail Indicator). Furthermore, we also make nowcasts where we exploit weekly data on redundancy notices.[3] We take the same approach as for the Weekly Economic Index and scale the weekly numbers to monthly numbers by assuming that the observed daily mean persists throughout the remainder of the month. Figure 5 shows the nowcasts.

The first panel of the figure shows that before the effects of the pandemic started to show in the data, all models produced nowcasts for 2020Q2 that were strikingly similar. However, as March progressed nowcasts started to diverge. Adding the new indicators lowered the nowcast by approximately 2 percentage points during the second half of March, and using weekly redundancy notices as a proxy for the monthly observation pulled the nowcast down even further and earlier in March and early April. The model with weekly redundancy notices subsequently revised its nowcast upwards in April, which is explained by each weekly release of newly reported redundancy notices in April being lower than the previous week's. The proxy for the monthly observation, obtained from the weekly data, was thus revised downwards each week. The two models with the Weekly Economic Index and all three indicators, respectively, reached the same nowcast one week later, at which point the model with weekly redundancy notices had started to revise its nowcast upwards. The benchmark model, without weekly redundancy notices and any of the three new indicators, reached its most pessimistic point an additional week later. While the precise trajectories in the figure

[3]The Swedish Public Employment Service began releasing weekly figures of reported redundancy notices on March 14. Each Monday, the reported number of redundancy notices for the preceding week was reported.




(a) Nowcast for 2020Q2




(b) Nowcast for 2020Q3

Figure 5: Nowcasts of GDP Growth During 2020Q2 and 2020Q3. 'Benchmark' is the benchmark specification of the dynamic factor model, which is subsequently augmented with: the Weekly Economic Index ('+WEI'); the Weekly Economic Index, Activity Indicator, and Retail Indicator ('+All'); all three indicators and weekly observations of redundancy notices ('+All, WRN').

are somewhat different, all versions here considered interpreted the inflow of data as more and more pessimistic until reaching a tipping point in April. Beyond that point, they all displayed a more optimistic interpretation of the economic outlook. While the magnitudes are arguably different, this interpretation of the economic situation is in accordance with that made by the National Institute of Economic Research, which revised its forecasts in June and August upwards due to more positive signals since the middle of April. The additional indicators helped in May–August to keep the nowcast and backcast for 2020Q2 at a lower level, which ultimately also turned out to be closer to the outcome of -8.3. The nowcasts produced by the MIDAS and bridge equation models were throughout the quarter much less volatile—a pattern also demonstrated historically in Figure 4—but also much further from the outcome, ending at approximately -1 as opposed to the dynamic factor models whose final predictions came in around -4.

The second panel, Figure 5b, displays the evolution of the nowcast for 2020Q3. Overall, all models interpreted new data from the second half of the quarter and onwards as more positive with an upwards trending development. The nowcasts from the various versions of the dynamic factor model are both qualitatively and quantitatively very similar, predicting GDP to grow by a little more than 4 percent in 2020Q3 from early September. Moreover, the models' nowcasts were closely in line with the two nowcasts made by the National Institute of Economic Research in August and September. The bridge equation model again shows a much more slowly changing prediction that in the end failed to pick up the magnitude of the economic recovery, whereas the MIDAS model displays more substantial upwards revisions in September and October that eventually leads to predictions that are closer to the outcome than the bridge equation model's. In comparison with the dynamic factor model, however, its forecasts are still associated with notably larger forecast errors.

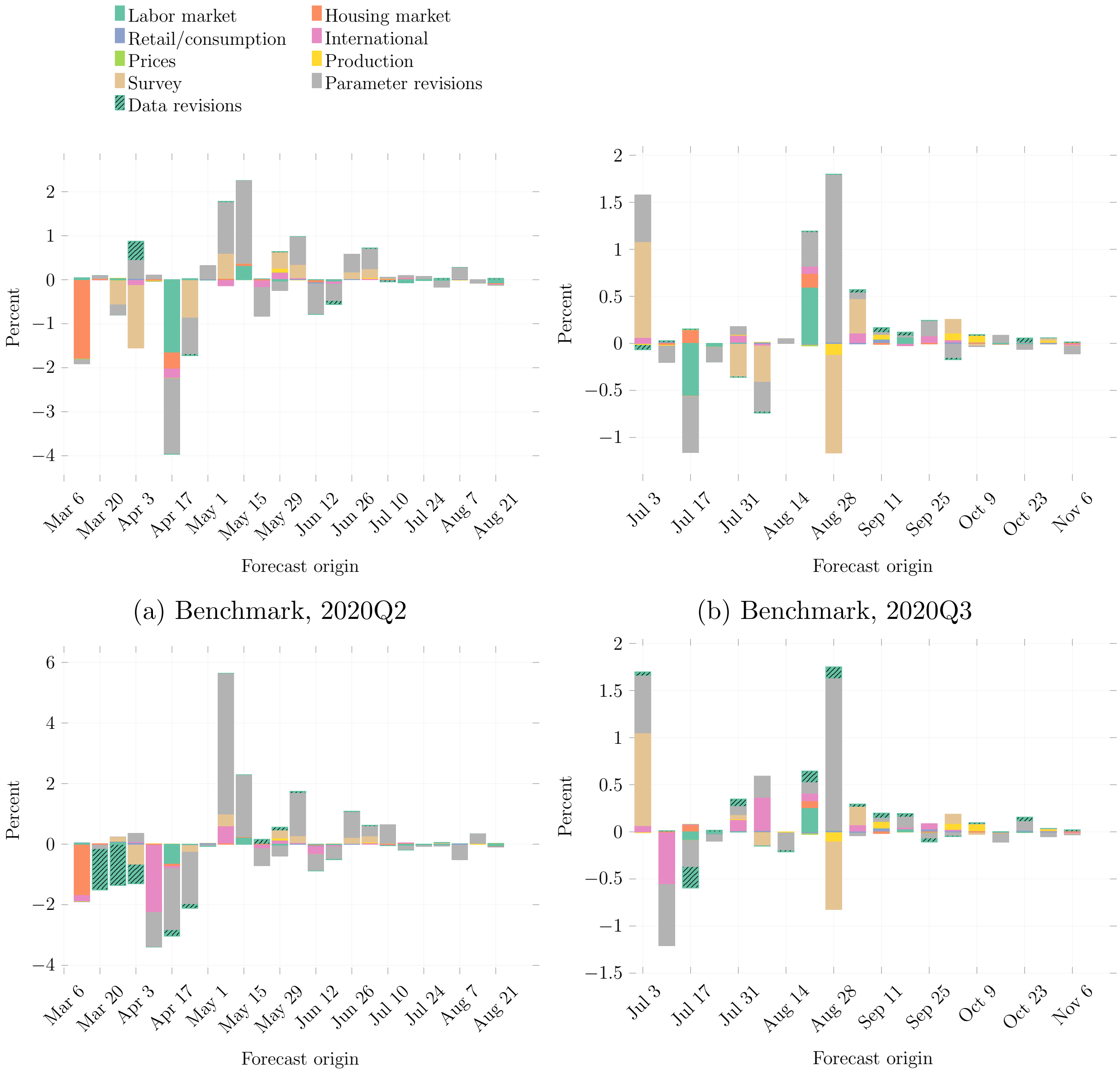
## 5.2 Decomposing the Nowcasts

In this section we decompose the revisions of the nowcasts into contributions to provide further understanding of the nowcasts in Figure 5.[4] The contributions to the revisions are divided into three main categories: parameters, historical data, and new data releases. The contributions from new data releases are further decomposed into contributions from different categories of variables following the categorization in Table 2 and the method by Bańbura and Modugno (2014). The computation of the contribution from a new data release has the form

$$\text{Contribution from variable } j = \text{weight}_j \times (\text{Outcome}_j - \text{Forecast}_j).$$

That is, how much a new data release for variable $j$ contributes to the revision of the forecast for GDP growth depends on the weight of variable $j$, itself a function of the model dynamics, and if the new data release is a surprise in the sense that the new datum deviates from the model's previous forecast of it. An important implication of the decomposition is that the order of releases matter, since new data for variable $j$ can change $\text{Forecast}_{j'}$ for variable $j'$. Thus, data that are released with significant publication delays will only warrant a revision of the GDP forecast if these releases were unexpected given data already available.

[4] More details are provided in Appendix B.




(a) Benchmark, 2020Q2

(b) Benchmark, 2020Q3

(c) Model with New Indicators, 2020Q2

(d) Model with New Indicators, 2020Q3

Figure 6: Contributions to Revisions of the Dynamic Factor Models' Nowcasts

The decompositions are presented in Figure 6. Figure 6a shows the contributions to the benchmark model's nowcast for 2020Q2, which makes it clear that the major changes observed in Figure 5a have primarily been driven by parameter revisions and survey data. Unsurprisingly, survey data accounts for most of the revisions due to new data releases: because it is released prior to all other data, contributions computed for other categories are calculated given the already observed surveys. Nevertheless, we see some larger revisions that stem from variables covering the housing market and the labor market.

Comparing Figure 6a with Figure 6c, we see that in the latter large contributions from international variables is now observed. The explanation is that the Weekly Economic Index, measuring real economic activity in the US, is included among the new indicators used for Figure 6c. Because it is released prior to the survey data, it assumes some of the contribution that was previously assigned to the survey category. The contributions from historical data revisions that is now observable in Figure 6c can be attributed to the monthly observation of the Weekly Economic Index being revised throughout each month as new weekly releases are available.

The decompositions of revisions of nowcasts for 2020Q3, displayed in Figure 6b and 6d, show similar patterns with survey data explaining most of the revisions to the benchmark model's nowcast. When the new indicators are added, the international category picks up some of these revisions at an earlier stage through the Weekly Economic Index. Comparing the two figures, it is also evident that the inclusion of the Weekly Economic Index picks up some of the revisions previously attributed to the labor market as seen on July 17 and August 21.

### 5.3 Nowcasts Without Re-Estimation

The previous section showed that nowcasts were subject to large revisions during the peak of the pandemic solely due to parameters being re-estimated. A natural question to ask is therefore: does re-estimation during the Covid-19 pandemic lead to inferior nowcasts? Schorfheide and Song (2020) estimate their mixed-frequency BVAR on the January 31, 2020 vintage so that shock variances in the period following the most acute phase of the pandemic are of the same magnitude as in the pre-pandemic era. Lenza and Primiceri (2020) develop a method for estimating VAR models that explicitly accounts for a sudden increase in shock variances, whereafter the temporary increase decays, and Antolín-Díaz et al. (2020) explicitly account for outliers. Similarly, the Federal Reserve Bank of New York ceased in 2020Q2 to re-estimate its nowcasting model, instead using the model estimated in the first quarter of 2020.[5] On the other side of the spectrum, the GDPNow model developed by Higgins (2014) and published online by the Federal Reserve Bank of Atlanta previously removed outliers in estimation, but since April 30, 2020, such preprocessing of the data is no longer conducted (Federal Reserve Bank of Atlanta, 2020).

To see the effect of re-estimation and parameter instability on our nowcasts, we show in Figure 7 nowcasts for 2020Q2 and 2020Q3. The models that are not subject to re-estimation are estimated on March 6. We compute nowcasts every Friday from the first week of the

[5]As of November 3, 2020, the online Nowcasting Report, available at `https://www.newyorkfed.org/research/policy/nowcast.html`, reads: "The Nowcast estimates are [...] based on the same parameters used during 2020:Q2, based on data through the end of 2020:Q1".

nowcast quarter until the publication of the outcome. The nowcasts are presented in Figure 7.

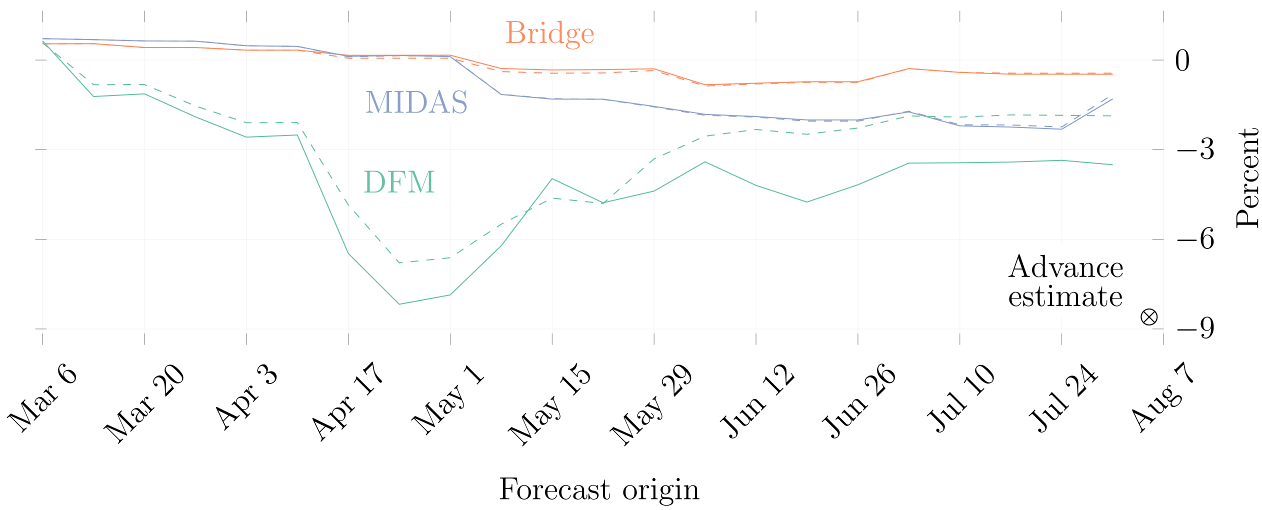
The figure shows that the estimation sample chosen does matter. The MIDAS and bridge equation models' nowcasts are only marginally affected by re-estimation for 2020Q2, whereas the dynamic factor model displays a much larger discrepancy between nowcasts derived from a repeatedly re-estimated model, and a model that keeps the model estimated on March 6 fixed. The re-estimated model returns more negative nowcasts already from the beginning of March and onwards, and the associated forecasting errors are smaller if the model is re-estimated. In 2020Q3, the difference between re-estimated and fixed-estimation forecasts decreases up through the middle of the quarter, at which point the fixed-estimation forecast settles at 1 percent, while the re-estimated model's forecasts increase to approximately 4.5 percent—only 0.2 percentage points from the official advance estimate published in early November. Thus, we find that whether one opts to re-estimate the dynamic factor model does matter, and that in our case re-estimation leads to nowcasts and backcasts that, in hindsight, are more reasonable. For our MIDAS and bridge equation models, the differences are negligible for the 2020Q2 forecasts, but the re-estimated MIDAS model eventually yields smaller forecasting errors for 2020Q3 than its fixed-estimation counterpart.

# 6 Conclusion

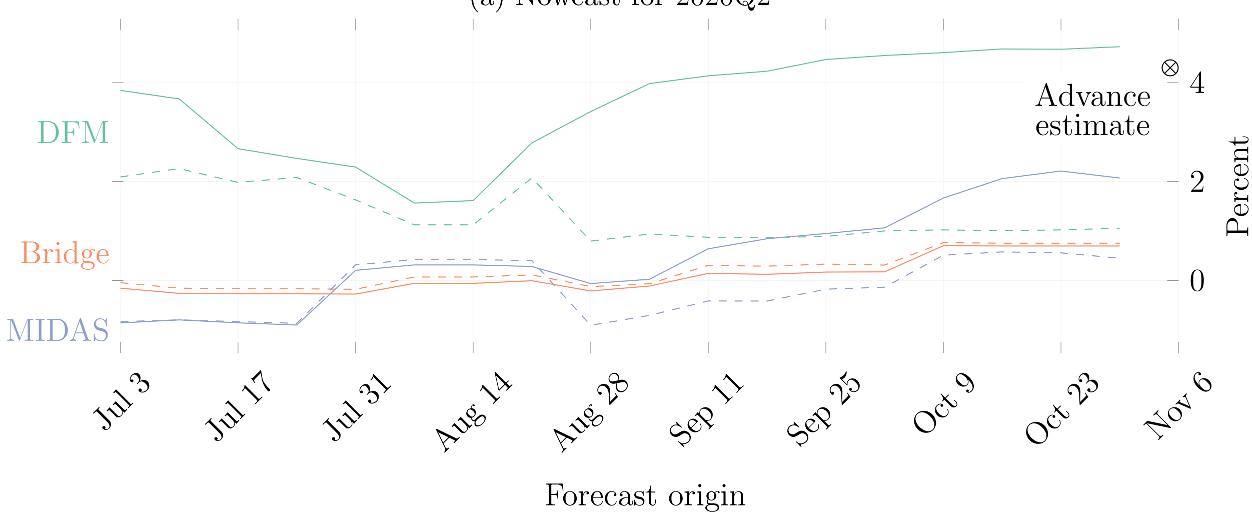
We have examined the use of short-term forecasting models for nowcasting Swedish GDP growth. Our key results show that simple specifications of MIDAS regressions and bridge equation models outperform a dynamic factor model in an evaluation period set to 2010Q1–2019Q4. We find that the three methods are idiosyncratic in their preference for data: the most satisfactory forecasting results for each method are obtained from a large version of the dynamic factor model, a medium-scale set of aggregated MIDAS regressions, and a small bridge equation model. These results point to the importance of leveling conditions to compare forecasting models fairly—by evaluating three different sizes of models for each method, we decrease the risk of biasing our results in favor of a particular method. We further examine various specification choices and find only marginal differences in historical forecast performance, suggesting that our results are not driven by specific details of our benchmark specifications.

For the most part, the dynamic factor model's nowcasts are qualitatively similar to the MIDAS and bridge equation models' nowcasts, but the latter models exhibit a larger degree of inertia and less volatility over time. When we use our models for nowcasting 2020Q2 and 2020Q3, the sluggishness becomes immediately clear: the MIDAS and bridge equation models react slowly during 2020Q2, and not to the same degree as the dynamic factor model. For 2020Q3, the same behavior is evident with the forecast made by the dynamic factor model eventually stabilizing close to the advance estimate, and remarkably close to professional forecasts published around the same time. We therefore find a clear discrepancy between, on the one hand, better historical forecasting performance, and, on the other hand, usefulness in the pandemic-induced downturn and subsequent recovery in 2020Q2 and 2020Q3.

Our results emphasize the need to consider multiple models, a notion that has been




(a) Nowcast for 2020Q2




(b) Nowcast for 2020Q3

Figure 7: Forecasts of GDP Growth at Different Points in Time. Solid lines are forecasts made by re-estimated models, and dashed lines show forecasts made by models estimated on March 6.

stressed by many, see for example Anesti et al. (2017); Bańbura and Saiz (2020) for recent descriptions of how the Bank of England and the European Central Bank maintain multiple models in informing their forecasts. The clear divide in performance that we document between normal and crisis times cannot be easily remedied by model combination approaches that rely on average historical performance: because change is sudden, a successful combination strategy needs to quickly redistribute its weight, for example through outcome-dependent weighting (Kapetanios et al., 2015) or pure judgmental assessment of how much each model should be incorporated into the combined forecast (Anesti et al., 2017). The idea of using similar historical episodes in producing forecasts at a specific point in time is another promising alternative that could potentially combine good predictive performance historically with rapid and accurate responses during turbulent times, as Dendramis et al. (2020); Foroni et al. (2020) discuss. Lerch et al. (2017) describe the danger of evaluating models conditional on extreme events, and one should therefore be wary of arguing that our dynamic factor model is a better model in general because it is more successful in 2020Q2–2020Q3. Chauvet and Potter (2013) document clear divides in forecasting performance between periods of boom and bust, and our results can be understood from the perspective of model performance varying with the business cycle. In light of this, our results thus make a strong case for maintaining multiple models to diversify risks, and closely monitor disagreements and rapid changes that occur both jointly and idiosyncratically.

# References


Andersson, M. K. and den Reijer, A. H. J. (2015). Nowcasting. *Sveriges Riksbank Economic Review*, 1:75–89.

Anesti, N., Hayes, S., Moreira, A., and Tasker, J. (2017). Peering into the Present: The Bank's Approach to GDP Nowcasting. *Bank of England Quarterly Bulletin*, 57(2):122–133.

Angelini, E., Rünstler, G., and Bańbura, M. (2008). Estimating and forecasting the euro area monthly national accounts from a dynamic factor model. Working Paper Series 953, European Central Bank.

Antolín-Díaz, J., Drechsel, T., and Petrella, I. (2020). Advances in Nowcasting Economic Activity: Secular Trends, Large Shocks and New Data. Unpublished manuscript.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Now-Casting and the Real-Time Data Flow. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, pages 195–237. Elsevier, `doi:10.1016/B978-0-444-53683-9.00004-9`.

Bańbura, M. and Modugno, M. (2014). Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data. *Journal of Applied Econometrics*, 29(1):133–160, `doi:10.1002/jae.2306`.

Bańbura, M. and Saiz, L. (2020). Short-Term Forecasting of Euro Area Economic Activity at the ECB. *ECB Economic Bulletin*, (2):84–95.

Bates, J. M. and Granger, C. W. J. (1969). The Combination of Forecasts. *Journal of the Operational Research Society*, 20(4):451–468, doi:10.1057/jors.1969.103.

Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M., and Tambalotti, A. (2017). Macroeconomic Nowcasting and Forecasting with Big Data. Staff Reports 830, Federal Reserve Bank of New York.

Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24(1):49–64, doi:10.1007/BF00117832.

Bundesbank (2013). Forecasting Models in Short-Term Business Cycle Analysis – A Workshop Report. Monthly Report September 2013, Deutsche Bundesbank.

Chauvet, M. and Potter, S. (2013). Forecasting Output. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, chapter 3, pages 141–194. Elsevier, doi:10.1016/B978-0-444-53683-9.00003-7.

Chernis, T. and Sekkel, R. (2017). A Dynamic Factor Model for Nowcasting Canadian GDP Growth. *Empirical Economics*, 53(1):217–234, doi:10.1007/s00181-017-1254-1.

Cimadomo, J., Giannone, D., Lenza, M., Sokol, A., and Monti, F. (2020). Nowcasting with large Bayesian vector autoregressions. Working Paper Series 2453, European Central Bank.

Claeskens, G., Magnus, J. R., Vasnev, A. L., and Wang, W. (2016). The forecast combination puzzle: A simple theoretical explanation. *International Journal of Forecasting*, 32(3):754–762, doi:10.1016/j.ijforecast.2015.12.005.

D'Agostino, A., Giannone, D., Lenza, M., and Modugno, M. (2016). Nowcasting Business Cycles: A Bayesian Approach to Dynamic Heterogeneous Factor Models. In *Dynamic Factor Models*, volume 35, pages 569–594. Emerald Publishing Ltd.

den Reijer, A. and Johansson, A. (2019). Nowcasting Swedish GDP with a large and unbalanced data set. *Empirical Economics*, 57(4):1351–1373, doi:10.1007/s00181-018-1500-1.

Dendramis, Y., Kapetanios, G., and Marcellino, M. (2020). A similarity-based approach for macroeconomic forecasting. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(3):801–827, doi:10.1111/rssa.12574.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, 2 edition.

ECB (2008). Short-Term Forecasts of Economic Activity in the Euro Area. *ECB Monthly Bulletin April 2008*, pages 69–74.

Federal Reserve Bank of Atlanta (2020). Modifications to GDPNow Model. https://www.frbatlanta.org/-/media/documents/cqer/researchcq/gdpnow/ModificationsToGDPNowModel.pdf. Accessed: November 3, 2020.

Foroni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1):57–82, doi:10.1111/rssa.12043.

Foroni, C., Marcellino, M., and Stevanović, D. (2020). Forecasting the Covid-19 recession and recovery: lessons from the financial crisis. Working Paper Series 2468, European Central Bank.

Galli, A., Hepenstrick, C., and Scheufele, R. (2019). Mixed-Frequency Models for Tracking Short-Term Economic Developments in Switzerland. *International Journal of Central Banking*, 15(2):151–178.

Geweke, J. (1977). The Dynamic Factor Analysis of Economic Time-Series Models. In *Latent Variables in Socio-Economic Models*. North-Holland, New York.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 26(1):53–90, doi:10.1080/07474930600972467.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The Real-Time Informational Content of Macroeconomic Data. *Journal of Monetary Economics*, 55(4):665–676.

Higgins, P. C. (2014). GDPNow: A Model for GDP "Nowcasting". FRB Atlanta Working Paper 7, Federal Reserve Bank of Atlanta.

Kapetanios, G., Mitchell, J., Price, S., and Fawcett, N. (2015). Generalised density forecast combinations. *Journal of Econometrics*, 188(1):150–165, doi:10.1016/j.jeconom.2015.02.047.

Karlsson, S. (2013). Forecasting with Bayesian Vector Autoregression. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, chapter 15, pages 791–897. Elsevier, doi:10.1016/B978-0-444-62731-5.00015-4.

Kuzin, V., Marcellino, M., and Schumacher, C. (2013). Pooling Versus Model Selection for Nowcasting GDP with Many Predictors: Empirical Evidence for Six Industrialized Countries. *Journal of Applied Econometrics*, 28(3):392–411, doi:10.1002/jae.2279.

Lenza, M. and Primiceri, G. E. (2020). How to Estimate a VAR After March 2020. Working Paper Series 2461, European Central Bank.

Lerch, S., Thorarinsdottir, T. L., Ravazzolo, F., and Gneiting, T. (2017). Forecaster's Dilemma: Extreme Events and Forecast Evaluation. *Statistical Science*, 32(1):106–127, doi:10.1214/16-STS588.

Lewis, D., Mertens, K., and Stock, J. H. (2020). U.S. Economic Activity during the Early Weeks of the SARS-Cov-2 Outbreak,. Staff Reports 920, Federal Reserve Bank of New York.

Luciani, M. and Ricci, L. (2014). Nowcasting Norway. *International Journal of Central Banking*, 10(4):215–248.

Mariano, R. S. and Murasawa, Y. (2003). A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, 18(4):427–443, doi:10.1002/jae.695.

Ollivaud, P., Pionnier, P.-A., Rusticelli, E., Schwellnus, C., and Koh, S.-H. (2016). Forecasting GDP During and After the Great Recession: A Contest Between Small-Scale Bridge and Large-Scale Dynamic Factor Models. OECD Economics Department Working Papers 1313, OECD Publishing, doi:10.1787/5jlv2jj4mw40-en.

Schorfheide, F. and Song, D. (2015). Real-Time Forecasting With a Mixed-Frequency VAR. *Journal of Business & Economic Statistics*, 33(3):366–380, doi:10.1080/07350015.2014.954.

Schorfheide, F. and Song, D. (2020). Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic. Working Papers 20-26, Federal Reserve Bank of Philadelphia, doi:10.21799/frbp.wp.2020.26.

Siliverstovs, B. (2020). Assessing nowcast accuracy of US GDP growth in real time: the role of booms and busts. *Empirical Economics*, 58(1):7–27, doi:10.1007/s00181-019-01704-6.

Smith, J. and Wallis, K. F. (2009). A Simple Explanation of the Forecast Combination Puzzle. *Oxford Bulletin of Economics and Statistics*, 71(3):331–355, doi:10.1111/j.1468-0084.2008.00541.x.

Statistics Sweden (2020a). Major economic differences between activities in the wake of the coronavirus outbreak. Press release, 2020-04-24. https://www.scb.se/en/About-us/news-and-press-releases/major-economic-differences-between-activities-in-the-wake-of-the-coronavirus-outbreak

Statistics Sweden (2020b). Statistics Sweden to publish new activity indicator. Press release, 2020-06-04. https://www.scb.se/en/About-us/news-and-press-releases/statistics-sweden-to-publish-new-activity-indicator/.

Timmermann, A. (2006). Forecast Combinations. volume 1 of *Handbook of Economic Forecasting*, chapter 4, pages 135–196. Elsevier, doi:10.1016/S1574-0706(05)01004-9.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2):241–259, doi:10.1016/S0893-6080(05)80023-1.

# A Estimating the Dynamic Factor Model

## A.1 State-Space Formulation

We define the dynamic factor model as a standard linear state-space model with normally distributed errors to seamlessly handle missing values and mixed frequencies. The state-space model is

$$\begin{aligned}\alpha_{t+1} &= A\alpha_t + Be_t \\ y_t &= C\alpha_t,\end{aligned}$$

where $e_t \sim N(0, I)$ and the state vector $\alpha_t$ is

$$\alpha_t = \begin{pmatrix} f_t \\ \vdots \\ f_{t-\bar{r}} \\ x_t \\ \vdots \\ x_{t-\bar{p}} \end{pmatrix}$$

where $\bar{r} = \max(r, q, 4)$ and $\bar{p} = \max(p, 4)$. The factor model can be rewritten as

$$\begin{aligned}x_{i,t} &= \lambda_{i,0} f_t + \sum_{j=1}^{r} \lambda_{i,j} f_{t-j} + \sum_{k=1}^{p} \phi_{i,k} x_{i,t-k} + \varepsilon_{i,t} \\ &= \lambda_{i,0} \sum_{j=1}^{q} \varphi_j f_{t-j} + \sum_{j=1}^{r} \lambda_{i,j} f_{t-j} + \sum_{k=1}^{p} \phi_{i,k} x_{i,t-k} + \varepsilon_{i,t} + \lambda_{i,0} u_t \\ &= \sum_{j=1}^{r} (\lambda_{i,j} + \lambda_{i,0}\varphi_j) f_{t-j} + \sum_{k=1}^{p} \phi_{i,k} x_{i,t-k} + \varepsilon_{i,t} + \lambda_{i,0} u_t,\end{aligned}$$

where $q \leq r$ and $\varphi_j = 0$ if $j > q$. Let $\bar{\lambda}_{i,j} = \lambda_{i,j} + \lambda_{i,0}\varphi_j$. The transition matrix $A$ is constructed as

$$A = \begin{pmatrix} \underset{\bar{r}\times\bar{r}}{A_1} & \underset{\bar{r}\times n\bar{p}}{0} \\ \underset{n\bar{p}\times\bar{r}}{A_2} & \underset{n\bar{p}\times n\bar{p}}{A_3} \end{pmatrix}, A_1 = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{\bar{r}-1} & \varphi_{\bar{r}} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} \bar{\lambda}_{1,1} & \bar{\lambda}_{1,2} & \cdots & \bar{\lambda}_{1,\bar{r}} \\ \bar{\lambda}_{2,1} & \bar{\lambda}_{2,2} & \cdots & \bar{\lambda}_{2,\bar{r}} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\lambda}_{n,1} & \bar{\lambda}_{n,2} & \cdots & \bar{\lambda}_{n,\bar{r}} \\ 0_{n(\bar{r}-1)\times 1} & 0_{n(\bar{r}-1)\times 1} & \cdots & 0_{n(\bar{r}-1)\times 1} \end{pmatrix}, A_3 = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{\bar{p}-1} & \Phi_{\bar{p}} \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{pmatrix}$$

Let $e_t = (u_t, \sigma_1^{-0.5}\varepsilon_{1,t}, \ldots, \sigma_n^{-0.5}\varepsilon_{n,t})'$. Then

$$B = \begin{pmatrix}\lambda_0 & \mathrm{diag}(\sigma_1, \ldots, \sigma_n).\end{pmatrix}$$

The matrix $C$ that relates the observed values to the state vector imposes the following relation for monthly variables ($i = 1, \ldots, n-1$):

$$y_{i,t} = \begin{cases} x_{i,t}, & \text{if } y_{i,t} \text{ is observed} \\ \varnothing, & \text{otherwise.} \end{cases}$$

For the final quarterly variable ($i = n$), the relation is

$$y_{i,t} = \begin{cases} \frac{1}{3}(x_{i,t} + 2x_{i,t-1} + 3x_{i,t-2} + 2x_{i,t-3} + x_{i,t-4}), & \text{if } y_{i,t} \text{ is observed} \\ \varnothing, & \text{otherwise.} \end{cases}$$

## A.2 Prior Distributions

The prior distribution for parameters relating factor lags to observables is $\lambda_{i,0} \sim \mathrm{N}(0.5, 1)$, and $\lambda_{i,j} \sim \mathrm{N}\left(0, (1+j)^{-1}\right)$ for $j = 1, \ldots, r$. Priors for parameters for own lags of observables follow a similar form, and are $\phi_{i,k} \sim \mathrm{N}\left(0, (1+k)^{-1}\right)$ $(k = 1, \ldots, p)$, but equipped with the additional constraint that the lag polynomial has stable roots. Finally, autoregressive parameters in the factor equation are assumed to define a stable lag polynomial, with prior distributions given by $\varphi_1 \sim \mathrm{N}(0.5, 0.5)$ and $\varphi_\ell \sim \mathrm{N}\left(0, (1+\ell)^{-1}\right)$ for $\ell = 2, \ldots, q$. All data enter the model in standardized form, with zero mean and unit variance.

## A.3 Posterior Sampling

The model is estimated using Markov Chain Monte Carlo (MCMC) and Gibbs sampling. The Gibbs sampler cycles through the following steps.

$p(f, X|Y, \lambda, \phi, \varphi, \sigma^2)$ We use disturbance smoothing using the state-space formulation of the dynamic factor model laid out in the previous section to sample from the posterior distribution of factors and missing observations. See Durbin and Koopman (2012) for a detailed description.

$p(\lambda, \phi, \sigma^2|X, f)$ The posterior conditional on $X$ and $f$ reduces to a standard linear regression problem with a normal-inverse gamma posterior. Each equation is conditionally independent of the others, and we therefore sample $p(\lambda_i, \phi_i, \sigma_i^2|X_i, f)$ equation-by-equation.

$p(\varphi|f)$ We sample from the posterior of $\varphi$ conditional on $f$ using Bayesian linear regression with known variance.

# B Decomposing Forecast Revisions in the Dynamic Factor Model

Let: $t_2 > t_1$ be two time points, $\Omega_{t_1}$ and $\Omega_{t_2}$ be the information sets at the respective points in time, and $\theta_{t_2}$ and $\theta_{t_1}$ be parameters estimated at each time point. For now, we suppose that these are known and fixed. Because of historical data revisions, $\Omega_{t_1}$ is not necessarily a subset of $\Omega_{t_2}$. Partition $\Omega_{t_2}$ into two disjoint subsets, $\Omega_{t_2} = \Omega_{t_2}^{(1)} \cup \Omega_{t_2}^{(2)}$ so that $\Omega_{t_2}^{(1)} \cap \Omega_{t_2}^{(2)} = \emptyset$, where $\Omega_{t_2}^{(1)}$ refers to the same observations as $\Omega_{t_1}$. Thus, if no historical data revisions have taken place, then $\Omega_{t_1} = \Omega_{t_2}^{(1)}$. Let $t$ be the period that is the forecasting target. We can then write the forecast revision as consisting of three parts according to

$$
\begin{aligned}
&E(y_t|\theta_{t_2}, \Omega_{t_2}) - E(y_t|\theta_{t_1}, \Omega_{t_1}) = \\
&\underbrace{E(y_t|\theta_{t_2}, \Omega_{t_1}) - E(y_t|\theta_{t_1}, \Omega_{t_1})}_{\text{Parameters}} + \underbrace{E(y_t|\theta_{t_2}, \Omega_{t_2}^{(1)}) - E(y_t|\theta_{t_2}, \Omega_{t_1})}_{\text{Historical data}} + \underbrace{E(y_t|\theta_{t_2}, \Omega_{t_2}) - E(y_t|\theta_{t_2}, \Omega_{t_2}^{(1)})}_{\text{News}}.
\end{aligned}
$$

The first part measures the contribution to the overall forecast revision that stems from a change in parameters, since it is computed as the difference between the forecast given the new parameters and the old data minus the forecast given the old parameters and the old data. The second term is computed as the difference between the forecast produced using revised and un-revised data while using the new parameters in both cases. The final term is the news part and represents the revision caused by the new data in $\Omega_{t_2}^{(2)}$. Bańbura and Modugno (2014) analyze this term in more detail and derive a way of decomposing the overall news revision into contributions from each new data release included in $\Omega_{t_2}^{(2)}$. To calculate the contributions in practice, we compute the various expectations using Kalman filtering with each $\theta$ replaced by a draw from the posterior distribution.