WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Nowcasting World GDP Growth with High-Frequency Data

Caroline Jardet[1] and Baptiste Meunier[2]

December 2020, WP #788

Abstract

The Covid-19 crisis has shown how high-frequency data can help tracking economic turning points in real-time. Our paper investigates whether high-frequency data can also improve the nowcasting performances for world GDP growth on quarterly or annual basis. To this end, we select a large dataset of 151 monthly and 39 weekly series for 17 advanced and emerging countries representing 68% of world GDP. Our approach builds on a Factor-Augmented MIxed DAta Sampling (FA-MIDAS) which allows us to take advantage of our large database and to combine different frequencies. Models that include weekly data significantly outperforms other models relying on monthly or quarterly indicators, both in- and out-of-sample. Breaking down our sample, we show that models with weekly data have similar nowcasting performances relative to other models during "normal" times but strongly outperform them during "crisis" episodes (2008-2009 and 2020). We finally construct a nowcasting model of annual world GDP growth incorporating weekly data which give timely (one every week) and accurate forecasts (close to IMF and OECD projections, but with a 1 to 3 months lead). Policy-wise, this model can provide an alternative "benchmark" projection for world GDP growth during crisis episodes when sudden swings in the economy make the usual "benchmark" projections (from the IMF or the OECD) rapidly outdated.

Keywords: Nowcasting, Mixed-Frequency Data, High-Frequency Data, World GDP, Large Factor Models

JEL classification: C53, C55, E37

[1] Banque de France, International macroeconomics division, caroline.jardet@banque-france.fr
[2] Banque de France, International macroeconomics division, baptiste.meunier@banque-france.fr
We are grateful to Laurent Ferrara and Clément Marsilli for their seminal nowcasting model of world GDP – still in use at the Banque de France and on which this paper largely builds, Claudia Braz, Franziska Ohnsorge, Sebastian Barnes as well as participants to the 2020 BdF-PSE meeting, the ECB high-frequency workshop, the October NCB G4 international economy meeting, the November extended Global Economy Meeting (GEM) workshop, and internal BdF seminars for useful comments. We thank Aude Le Métayer and Fabien Lebreton for excellent research assistance.

Non-Technical Summary

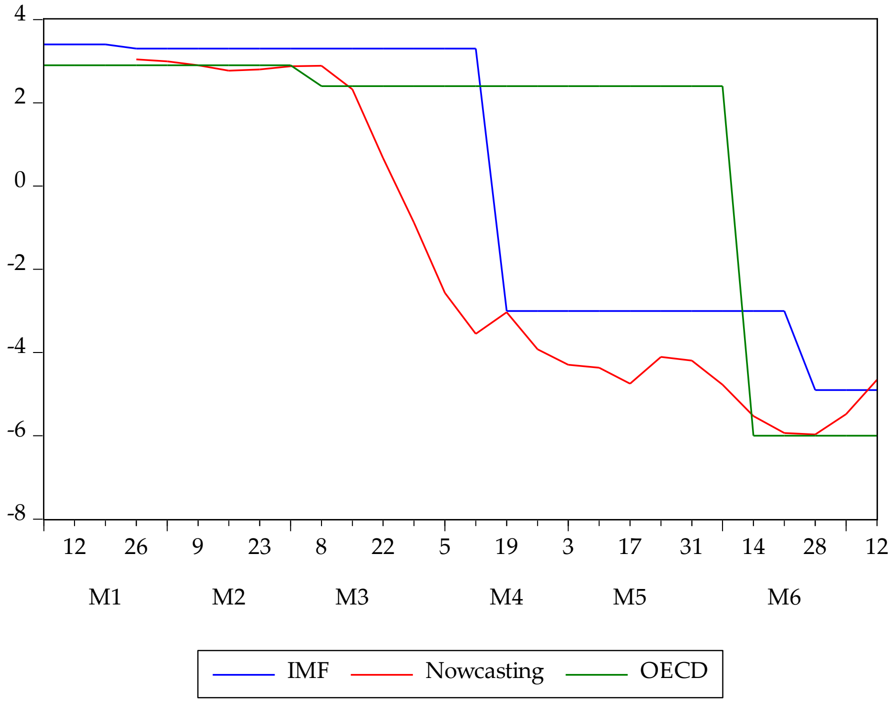
The sudden shock of the Covid-19 crisis has put new emphasis on high-frequency data and a number of weekly, daily, or even hourly data have been extensively used to assess in real-time the impact of the Great Lockdown. In the meantime, world GDP forecasts provided by international organizations such as the OECD and the IMF – which are widely used as "benchmark" projections by economists – have been outdated shortly after their releases given the large and sudden swings in global economic conditions. This has left most macroeconomists with a lack of "benchmark" projections for world GDP. The purpose of this paper is to assess whether high-frequency data can be used to nowcast world GDP growth and therefore provide an alternative "benchmark" projection during those particular times.

Our approach builds on the Factor-Augmented MIxed DAta Sampling (FA-MIDAS) proposed by Marcellino and Schumacher (2008). This set-up suits best our purpose since: *(i)* we mix multiple frequencies by forecasting annual/quarterly GDP growth using monthly macroeconomic variables (e.g. retail sales, PMIs) and weekly indicators (e.g. US jobless claims, stock market indexes); and *(ii)* rely on the aggregation of multiple national variables to make up for the lack of timely global variables. In this vein, we select a large dataset of 151 monthly variables and 39 weekly series out of around 500 monthly and 250 weekly potential regressors. This dataset covers 17 countries representing 68% of world GPD. We then apply a principal component analysis to extract the common trends from these cross-national datasets at both monthly and weekly frequencies. Finally, we run MIDAS regressions to forecast annual/quarterly world GDP growth using monthly and weekly factors as explanatory variables. To test whether high-frequency data enhance performances, we compare performances across different models including (or not) subsets of weekly data, monthly data, or an AR annual/quarterly term. While comparisons are based on in-sample and out-of-sample RMSE, we also test more formally for significant differences in predictive accuracy using the tests developed by Diebold and Mariano (1995).

We find evidence that high-frequency data improve nowcasting performance. Accuracy – both in-sample and out-of-sample – is significantly enhanced when a model includes weekly data relatively to models based only on monthly or quarterly indicators. Performance is at its peak when weekly data is included separately in a specific weekly factor rather than when averaged and incorporated in the monthly factor. This finding is robust to different MIDAS specification and to changes in the dataset. In line with the intuition that high-frequency might be only of second-order when economic conditions are stable, we find that models with weekly data greatly enhance in- and out-of-sample performances during "crisis" episodes (2008-2009 and 2020) but have performances similar to other models during other "normal" periods.

Finally, coming back to our purpose of providing an alternative "benchmark" projection at higher-frequency, we build a nowcasting model for annual world GDP growth using our weekly data. The real-time experience during the Covid-19 crisis shows that this model provides timely estimates for world GDP growth with a 2-3 months lead on IMF and OECD releases (cf. **Figure 1**). It can therefore serve as an alternative "benchmark" projection during those "crisis" episodes when institutional projections are rapidly outdated and when the predictive power of high-frequency data is the greater.

Figure 1. Real-time nowcast and forecasts for the 2020 annual growth rate of world GDP



Prévoir le PIB mondial avec les données haute-fréquence

Résumé

La crise de la Covid-19 a montré que les données à haute-fréquence permettaient un suivi en temps réel de l'activité. Ce papier examine si ces données peuvent aussi améliorer les prévisions du PIB mondial. Nous sélectionnons une large base de 151 séries mensuelles et 39 hebdomadaires couvrant 17 pays avancés et émergents représentant 68% du PIB mondial. Notre approche utilise un *Factor-Augmented MIxed DAta Sampling* (FA-MIDAS) qui permet de combiner différentes fréquences et d'utiliser un grand nombre de variables. Nous montrons que les modèles utilisant des données hebdomadaires ont de meilleures performances que ceux limités aux données mensuelles ou trimestrielles, en estimation comme en prévision. Plus précisément, ces modèles à données hebdomadaires ont des performances similaires aux autres en temps « normal » mais les surclassent largement pendant les périodes de « crise » (2008-2009 et 2020). Nous construisons enfin un modèle de prévision en temps réel du PIB mondial qui peut fournir un point de comparaison aux économistes pendant les périodes de crise – quand les prévisions des institutions (FMI, OCDE) deviennent rapidement obsolètes. Pendant la crise de la Covid-19, notre modèle permet en effet d'anticiper de 1 à 3 mois les prévisions de PIB mondial de ces institutions.

Mots-clés : prévision, temps réel, données haute fréquence, mélange de fréquences

Introduction

The sudden shock of the Covid-19 crisis – with some economies shutting down almost entirely in a matter of days – has put new emphasis on high-frequency data. Weekly, daily, or even hourly data have been extensively used to assess in real-time the impact of the Great Lockdown. A number of innovative dataset have emerged: for example real-time marine traffic was used to track world trade (Cerdeiro *et al.*, 2020), hourly electricity consumption to estimate the loss of industrial activity in Europe (Cicala, 2020), daily credit card spending to quantify the consumption shock (Carvalho *et al.*, 2020), or weekly labour market statistics to model changes in US employment (Coibion *et al.*, 2020). Particularly illustrative of this search for original data, Chetty *et al.* (2020) have developed partnerships with private firms to tap a vast amount of highly granular data at city-level on US employment, household spending and mobility.

In the meantime, world GDP forecasts provided by international organizations such as the OECD or the IMF – which are widely used by economists as "benchmark" projections – have appeared to be lagging behind. These institutions did assemble plausible scenarios and release projections. However, they could only be updated every two or three months – which made them rapidly outdated given the large and sudden changes in economic conditions. Thus, OECD projections as of March 2020 still assumed a positive growth (+2.4%) for world GDP in 2020. Two weeks later when most Western countries entered in a strict lockdown, this scenario was already outdated. Mid-April, the IMF's WEO projected world GDP at -2.9% for 2020. It also rapidly appeared quite optimistic as the virus spread in new geographies and lockdown measures were reinforced. Forecasts were not updated before mid-June, where they stood at -6.0% (OECD) and -4.9% (IMF). Most macroeconomists – who are "projections-takers" – were then facing a lack of "benchmark" projection for world GDP – as those of usual "projection-issuers" became quickly obsolete.

The purpose of this paper is to assess whether high-frequency data can be used to nowcast world GDP growth and therefore provide a timely alternative "benchmark" projection. Our approach builds on the Factor-Augmented MIxed DAta Sampling (FA-MIDAS) proposed by Marcellino and Schumacher (2008). This set-up suits best since: *(i)* we mix multiple frequencies when forecasting annual/quarterly world GDP growth with monthly and weekly series; and *(ii)* we rely on the aggregation of multiple national variables to make up for the lack of global variables. In this vein, we select a large dataset of 12 monthly indicators for 17 representative countries – out of around 500 series for 40 countries. We apply a similar selection to weekly data and build a dataset of 39 weekly series pertaining to financial indicators (e.g. stock market indexes, VIX) and variables for the "real" economy (e.g. US jobless claims, Baltic dry index). We then apply a principal component analysis to extract the common trend out of the monthly and weekly cross-national datasets. Finally, we run MIDAS regressions to model quarterly (or annual) world GDP growth with monthly and weekly factors as explanatory variables.

Comparing across different model specifications, we find evidence that high-frequency data can improve nowcasting performance. Accuracy – both in-sample and out-of-sample – is significantly enhanced when the model includes weekly data relatively to models based only on monthly or quarterly indicators. Performance is at its peak when weekly data is included separately in a weekly factor rather than when it is incorporated in the monthly factor. This finding is robust to different MIDAS specification and to changes in the dataset. More precisely, we find that models with weekly data greatly enhance in- and out-of-sample performances during "crisis" episodes but have performances similar to other models during "normal" periods. Finally, we build a nowcasting model for annual world GDP growth with weekly data. The real-time experience during the Covid-19 crisis shows that this model provides timely estimates of world GDP with a 2-3 months lead on IMF and OECD releases. It might therefore serve as an alternative "benchmark" projection during crisis episodes when institutional projections are rapidly outdated.

Our paper contributes to the literature on MIDAS regression. Based on the seminal work by Ghysels *et al.* (2004), a number of paper have shown that this specification – by allocating different weights to the different lags of high-frequency regressors – performs better than a flat aggregation where high-frequency regressors are averaged at lower-frequency. However, three-frequency models (e.g. Chernis *et al.*, 2020) have remained limited with most of the literature using monthly or quarterly data. Only few have turned to weekly or daily variables and, if so, have generally relied only on financial variables (e.g. Andreou *et al.*, 2013; Banbura *et al.*, 2013; or Ferrara *et al.*, 2014). Our papers extends this literature by including weekly data for the "real" economy, in line with recent papers that deliberately move away from financial series and turn to variables for the "real" economy: e.g. Fed's "Weekly Economic Index" of Lewis *et al.* (2020) or the index for economic growth in low-income countries of Stanger (2020).

Our paper also contributes to the on-going debate of whether high-frequency data enhance forecasting performances. Ferrara *et al.* (2020) showed that a nowcasting model based on high-frequency data produced more accurate forecasts as of end-March 2020 for US growth in Q1 than other models based on standard macroeconomic information. On the other hand, others such as the INSEE (2020) have found no significant gains from including high-frequency data. Our paper shows that nowcasting performances are greatly enhanced by the inclusion of weekly data during "crisis" episodes but not during "normal" times. While the timely signal provided by weekly data allows for a swift detection of turning point when global dynamics experience dramatic changes, the contribution of weekly data is only of second-order when economic conditions are stable.

Our paper finally extends the literature aiming at forecasting world GDP. While some papers have used bridge models (e.g. Golinelli and Parigi, 2014), our paper is in line with those based on large datasets such as Matheson (2011). Closest to us is Ferrara and Marsilli (2019) on which our paper largely builds. Similar to them, we use a FA-MIDAS approach on a large cross-national database to nowcast world GDP; main differences relates to: *(i)* the inclusion of weekly data; *(ii)* the inclusion of a step for selecting variables; *(iii)* the research question on whether high-frequency enhance nowcasting performance; and *(iv)* the purpose to also track GDP at a quarterly frequency.

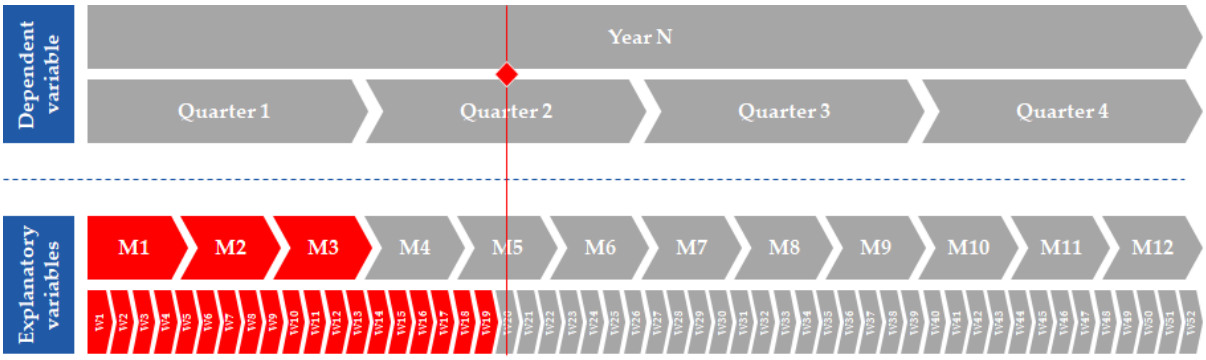
The rest of paper is organised as follows: section 1 presents the data and statistical issues, section 2 details the econometric framework, and results are discussed in section 3.

Section 1: Data

1.1. Dependent variable

The purpose of this paper is the nowcasting of global GDP quarterly (or annual) growth rate y_t. The quarterly GDP series is taken from ECB's macroeconomic projections[1]. Our approach relies on exploiting the infra-quarterly (or infra-annual) information available through monthly or weekly indicators as represented in **Figure 1**: the red square figures a given date – around May 10th in this example – and available information appear in red. Official quarterly growth rates for Q1 are not yet available (they are published at best 45 days after quarter end), monthly indicators are available only until month 3 (they are generally published around 20 days after month end) but weekly data are available up to the preceding week. This shows how timeliness is a strong comparative advantage of weekly data and the main reason why we have considered incorporating it in a nowcasting model.

Figure 1. Approach to nowcasting annual / quarterly world GDP growth



Facing a lack of "world" variables – or their lack of timeliness when such series exist, our approach rely on comparable statistics for a number of representative countries. Our idea is to build a large cross-national dataset from which we can aggregate the information into a few factors by a principal component analysis (PCA)[2]. Once extracted, this common trend can be viewed as a global variable.

1.2. Monthly data

In factor models, Bai and Ng (2008) show that forecasting performances are significantly improved when selecting fewer but more informative predictors[3]. Against this background,

[1] Data is taken from ECB's June 2020 macroeconomic projection exercise.

[2] "Sparse" methods – that select a few explanatory variables out of a large pool of potential regressors – such as the lasso (Tibshirani, 1996) might also be have been applied to this large dataset. However, Giannone *et al.* (2017) find that this type of models perform better only if it can be *a priori* assumed that the data generation process depends only on a small number of regressors. Since it might not *a priori* be the case for world GDP, we instead rely on a "dense" model (the FA-MIDAS) whose principle is that all potential regressors might contribute to prediction, even though their individual impact could be small.

[3] The number of explanatory variables included in the factor model after selection however remains relatively high compared to the number of explanatory that would have been chosen by "sparse" methods (see above).

we implement a data selection process based on four criteria: limited publication lags, sufficient timespan, cross-country availability, and correlation with our target variable[4]. The last criteria is close to the "supervised principal component" method proposed by Bair *et al.* (2006) but we select regressors based on their univariate correlation with our target variable. Due to our additional constraint of cross-country availability, we do not look at the correlation for a unique variable (e.g. Indian retail sales) but at the correlation for the common trend of a cross-country group of similar variables (e.g. the factor obtained by running a PCA on retail sales across all countries). We explore an alternative selection process in which we relax this cross-country availability constraint in **Annex 5**: the selected variables remain broadly the same. In parallel, we also select countries based on three criteria: weight in the global economy, representativeness for a type of economy or geography (e.g. commodity exporters, Africa), and correlation between the domestic cycle and the global cycle. At the end, our sample includes 9 advanced and 8 emerging countries accounting for 68% of world GDP in PPP terms. Among the more than 500 series tested, our dataset finally includes 151 monthly series detailed in **Table 1**.

Table 1. Monthly variables

	Advanced economies									*Emerging economies*							
	US	Japan	Germany	UK	France	Italy	Spain	Canada	Norway	China	India	Russia	Brazil	Mexico	Korea	South Africa	Hong Kong
PPP weight (2019)	*15*	*3.7*	*3.0*	*2.2*	*2.2*	*1.7*	*1.4*	*1.4*	*0.3*	*19*	*8.5*	*2.7*	*2.6*	*1.9*	*1.6*	*0.6*	*0.4*
Retail sales	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
Households' confidence	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	
Car registrations	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of employees	✓	✓	✓	✓		✓	✓	✓	✓			✓	✓		✓		✓
Unemployment rate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
Industrial production	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Manufacturing PMIs	(1)	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓		✓		(2)
Services PMI "headline"	✓		✓	✓	✓	✓	✓			✓		✓					(2)
Composite PMI "headline"	✓		✓	✓	✓	✓	✓			✓		✓					(2)

(1) "Output" and "new export orders" indexes not available

(2) "New export orders" index not available; no sectoral decomposition available

[4] To avoid potential COvid-19 related bias, we exclude 2020 and compute correlations over 1998-2019.

These 151 series relate to 12 different monthly indicators pertaining to various aspects of the economy: households' consumption (retail sales, households' confidence, and car registrations); labour market (number of employees and unemployment rates); and activity in the private sector (industrial production). Our dataset also include Purchase Managers' Index (PMI) which have the double advantage of timeliness – they are generally released the day after month end – as well as high correlation (d'Agostino and Schnatz, 2012) and significant predictive power (e.g. Harris, 1991) for GDP. Closer to us, Lahiri and Monokroussos (2011) find evidence that PMIs improve forecast accuracy in dynamic factor models. In addition to "headline" indexes for both manufacturing and services sectors, we include relevant index for trade ("new export orders" manufacturing PMI) and those capturing the tensions in the production apparatus at an early stage ("new orders" and "output" manufacturing PMIs).

1.3. Weekly data

Similar to monthly data, our selection of weekly series is based on multiples criteria: sufficient timespan, availability, timeliness (publication at least at the end of the following week), and correlation with the target variable. The first criteria has been a severe limitation to the use of innovative "big data", most of them having only a very recent timespan (e.g. since 2009 for marine traffic, since 2020 for mobility data) which makes it challenging for econometrics. The second criteria has also been a strong constraint as multiple data owned by the private sector are either confidential (e.g. credit card data) or costly. In the end, we have tested around 250 series and selected 39 in our dataset.

A first set of weekly data relates to market-based variables – available at a daily frequency but averaged over the week. We include stock market indexes and nominal effective exchange rates for countries in our sample. We also include the Standard & Poor's 1200 index, the VIX computed on the Chicago Board Exchange, and the corporate spread for the US (difference between the 10-year Treasury rate and the BAA rate for corporates).

While the literature generally only considers financial variables at high-frequency (e.g. Andreou *et al.*, 2013), our second set of weekly variables pertains to the "real" economy. We include US new jobless claims, kerosene and gasoline consumption in the US, steel production and capacity utilization in the steel industry also in the US, and finally the US business condition index of Aruoba *et al.* (2009). This might appear US-centric but on the other hand, Kindberg-Hanlon and Sokol (2018) find evidence of a high correlation between US data and world GDP growth. The authors show that only PMIs and industrial production data – already included in our monthly dataset – display a higher degree of correlation with world GDP growth. In addition, we also append indicators pertaining to global dynamics such as the Baltic dry index or commodity prices (Brent oil, gold, platinum, and wheat). Interestingly, Chiu *et al.* (2020) find a high correlation between the Baltic Dry Index and business cycles across BRICS, demonstrating that this data can not only account for trade dynamics but also for economic cycles in emerging economies.

1.4. Handling high-frequency data

An issue arising for monthly data with asynchronous publication lags is that the dataset has a "ragged-edge" pattern: monthly indicators can have different missing elements at the end of the sample – making the monthly dataset unbalanced. To address this, we use the "vertical realignment" procedure introduced by Altissimo *et al.* (2006). For every series, the last available point is taken as the contemporaneous value and the entire series is realigned accordingly. Formally, for a series x_t whose last observation at a contemporaneous date T is at $T-k$, the series becomes $\bar{x}_t = x_{t-k}$. While several issues may be induced by this method – most notably the availability of data determines dynamic cross-correlation between variables and can then change over time, Marcellino and Schumacher (2008) empirically test other methods for handling ragged-edge data[5] and find no substantial changes on the nowcasting performance across methods.

While the "ragged-edge" pattern does not affect our timely weekly dataset, stationarity and seasonality are major concerns[6]. Both issues might be alleviated by taking the annual growth rate of the series as in Ferrara and Simoni (2019) or Lewis *et al.* (2020). While statistically correct, this approach introduces a base effect which might be problematic in our nowcasting: as most indicators suffered a dramatic drop in March 2020, the jump in March 2021 will be symmetrically dramatic and would put at risk the viability of our nowcasting at that point. More broadly, Ladiray *et al.* (2018) discuss the drawbacks of taking annual growth for weekly data and point out that it not only include a phase shift by design but also can introduce spurious cycles.

To alleviate these concerns, we use a two-step procedure to obtain de-seasonalised and stationary weekly indicators. In the first step, the series are transformed in their average weekly variation over the last four weeks. This transformation – equivalent to a moving monthly growth – have the double advantage of making the series stationary and correcting for infra-monthly seasonality. Then the transformed series is regressed on monthly dummies; the final series is the residual of the regression. This last step allows us to correct for any monthly (or lower-frequency) seasonality. All-in-all, this procedure allows us to get de-seasonalised and stationary weekly series while alleviating aforementioned concerns on phase shift or spurious cycles.

[5] The authors also test for the EM-algorithm of Stock and Watson (2002) and the Kalman smoother estimates of Doz *et al.* (2006).

[6] It is worth nothing that standards methods for de-seasonalisation (Census X-11 and following) are not applicable to weekly data.

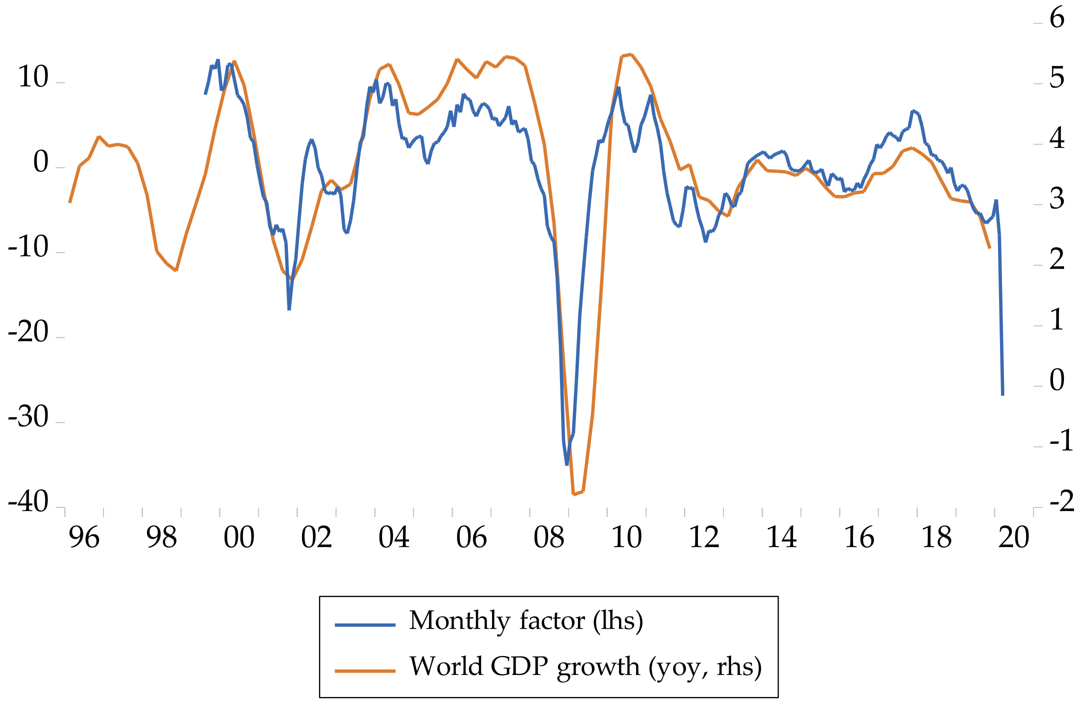
Section 2: Econometric framework

Our econometric framework is based on the two-step FA-MIDAS approach proposed by Marcellino and Schumacher (2008) and used in Ferrara and Marsilli (2019). The first step is a principal component analysis on our large monthly and weekly datasets. Doing so, we extract the common trends at both weekly and monthly frequencies. Formally, we assume that our balanced dataset X_T can be represented according to a factor structure with a r-dimensional factor vector F_T, Λ the loadings matrix and idiosyncratic components ξ_T not explained by the common factors. The common components $(\Lambda \cdot F_T)$ and the idiosyncratic components are mutually orthogonal.

$$X_T = \Lambda \cdot F_T + \xi_T$$

The first factors for the monthly and weekly datasets are displayed in **Figures 2 and 3**. Both appear to track adequately world GDP and to be leading indicators of turning points in the global economy with a 2-3 month lead. While figure 3 seems to indicate that the weekly factor detects adequately dramatic swings in the world economy (as in 2009 and 2020), we present the weekly factor smoothed with a Hodrick-Prescott (HP) filter in **Figure 4**[7]. This HP-filtered factor shows more clearly the capacity of the weekly factor to also track shallower economic changes.

Figure 2. Monthly factor and world GDP growth



[7] HP-filtered series is only displayed for illustrative purpose. Throughout our estimations, we use the raw factor.

Figure 3. Weekly factor and world GDP growth

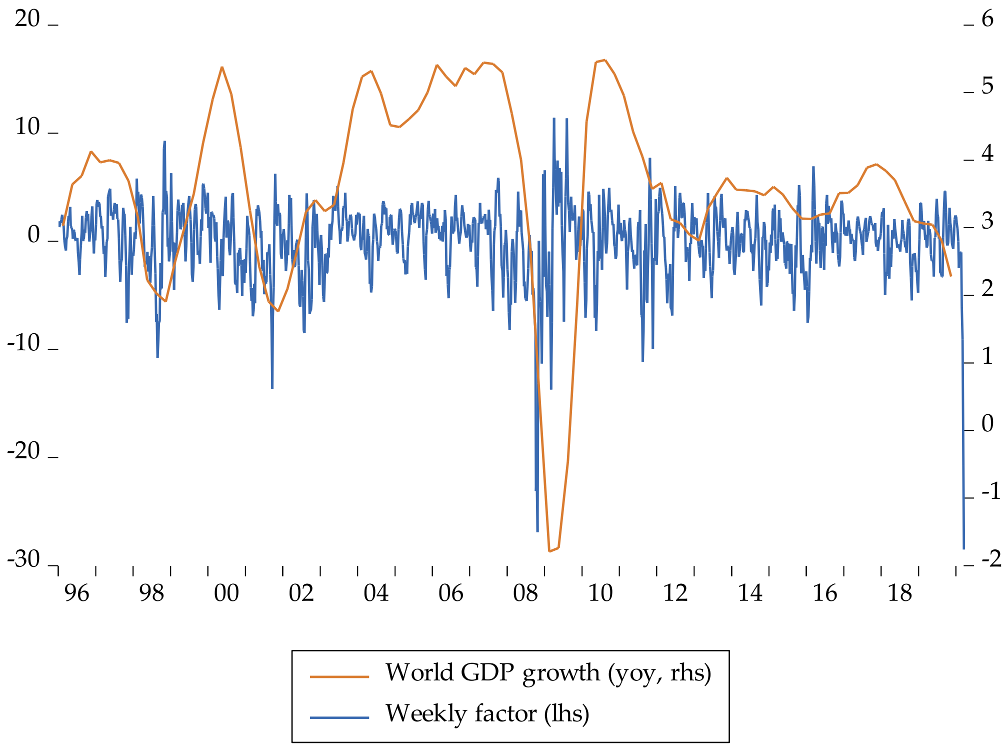
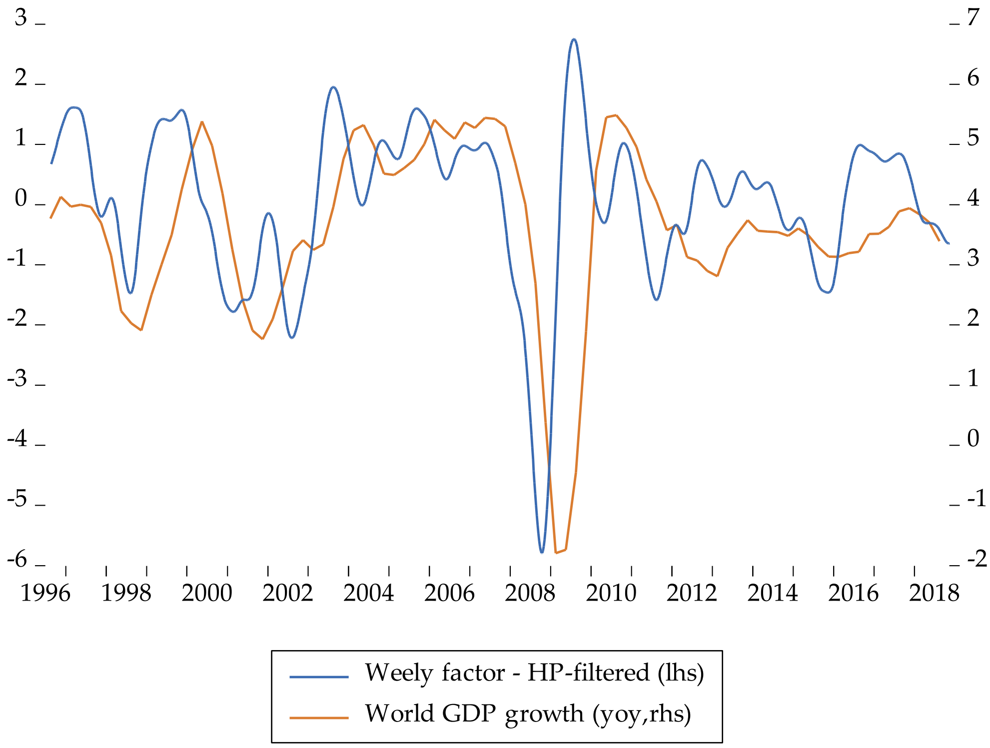


Figure 4. Weekly factor (HP-filtered) and world GDP growth



Once monthly and weekly factors have been extracted, the second step is the modelling in a MIDAS specification in which the dependent variable is the quarterly growth rate of world GDP y_t. Explanatory variables are a quarterly constant β_0 as well as the aforementioned monthly $f^m_{t/_3}$ and weekly $f^w_{t/_{13}}$ factors. K represents the number of high-frequency lags and θ is a vector of parameters to be estimated in the MIDAS weighting function g.

$$y_t = \beta_0 + g\left(f^m_{t/_3}, \theta^m, K^m, \dots\right) + g\left(f^w_{t/_{13}}, \theta^w, K^w, \dots\right) + \varepsilon_t$$

We run regression using three types of MIDAS weighting functions g:

- "Almon" of degree p: the coefficients at each lag of the high-frequency regressor are modelled through a polynomial function of degree $p-1$. This specification is the most parsimonious since only p parameters are estimated and our baseline specification in the rest of the paper is an "Almon" of degree 3. Formally, the weighting function is:

$$g(f_t, \theta, K, p) = \sum_{k=0}^{K} c(k,\theta) \cdot f_{t-k} \; where \; c(k,\theta) = \sum_{j=0}^{p-1} k^j \cdot \theta_j$$

- U-MIDAS (Foroni *et al.*, 2012): no particular structure is applied to individual coefficients for high-frequency lags. The model has then to estimate a large number of $K^m + K^p$ parameters.

- "Step": coefficients for high-frequency lags follow a step function: they are grouped by η sharing the same value θ_i. For example if $\eta = 5$, the first five lags will all have a coefficient θ_0, the following five will have θ_1 and so on. While the number of parameters to be estimated $(ceil(K/\eta))$ increases with the number of lags K, it is reduced by roughly η compared to the estimation of a U-MIDAS. Formally:

$$g(f_t, \theta, K, \eta) = \sum_{k=0}^{K} \theta_{floor(k/\eta)} \cdot f_{t-k}$$

For "Almon" and "step" MIDAS, the optimal number of lags is determined using the minimal sum-of-squared residuals as the selection criterion. Such a criteria however cannot apply in a "U-MIDAS"; for this specification we then use the optimal number of lags determined for the "Almon" MIDAS – the most parsimonious model – to mitigate overfitting concerns.

To test whether high-frequency data improves predictive accuracy of the nowcasting, we compare six different models:

- **Model 1** includes the monthly factor (f_t^m) and two weekly factors[8], one pertaining to financial variables ($f_t^{w,fin}$) and the other to variables for the "real" economy ($f_t^{w,real}$);

- **Model 2** includes the monthly factor (f_t^m) but a unique weekly factor incorporating both financial and "real" variables ($f_t^{w,all}$). Comparing model 1 *vs.* model 2 allows us to test whether separating financial and "real" variables can improve performances;

- **Model 3** includes the monthly factor (f_t^m) and only the weekly factor based on financial variables ($f_t^{w,fin}$). Comparing model 1 (or 2) *vs.* model 3 allows us to test whether including variables for the "real" economy brings significant value *vs.* a nowcasting based only on financial variables – as is generally done in the literature;

- **Model 4** includes a unique monthly factor which also incorporates the weekly variables but averaged over the month ($f_t^{m,w}$). Comparing model 1 (or 2 or 3) *vs.* model 4 allows us to test whether a three-frequency model performs better than a two-frequency model where weekly indicators are averaged at a monthly frequency;

- **Model 5** includes the baseline monthly factor (f_t^m) based only on monthly indicators. Comparing model 1 (or 2 or 3) *vs.* model 5 allows us to test whether weekly data can improve nowcasting performances. In addition, comparing model 4 *vs.* model 5 allows us to test whether weekly data – even averaged at a monthly frequency – can still improve accuracy;

- **Model 6** is an AR model with the latest available data point of quarterly world GDP.

Model comparisons are based on root mean squared errors (RMSE). Errors are computed over 2005Q1 to 2020Q1 – so as to capture both the Great Financial Crisis and the Great Lockdown. Out-of-sample errors relates to the one-period ahead forecast errors: the initial estimation sample is 1998Q4 to 2004Q4, and therefore our initial forecast is 2005Q1; then the estimation sample is extended quarter after quarter following an "expanding window" procedure.

To reflect the fact that nowcasting is done in real-time, estimations are replicated at the first, second, and third month of the quarter. The information available differs for each month m: we generally consider that the nowcast takes place around the 15th of month $m+1$. Quarterly variable (AR term) is then available from the 2nd month onwards – in line with an average publication lag of 45 days for GDP statistics. Monthly variables are available up to the month[9]

[8] Instead of separating into two subsets prior to running the PCA, another approach might have been to take two factors after running the PCA on the entire weekly dataset. This data-driven approach – explored in **Annex 3** – however do not yield better performances than the *ex ante* separation between two subsets.

[9] The "vertical realignment" procedure of Altissimo *et al.* (2006) is particular convenient at this stage since it transposes past publication lags. If publication lags are constant over time, the value at any month m_0 in the past of the factor estimated contemporaneously at $m = m_0 + \tau$ is the exact same value as what would have been estimated with information available at m_0 – not accounting for the revisions between m_0 and m.

m. Finally, given their timeliness, weekly variables are available up the 2nd week of the month $m + 1$. Data availability is recapitulated in **Table 2** below.

Table 2. Data availability for each month in the quarter

	1st month	*2nd month*	*3rd month*
Quarterly variable (AR term)	2 quarters lag	1 quarter lag	1 quarter lag
Monthly variables	Up to 1st month	Up to 2nd month	Up to 3rd month
Weekly variables	Up to 2nd week of 2nd month	Up to 2nd week of 3rd month	Up to final week of 3rd month

Section 3: Results

3.1. Baseline results

In-sample results for the baseline specification are reported in **Table 3** below. They suggest that weekly data improves performances. Models with weekly factors (1, 2, and 3) perform better than models with only a monthly factor (4 and 5) at all months of the quarter. They also largely outperform the benchmark AR model (6). Weekly data also seem to enhance performances when averaged at monthly frequency: model 4 – where weekly data is averaged over the month and included in the monthly factor – slightly outperform model 5 which rely only on monthly data. Among models with weekly data, it appears that model 1 – which includes two distinct weekly factors, one for financial indicators and the other for variables of the "real" economy – performs better than models with a unique weekly factor (model 2). Finally, the results indicate that including weekly series for the "real" economy (as in models 1 and 2) yields better performances than forecasts relying solely on financial variables at the weekly frequency (as in model 3).

Table 3. In-sample performances (RMSE) across models and months of the quarter

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
1st month	**0.442**	0.476	0.483	0.529	0.530	0.689
2nd month	**0.211**	0.270	0.282	0.463	0.466	0.692
3rd month	**0.202**	0.230	0.232	0.271	0.285	0.692

Grey cells indicate best performance for a given month

Out-of-sample results in **Table 4** below tend to the same conclusion. Models with weekly data (1, 2, and 3) still outperform those with only monthly factors (model 4 and 5) or an AR term (6). This supports the conclusion that weekly data can improve the nowcasting performances in a MIDAS set-up – both in-sample and out-of-sample.

Table 4. Out-of-sample performances (RMSE) across models and months of the quarter

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
1st month	0.599	**0.557**	0.559	0.559	0.562	0.868
2nd month	**0.422**	0.455	0.468	0.512	0.516	0.589
3rd month	**0.351**	0.430	0.429	0.425	0.447	0.589

Grey cells indicate best performance for a given month

To confirm these results, we formally test for significant difference in predictive accuracies across the different models using the Diebold and Mariano (1995) test. The results – reported in **Table A1.1 in Annex 1** – confirm that high-frequency data significantly improve forecasting performance. Models with weekly factors have a significantly better predictive accuracy than models with only monthly factors or quarterly variables (i.e. model 1 outperforms models 4, 5 and 6). In addition, high-frequency data enhance accuracy not only when they are included at weekly frequency (as in model 1) but also when they are averaged and incorporated in the monthly factor (as in model 4): not only model 1 outperforms model 5, but model 4 also does. Among models with weekly data, the results also shows that incorporating two separate weekly factors (model 1) perform significantly better than when aggregating weekly data into a single factor (model 2) or restricting the weekly dataset to financial variables (model 3).

3.2. Alternative specifications and datasets

For robustness, we run the same regressions using alternative MIDAS specifications: "Step" and "U-MIDAS". The results are presented in **Table A1.2 in Annex 1** and confirm those discussed in the previous section. Models with weekly factors (1, 2, and 3) outperform others both in-sample and out-of-sample. Among models with weekly factors, models incorporating weekly indicators for the "real" economy (models 1 and 2) show greater accuracy than the model 3 which only include financial variables at the weekly frequency.

Our baseline results are also robust to changes in the dataset. In **Annex 2**, we explore the role of PMIs in the monthly factor. While we find evidence that PMIs explain a large part of the monthly factor[10], we show that an alternative factor computed without the PMIs presents the same evolution as our baseline factor. More importantly, when running the same regressions with the alternative monthly factor that excludes PMIs, our findings remain valid.

Finally, in the **Annex 4** we disentangle between emerging and advanced economies and perform the FA-MIDAS approach on the two subsets. Our baseline results remain valid on both sub-groups: models with weekly data still outperform others in- and out-of-sample.

3.3. "Crisis" vs. "normal" periods

There is however a general trade-off between timeliness and accuracy for high-frequency data as put forward by Ahnert and Bier (2001). High-frequency data have the potential to provide a very timely signal which might significantly enhance the nowcasting performance when economic conditions suddenly deteriorate. In other terms, during "crisis" episodes, weekly data can make up for the long publication lags in standard monthly indicators. During "normal" periods however, the contribution of high-frequency data might only be of second-order relatively to monthly data. In addition, high-frequency data are generally not statistically adjusted and can prove rather noisy.

[10] This is expected: when performing a PCA on a dataset combining non-transformed series (PMIs) and transformed (taken as first differences or monthly growth rates), the non-transformed series generally explain a large part of the variance. However in our dataset, some other series – notably statistics for industrial production – are also strongly correlated with the factor.

Against this background, we distinguish our sample between "normal" and "crisis" episodes, and then compute in-sample and out-of-sample RMSE for both sub-samples[11]. Results are reported in **Table 6** where it appears that models with weekly data (1, 2 and 3) heavily outperform others during "crisis" episodes. During those periods, the upside of providing a very timely signal exceeds the downside of the noise in the weekly data. However, during "normal" periods, models with weekly factors (models 1, 2, and 3) have performances similar to those of models relying on monthly data (4 and 5) or on quarterly data once the data point for the preceding quarter is published (after month 2 in model 6).

Table 6. RMSE across models, months of the quarter, and periods

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Crisis episodes						
In-sample						
1st month	**1.057**	1.161	1.186	1.363	1.364	1.635
2nd month	**0.264**	0.387	0.431	1.102	1.139	1.433
3rd month	**0.271**	0.362	0.359	0.547	0.584	1.433
Out-of-sample						
1st month	1.540	**1.439**	1.444	1.451	1.456	1.709
2nd month	**1.007**	1.140	1.179	1.305	1.324	1.545
3rd month	**0.803**	1.064	1.066	1.063	1.130	1.545
Non-crisis episodes						
In-sample						
1st month	0.236	0.239	0.238	**0.204**	0.206	0.300
2nd month	**0.202**	0.248	0.252	0.251	0.232	0.203
3rd month	**0.189**	0.203	0.207	0.198	0.206	0.203
Out-of-sample						
1st month	0.235	0.212	0.211	**0.205**	0.208	0.318
2nd month	0.227	0.206	0.207	0.210	0.203	**0.197**
3rd month	0.211	0.205	0.202	0.194	**0.192**	0.197

Grey cells indicate best performance for a given month

[11] Regressions are still estimated over full sample – or for out-of-sample over the sample preceding target quarter.

3.4. Annual GDP forecasts

Back to the issue highlighted in the introduction, the main concern for "projection-takers" macroeconomists during the Covid-19 crisis has been the rapid obsolescence of "benchmark" world GDP forecasts. Building on our results that weekly data significantly improve forecast performance, we construct a nowcasting model for the annual growth rate of world GDP based on our baseline MIDAS specification. This model includes the latest available data point for quarterly world GDP, the monthly factor (f_t^m) and a unique weekly factor ($f_t^{w,all}$) which combines financial indicators and variables for the "real" economy[12].

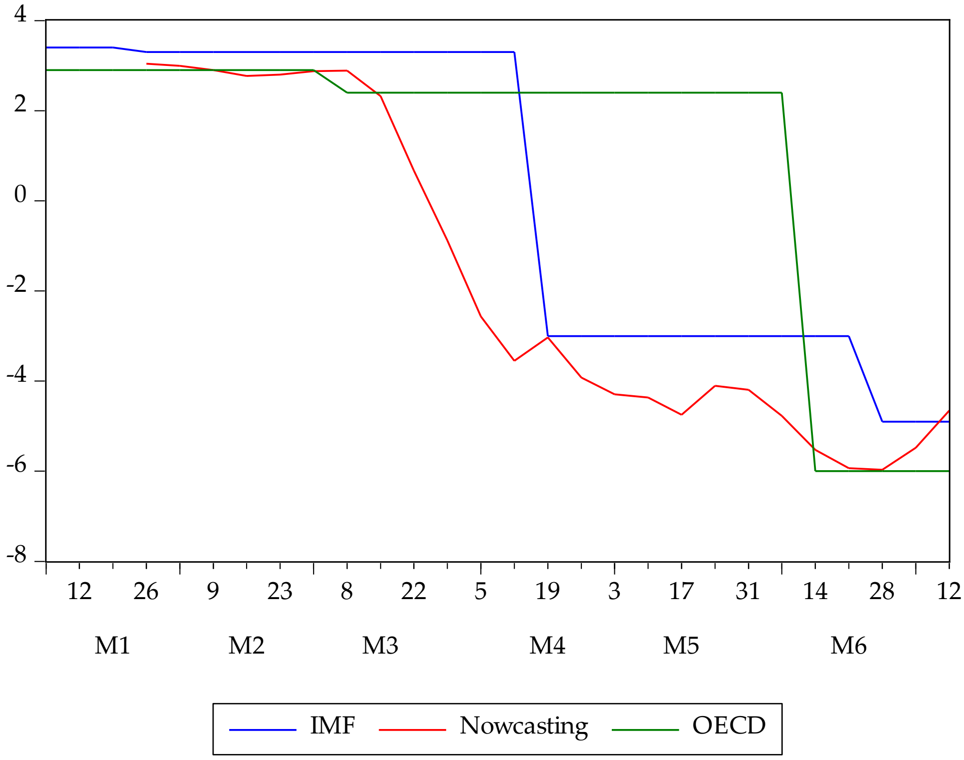
We start by comparing performances across different models to test if weekly data still enhance the predictive accuracy for annual forecasts. We run a "real-time" comparison: factors and regression are re-estimated at the end of each week with the information available at this stage, so it mimics what could have been done in real-time. In-sample RMSE are reported in **Table A1.3 in Annex 1** where we compare four models: our baseline model with the latest available data point of quarterly world GDP, the monthly factor, and weekly factor (model 2b); one without weekly data but still including quarterly and monthly regressors (model 4b); one with only latest available data point of quarterly world GDP as regressor (model 7); and an AR annual model (model 6b). The model 2b appears to outperform other models for all weeks of the year. In line with the intuition, the largest improvements in performance are met in the first weeks of the year when quarterly and monthly information are not yet available.

We finally compare in the **Figure 5** the "real-time" nowcasting results of our model to the projections released by the OECD and the IMF – updated every two or three months. To account for a certain volatility in the nowcasting results, the series in the graph is the moving average over four weeks of raw results. This figure shows that the nowcasting has the advantage of greater timeliness and decent accuracy compared with IMF's or OECD's forecasts. For example, as soon as of mid-April 2020, our nowcasting indicates that the recession in 2020 would fluctuate around -4% to -5% before IMF's and OECD's indicated the same estimates by mid-June.

This would suggest that this nowcasting model based on high-frequency data can provide an alternative "benchmark" projection during crisis episodes, when it is most needed due the obsolescence of usual benchmark projections and when high-frequency significantly and greatly enhance nowcasting performances.

[12] Although our previous results indicate that including two separate weekly factors yields better performances, we prefer using only a unique factor to alleviate overfitting concerns given the limited number of observations.

Figure 5. Real-time nowcast and forecasts for the annual growth rate of world GDP



Conclusion

Our paper provides evidence that including high-frequency (weekly) variables significantly improve the predictive accuracy in a MIDAS specification. In-sample and out-of-sample performances are greater for models with weekly data – a finding robust to different MIDAS specifications and changes in the dataset. This is formally confirmed by Diebold-Mariano tests which shows that these accuracy gains are significant. Higher accuracy is reached in a three-frequency model where weekly factors are introduced alongside a monthly factor. Finally, it appears that adding weekly variables for the "real" economy (consumption of oil products, new jobless claims, etc.) enhance nowcasting accuracy compared to models relying only on financial indicators at the weekly frequency.

These findings are driven by "crisis" episodes during which the benefits of including weekly data are highly significant in terms of nowcasting accuracy. During "normal" periods when the global economy does not experience sudden and abrupt changes however, models with weekly data have performances similar to those of models based on monthly data.

This drives us to build a real-time nowcasting of annual world GDP growth which has the double advantage of: *(i)* timeliness as it provides a new forecast every week while institutional forecasts are only updated three or four times per year; *(ii)* accuracy since it produce forecasts close to IMF's or OECD's projections but with a 1 to 3 month lead. This can provide an alternative "benchmark" projection of world GDP to macroeconomists during crisis episodes when the contribution of weekly data is highly significant and when the institutional projections – usually taken as "benchmark" by macroeconomists – are rapidly outdated given the dramatic changes that have occurred since their releases.

An avenue for future research could be the inclusion of privately-held high-frequency data such as the innovative dataset mentioned in the introduction (credit card data, marine traffic) even though limited timespan (e.g. mobility data only starts in January 2020) or costs (e.g. using marine traffic requires several thousand euros) are strong limitations. Including these innovative big data also makes the case for using machine learning techniques better suited to handle vast amounts of data such as the Bayesian MIDAS penalized regressions introduced by Mogliani and Simoni (in press).

Finally, our approach can be extended to nowcast GDP for a specific country or group of countries (e.g. euro area). Narrowing the target might increase the scope of available data since one of our constraint was to build on comparable statistics across several countries. In addition, making specific nowcast can pave the way for a "bottom-up" forecasting approach as the one we tentatively explored in the **Annex 4** which distinguish two separate nowcasting models for emerging and advanced economies.

References

d'Agostino A. and Schnatz B. (2012). "Survey-based nowcasting of US growth: a real-time forecast comparison over more than 40 years", *European Central Bank Working Paper Series*, No 1455

Ahnert H. and Bier W. (2001). "Trade-off between timeliness and accuracy", *Economisch Statistische Berichten*, No 4299

Altissimo F., Cristadoro R., Forni M., Lippi M., and Veronese G. (2006). "New Eurocoin: Tracking Economic Growth in Real Time", *CEPR Discussion Papers*, No 5633

Andreou E., Ghysels E., and Kourtellos A. (2013). "Should Macroeconomic Forecasters Use Daily Financial Data and How?", *Journal of Business and Economic Statistics*, 31(2), pp. 240-251

Aruoba S., Diebold F., and Scotti C. (2009). "Real-Time Measurement of Business Conditions", *Journal of Business and Economic Statistics*, 27(4), pp. 417-427

Bai J. and Ng S. (2008). "Forecasting economic time series using targeted predictors", *Journal of Econometrics*, 146(2), pp. 304-317

Bair E., Hastie T., Paul D., and Tibshirani R. (2006). "Prediction by supervised principal components", *Journal of the American Statistical Association*, 101(473), pp. 119-137

Bańbura M., Giannone D., Reichlin L., and Modugno M. (2013). "Now-casting and the real-time data flow", *European Central Bank Working Paper Series*, No 1564

Carvalho V., Garcia J., Hansen S., Ortiz Á., Rodrigo T., Rodríguez Mora J., and Ruiz J. (2020). "Tracking the COVID-19 Crisis with High-Resolution Transaction Data", *Cambridge-INET Working Papers*, No 2016

Cerdeiro D., Komaromi A., Liu Y., and Saeed M. (2020). "World Seaborne Trade in Real Time: A Proof of Concept for Building AIS-based Nowcasts from Scratch", *International Monetary Fund Working Papers*, No 20/57

Chiu C., Chen C., Kuo P.-L., and Wang M. (2020). "The Dynamic Relationships between the Baltic Dry Index and the BRICS Stock Markets: A Wavelet Analysis", *Asian Economic and Financial Review*, 10(3), pp. 340-351

Chernis T., Cheung C., and Velasco G. (2020). "A three-frequency dynamic factor model for nowcasting Canadian provincial GDP growth", *International Journal of Forecasting*, 36(3), pp. 851-872

Chetty R., Friedman J., Hendren N., Stepner M., and The Opportunity Insights Team (2020). "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data", *NBER Working Paper*, No 27431

Cicala S. (2020). "Early Economic Impacts of COVID-19 in Europe: A View from the Grid", mimeo

Coibion O., Gorodnichenko Y., and Weber M. (2020). "Labor Markets During the COVID-19 Crisis: A Preliminary View", *University of Chicago Working Papers*, No 2020-041

Diebold F. and Mariano R. (1995). "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 13, pp. 253-263

Doz, C., Giannone D., and Reichlin L. (2006). "A quasi maximum likelihood approach for large approximate dynamic factor models", *European Central Bank Working Paper Series*, No 674

Ferrara L. and Marsilli C. (2019). "Nowcasting global economic growth: A factor-augmented mixed-frequency approach", *The World Economy*, 42(3), pp. 846-875

Ferrara L., Marsilli C., and Ortega J.-P. (2014). "Forecasting growth during the Great Recession: is financial volatility the missing ingredient?", *Economic Modelling*, 36, pp. 44–50

Ferrara L. and Simoni A. (2019). "When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage", *Center for Research in Economics and Statistics (CREST) Working Papers*, No 2019-04

Ferrara L., Froidevaux A., and Huynh T-L. (2020). "Macroeconomic nowcasting in times of Covid-19 crisis: On the usefulness of alternative data", *Econbrowser*, https://econbrowser.com/archives/2020/03/guest-contribution-macroeconomic-nowcasting-in-times-of-covid-19-crisis-on-the-usefulness-of-alternative-data

Foroni C., Marcellino M., and Schumacher C. (2012). "U-MIDAS: MIDAS regressions with unrestricted lag polynomials", *CEPR Discussion Papers*, No 8828

Ghysels E., Santa-Clara P., and Valkanov R. (2004). "The MIDAS Touch: Mixed Data Sampling Regression Models", *CIRANO Working Paper*, No 2004-20

Giannone D., Lenza M., and Primiceri G. (2017). "Economic Predictions with Big Data: The Illusion Of Sparsity", *CEPR Discussion Papers*, No 12256

Golinelli R. and Parigi G. (2014). "Tracking world trade and GDP in real time", *International Journal of Forecasting*, 30(4), pp. 847-862

Harris E. (1991). "Tracking the Economy with the Purchasing Managers' Index", *Federal Reserve Bank of New York Quarterly Review*, Autumn

INSEE (2020). "Les données « haute fréquence » sont surtout utiles à la prévision économique en période de crise brutale", *Note de conjoncture de l'INSEE*, June 17th, https://www.insee.fr/fr/statistiques/4513034?sommaire=4473296

Kindberg-Hanlon G. and Sokol A. (2018). "Gauging the globe: the Bank's approach to nowcasting world GDP", *Bank of England Quarterly Bulletin*, Q3

Ladiray D., Palate J., Mazzi G., and Proietti T. (2018). *"Seasonal Adjustment of Daily and Weekly Data"*, in 16th Conference of the International Association of Official Statisticians, Paris, September 19th-21st, https://www.oecd.org/iaos2018/programme/IAOS-OECD2018_Item_1-A-1-Ladiray_et_al.pdf

Lahiri K. and Monokroussos G. (2011). "Nowcasting US GDP: The role of ISM Business Surveys", *University at Albany Discussion Papers*, No 11-01

Lewis D., Mertens K., and Stock J. (2020). "Monitoring Real Activity in Real Time: The Weekly Economic Index", *Liberty Street Economics*, No 20200330b, Federal Reserve Bank of New York

Matheson T. (2011). "New Indicators for Tracking Growth in Real Time", *International Monetary Fund Working Papers*, No 11/43

Marcellino M. and Schumacher C. (2008). "Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP", *CEPR Discussion Papers*, No 6708

Mogliani M. and Simoni A. (in press). "Bayesian MIDAS penalized regressions: estimation, selection, and prediction", *Journal of Econometrics*

Stanger M. (2020). "A Monthly Indicator of Economic Growth for Low Income Countries", *International Monetary Fund Working Papers*, No 20/13

Stock J. and Watson M. (2002). "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20, pp. 147-162

Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society*, 58(1), pp. 267–288

Annex 1: Additional tables

Table A1.1. Diebold-Mariano test results

	Model 1 *vs.* model 2	Model 1 *vs.* model 3	Model 1 *vs.* model 4	Model 1 *vs.* model 5	Model 4 *vs.* model 5	Model 1 *vs.* model 6
H_0 = both forecasts have the same accuracy						
1st month	0.87	0.87	0.62	0.84	0.36	**0.03**
2nd month	**0.04**	**0.03**	**0.09**	**0.10**	0.78	**0.00**
3rd month	**0.04**	**0.04**	**0.03**	**0.02**	**0.06**	**0.00**
H_0 = model A have lower accuracy than model B						
1st month	0.57	0.56	0.69	0.58	0.18	**0.02**
2nd month	**0.02**	**0.02**	**0.05**	**0.05**	0.60	**0.00**
3rd month	**0.02**	**0.02**	**0.01**	**0.01**	**0.03**	**0.00**

Results report p-value
Grey cells indicate where the null hypothesis can be rejected at a 10% significance

Table A1.2. Performances (RMSE) across models, specifications, and months

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
In-sample						
Step						
1st month	**0.507**	0.716	0.740	0.561	6.055	0.689
2nd month	**0.258**	0.355	0.376	0.563	0.529	0.692
3rd month	**0.246**	0.394	0.416	0.440	5.667	0.692
U-MIDAS						
1st month	**0.440**	0.489	0.498	0.530	0.531	0.689
2nd month	**0.201**	0.270	0.282	0.463	0.464	0.692
3rd month	**0.198**	0.221	0.223	0.259	0.273	0.692
Out-of-sample						
Step						
1st month	0.892	0.913	0.925	**0.599**	0.629	0.686
2nd month	**0.425**	0.441	0.466	0.553	0.574	0.589
3rd month	**0.392**	0.430	0.418	1.957	0.503	0.589

U-MIDAS						
1st month	**0.543**	0.558	0.561	0.561	0.563	0.686
2nd month	**0.405**	0.456	0.468	0.515	0.519	0.589
3rd month	**0.381**	0.406	0.413	0.423	0.447	0.589

Grey cells indicate best performance for a given month

Table A1.3. In-sample RMSE across models and weeks of the year

	Model 2b	Model 4b	Model 7	Model 6b
1st week	**0.394**	0.531	0.573	1.282
2nd week	**0.421**	0.531	0.573	1.282
3rd week	**0.411**	0.469	0.573	1.282
4th week	**0.402**	0.469	0.573	1.282
5th week	**0.365**	0.469	0.573	1.282
6th week	**0.382**	0.453	0.495	1.282
7th week	**0.345**	0.446	0.495	1.256
8th week	**0.290**	0.446	0.495	1.256
9th week	**0.332**	0.446	0.495	1.256
10th week	**0.306**	0.446	0.495	1.256
11th week	**0.189**	0.446	0.495	1.256
12th week	**0.214**	0.404	0.495	1.256
13th week	**0.210**	0.404	0.495	1.256
14th week	**0.279**	0.404	0.495	1.256
15th week	**0.304**	0.404	0.495	1.256
16th week	**0.244**	0.433	0.495	1.256
17th week	**0.292**	0.433	0.495	1.256
18th week	**0.308**	0.433	0.495	1.256
19th week	**0.307**	0.388	0.404	1.256
20th week	**0.295**	0.349	0.404	1.256
21st week	**0.286**	0.349	0.404	1.256
22nd week	**0.220**	0.349	0.404	1.256
23rd week	**0.209**	0.349	0.404	1.256
24th week	**0.197**	0.349	0.404	1.256

Table A1.3 (continued). In-sample RMSE across models and weeks of the year

25th week	**0.213**	0.308	0.404	1.256
26th week	**0.210**	0.308	0.404	1.256
27th week	**0.160**	0.308	0.404	1.256
28th week	**0.170**	0.308	0.404	1.256
29th week	**0.135**	0.178	0.404	1.256
30th week	**0.131**	0.178	0.404	1.256
31st week	**0.136**	0.178	0.404	1.256
32nd week	**0.115**	0.146	0.222	1.256
33rd week	**0.130**	0.141	0.222	1.256
34th week	**0.121**	0.141	0.222	1.256
35th week	**0.095**	0.141	0.222	1.256
36th week	**0.125**	0.141	0.222	1.256
37th week	**0.128**	0.141	0.222	1.256
38th week	**0.110**	0.146	0.222	1.256
39th week	**0.094**	0.146	0.222	1.256
40th week	**0.081**	0.146	0.222	1.256
41st week	**0.083**	0.146	0.222	1.256
42nd week	**0.082**	0.123	0.222	1.256
43rd week	**0.083**	0.123	0.222	1.256
44th week	**0.086**	0.123	0.222	1.256
45th week	**0.035**	0.054	0.065	1.256
46th week	**0.036**	0.059	0.065	1.256
47th week	**0.034**	0.059	0.065	1.256
48th week	**0.032**	0.059	0.065	1.256
49th week	**0.039**	0.059	0.065	1.256
50th week	**0.037**	0.059	0.065	1.256
51st week	**0.041**	0.053	0.065	1.256
52nd week	**0.040**	0.053	0.065	1.256

Grey cells indicate best performance for a given week

Annex 2: The role of PMIs in the monthly factor

When performing our principal component analysis (PCA), we combine non-transformed and transformed series. All series but PMIs are either transformed into month-on-month percentage change or first differences. PMIs are untouched since they are already computed as diffusion indices on surveys of how business conditions have evolved compared to last month. An issue that might arise when combining transformed and non-transformed is that the resulting factor might be almost exclusively driven by the non-transformed series. Factor loadings in **Table A2.1** below indicate that PMIs indeed explain a large part of the variance even though a number of other series also fit well (e.g. US industrial production, Chinese households' confidence, or German employment).

Table A2.1. Factor loadings

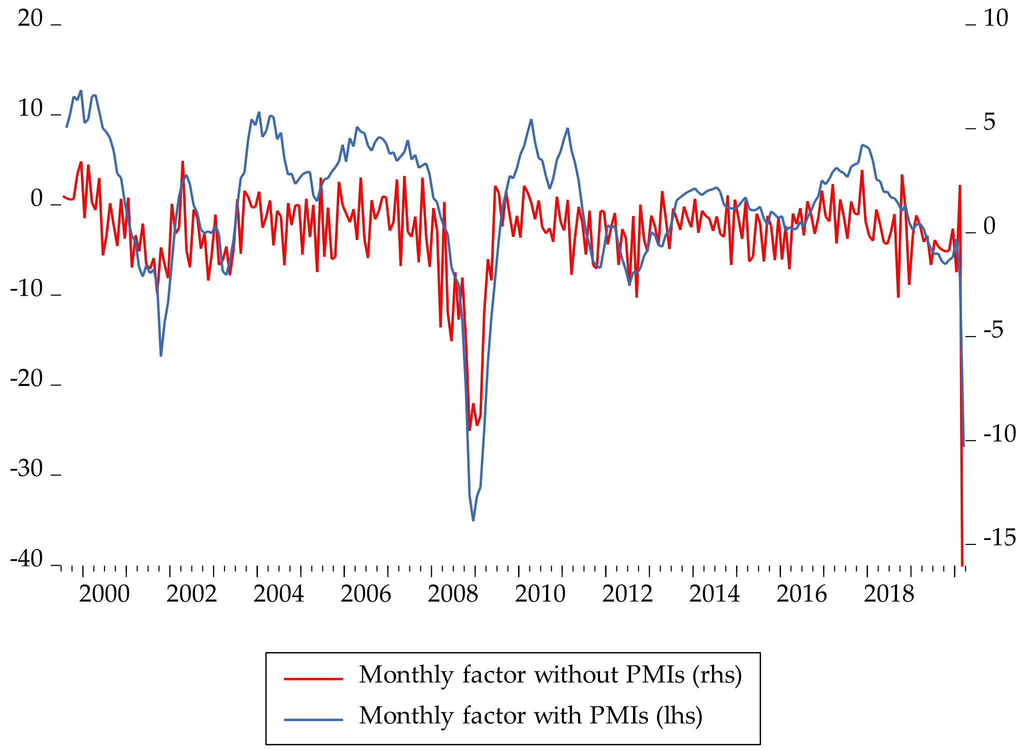
Series name	Loading	Series names (cont.)	Loading (cont.)	Series names (cont.)	Loading (cont.)
BR_CAR	0.022142	MX_CAR	0.036261	PMI_HK_MAN_NI	0.104437
BR_CNF	0.051912	MX_PRO	0.052720	PMI_HK_MAN_OB	0.102489
BR_PRO	0.059009	MX_RET	0.040631	PMI_HK_MAN_PM	0.102925
CH_CAR	-0.034171	NW_CAR	0.015779	PMI_IN_MAN_NE	0.103269
CH_CNF	0.046859	NW_EMP	0.031543	PMI_IN_MAN_NI	0.093834
CH_PRO	-0.009232	NW_PRO	-0.002278	PMI_IN_MAN_OB	0.099037
CH_RET	-0.007066	NW_RET	-0.010129	PMI_IN_MAN_PM	0.097411
CN_CAR	0.033900	NW_UNR	-0.023923	PMI_IT_COM_HE	0.116114
CN_CNF	0.025181	PMI_BR_MAN_NE	0.113903	PMI_IT_MAN_NE	0.115774
CN_EMP	0.063748	PMI_BR_MAN_NI	0.104900	PMI_IT_MAN_NI	0.118379
CN_PRO	0.066614	PMI_BR_MAN_OB	0.106695	PMI_IT_MAN_OB	0.116180
CN_UNR	-0.081979	PMI_BR_MAN_PM	0.098649	PMI_IT_MAN_PM	0.112762
DE_CAR	0.002741	PMI_CN_COM_HE	0.076702	PMI_IT_SER_HE	0.104653
DE_CNF	0.043167	PMI_CN_MAN_NE	0.106978	PMI_JP_MAN_NE	0.113448
DE_EMP	0.064232	PMI_CN_MAN_NI	0.087971	PMI_JP_MAN_NI	0.110533
DE_PRO	0.072788	PMI_CN_MAN_OB	0.077804	PMI_JP_MAN_OB	0.110040
DE_RET	0.002158	PMI_CN_MAN_PM	0.086697	PMI_JP_MAN_PM	0.111427
DE_UNR	-0.069217	PMI_CN_SER_HE	0.059093	PMI_KR_MAN_NE	0.085873
ES_CAR	-0.027556	PMI_DE_COM_HE	0.122171	PMI_KR_MAN_NI	0.096783
ES_CNF	0.031974	PMI_DE_MAN_NE	0.122981	PMI_KR_MAN_OB	0.092536
ES_EMP	0.027251	PMI_DE_MAN_NI	0.120081	PMI_KR_MAN_PM	0.091808
ES_PRO	0.066704	PMI_DE_MAN_OB	0.120394	PMI_RU_COM_HE	0.110233
ES_RET	0.003238	PMI_DE_MAN_PM	0.109788	PMI_RU_MAN_NE	0.098116
ES_UNR	-0.070491	PMI_DE_SER_HE	0.108191	PMI_RU_MAN_NI	0.100449
FR_CAR	-0.026526	PMI_ES_COM_HE	0.113163	PMI_RU_MAN_OB	0.105043
FR_CNF	0.022467	PMI_ES_MAN_NE	0.110849	PMI_RU_MAN_PM	0.104171

Table A2.1. Factor loadings (cont.)

FR_PRO	0.064631	PMI_ES_MAN_NI	0.113788	PMI_RU_SER_HE	0.101242
FR_RET	-7.39E-05	PMI_ES_MAN_OB	0.111878	PMI_US_COM_HE	0.094756
FR_UNR	-0.023072	PMI_ES_MAN_PM	0.107639	PMI_US_MAN_NI	0.105165
HK_CAR	0.009031	PMI_ES_SER_HE	0.107865	PMI_US_MAN_PM	0.100806
HK_EMP	0.066321	PMI_FR_COM_HE	0.117448	PMI_US_SER_HE	0.100465
HK_PRO	0.060080	PMI_FR_MAN_NE	0.122720	RS_EMP	0.037889
HK_RET	0.001127	PMI_FR_MAN_NI	0.118609	RS_PRO	0.033910
HK_UNR	-0.088604	PMI_FR_MAN_OB	0.118656	RS_UNR	-0.064550
IN_CAR	-0.032913	PMI_FR_MAN_PM	0.113749	SA_CAR	-0.042012
IT_CAR	-0.039141	PMI_FR_SER_HE	0.110442	SA_CNF	0.017457
IT_CNF	0.018742	PMI_GB_COM_HE	0.121658	SA_PRO	0.043775
IT_PRO	0.061003	PMI_GB_MAN_NE	0.103293	SA_RET	0.026430
IT_RET	0.050962	PMI_GB_MAN_NI	0.110252	UK_CAR	-0.003739
IT_UNR	-0.004454	PMI_GB_MAN_OB	0.112893	UK_CNF	0.024464
JP_CAR	0.020999	PMI_GB_MAN_PM	0.108912	UK_EMP	0.023726
JP_CNF	0.030921	PMI_GB_SER_HE	0.117827	UK_PRO	0.058665
JP_EMP	0.030668	PMI_GB_WHE_HE	0.122020	UK_RET	0.026958
JP_PRO	0.066480	PMI_GL_COM_HE	0.128824	UK_UNR	-0.091848
JP_RET	0.053427	PMI_GL_MAN_NE	0.129295	US_CAR	0.012625
JP_UNR	-0.034537	PMI_GL_MAN_NI	0.126551	US_CNF	0.031681
KO_CAR	-0.003546	PMI_GL_MAN_OB	0.127084	US_EMP	0.040787
KO_EMP	0.026657	PMI_GL_MAN_PM	0.124251	US_PRO	0.078187
KO_PRO	0.032110	PMI_GL_MAN_SD	-0.056321	US_RET	0.044365
KO_RET	-0.010346	PMI_GL_SER_HE	0.122883	US_UNR	-0.066298
KO_UNR	-0.015248				

We further check this issue by comparing our baseline factor with the factor that would have been obtained if PMIs were excluded from our dataset. The two factors obtained with and without PMIs are represented in **Figure A2.1** below. While both show the same sharp turning points around crisis episodes in 2009 and 2020, the factor including the PMIs appears smoother. The larger amount of noise in the non-PMI factor might make it harder to capture shallower movements in global dynamics. This is one of the reasons why our baseline approach in the paper rely on the PMI-based factor.

Figure A2.1. Monthly factors with and without PMIs



We finally run the same regressions as in the core paper but using the non-PMI factor. Results are reported in **Table A2.2** below. Our baseline findings remain valid when excluding PMIs. Models with weekly data and weekly factors (1, 2, and 3) still outperform other models both in-sample and out-of-sample. Accuracy gains for these models are greater during "crisis" episodes while during "normal" periods, models with weekly data and other modes based on monthly factor (4 and 5) or an AR term (6) have comparable performances. A slight difference with our baseline results is that for non-crisis episodes, the AR model (model 6) now perform better than models with monthly factors (4 and 5) which might suggest that PMIs helps in tracking the shallower changes in global dynamics.

Table A2.2. Performances (RMSE) across models, months, and periods

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
All sample						
In-sample						
1st month	**0.451**	0.492	0.501	0.587	0.539	0.689
2nd month	**0.236**	0.339	0.362	0.589	0.545	0.692
3rd month	**0.238**	0.250	0.254	0.374	0.277	0.692
Out-of-sample						
1st month	0.562	**0.542**	0.545	0.608	0.558	0.686

2nd month	**0.436**	0.496	0.508	0.608	0.562	0.589
3rd month	**0.359**	0.375	0.378	0.497	0.394	0.589
Crisis episodes						
In-sample						
1st month	**1.012**	1.150	1.181	1.481	1.342	1.635
2nd month	**0.237**	0.573	0.656	1.473	1.358	1.433
3rd month	**0.246**	0.324	0.336	0.762	0.420	1.433
Out-of-sample						
1st month	1.389	**1.345**	1.353	1.539	1.407	1.709
2nd month	**1.012**	1.211	1.247	1.529	1.422	1.545
3rd month	**0.711**	0.795	0.809	1.201	0.885	1.545
Non-crisis episodes						
In-sample						
1st month	0.282	0.280	0.279	0.256	**0.251**	0.300
2nd month	0.235	0.288	0.293	0.269	0.252	**0.203**
3rd month	0.237	0.237	0.239	0.271	0.248	**0.203**
Out-of-sample						
1st month	0.268	0.255	0.256	0.262	**0.243**	0.318
2nd month	0.254	0.249	0.250	0.268	0.243	**0.197**
3rd month	0.268	0.258	0.256	0.258	0.247	**0.197**

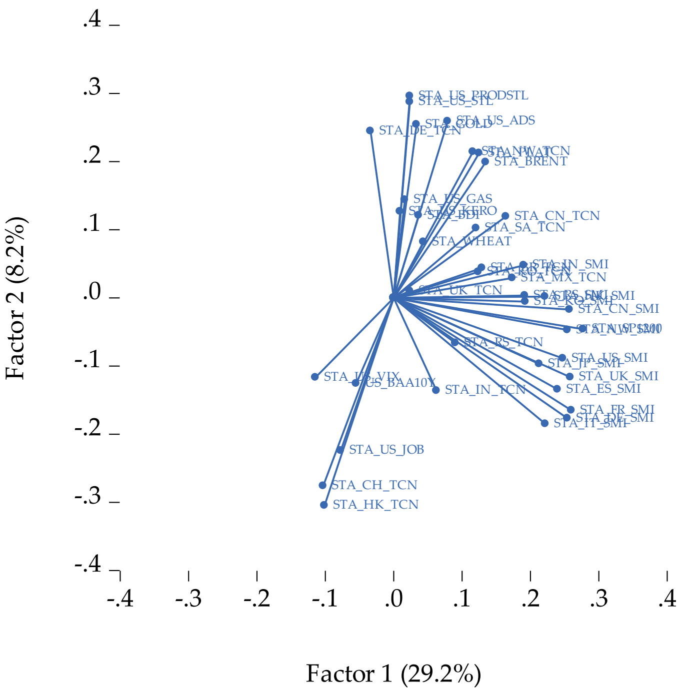
Grey cells indicate best performance for a given month

Annex 3: Alternative approach for the weekly factors

Our baseline approach regarding weekly data is to *ex ante* separate the dataset into a subset of financial series (VIX, stock market indices, NEER) and a subset of variables for the "real" economy (variables for the US economy and global variables). An alternative approach could be to separate *ex post* by taking the first and second factors obtained by running a PCA on the entire dataset. If the dataset includes two orthogonal subsets (financial *vs.* "real"), this approach would in principle clearly disentangle the two. Intuitively, we would expect one of the factor to be mostly related to financial variables while the other would pertain to the "real" economy.

A graph of the orthonormal loadings in **Figure A3.1** seems to corroborate this: variables for the "real" economy are more correlated with the 2nd factor on the y-axis (US steel production, ADS business conditions index, gas consumption, etc.) while most of stock market indices are highly correlated with the 1st factor on the x-axis. Interestingly, some of the NEER variables show greater correlation with the 2nd factor than with the supposedly-financial factor (e.g. Chinese, German, and Indian NEER).

Figure A3.1. Orthonormal loadings



Against this background, we now turn to a new **model 1b** which includes the monthly factor (f_t^m) and the two first weekly factors obtained by running a PCA on the entire weekly dataset. As stated above, the first factor ($f_t^{w,1}$ – on the x-axis) seems to relate more closely to financial variables while the second factor ($f_t^{w,2}$ – on the y-axis) is seemingly more correlated to the "real" economy. Comparing model 1 – in which the weekly factors are obtained by running two different PCAs on the two different "financial" and "real" subsets of weekly data – to model 1b allow us to test which approach yields better performance.

Results are reported in the **Table A3.1** below. They indicate that performances are slightly better both in-sample and out-sample for model 1 where the two subsets of weekly data are separated *ex ante*. This result is consistent with the findings of Cadima and Jolliffe (2001)[13] for which the selection of subsets in a large dataset should not be solely based on PCA. Therefore, in this paper, we distinguish weekly data *ex ante* into two subsets – and run two distinct PCAs – rather than letting the PCA perform itself an *ex post* division of the dataset.

Table A3.1. Performances (RMSE) across models and months of the quarter

	Model 1	Model 1b
In-sample		
1st month	**0.442**	0.442
2nd month	**0.211**	0.261
3rd month	**0.202**	0.227
Out-of-sample		
1st month	0.599	**0.546**
2nd month	**0.422**	0.459
3rd month	**0.351**	0.429

Grey cells indicate best performance for a given month

[13] Cadima J. and Jolliffe I. (2001). "Variable Selection and the Interpretation of Principal Subspaces," *Journal of Agricultural, Biological, and Environmental Statistics*, 6(1), pp. 62-79

Annex 4: Disentangling emerging and advanced economies

Given our large dataset covering both emerging and advanced economies, it is possible to perform two distinct nowcasting models for the two sub-groups. To that end, we distinguish our monthly dataset between emerging and advanced variables (resp. 57 and 100 series[14]). For weekly data, we separate only the financial subset between emerging and advanced variables. Weekly variables for the "real" economy – despite being mostly related to the US cycle – are included for both emerging and advanced economies given the pivotal role of the US economy in shaping the global cycle (see Kindberg-Hanlon and Sokol, 2018) and the ability of the Baltic Dry Index to track economic conditions across emerging economies (Chiu *et al.*, 2020). In the end, the weekly dataset for advanced economies includes 27 series and 26 for the dataset on emerging economies.

A first question relates to whether our results still hold when disentangling emerging and advanced economies. Model performances by country groups are reported in **Table A4.1**. Our baseline results are still robust to this major change in dataset: models with weekly data (1, 2, and 3) still outperform other models both in- and out-of-sample. Interestingly, these results indicate that nowcasting models are more accurate for advanced economies – which might not be surprising considering that weekly data are scarce for emerging economies and that even at monthly frequency, our dataset includes almost twice as many series for advanced economies than for emerging countries.

Table A4.1. Performances (RMSE) across models, months, and country groups

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Advanced economies						
In-sample						
1st month	**0.321**	0.330	0.339	0.391	0.401	0.618
2nd month	**0.216**	0.231	0.234	0.355	0.352	0.612
3rd month	**0.202**	0.215	0.214	0.265	0.277	0.612
Out-of-sample						
1st month	0.455	**0.443**	0.451	0.466	0.466	0.672
2nd month	0.360	0.360	**0.351**	0.393	0.387	0.557
3rd month	0.349	**0.338**	0.344	0.343	0.361	0.557
Emerging economies						
In-sample						

[14] This exceeds 151 since our dataset also includes 6 "global" PMIs variables that are included in both subsets.

1st month	**0.570**	0.616	0.640	0.690	0.688	0.814
2nd month	**0.293**	0.354	0.403	0.561	0.576	0.826
3rd month	**0.288**	0.349	0.383	0.450	0.538	0.826
Out-of-sample						
1st month	0.747	**0.710**	0.723	0.710	0.713	0.766
2nd month	**0.562**	0.597	0.630	0.649	0.652	0.700
3rd month	**0.549**	0.594	0.630	0.608	0.630	0.700

Grey cells indicate best performance for a given month

Given the accuracy of these sub-models, a second question is whether a "bottom-up" approach – in which GDP is first forecasted at country-group level and then aggregated at global level – would yield better results than our baseline "top-down" approach – in which world GDP is directly forecasted. The **Table A4.2** below summarizes the relative in-sample RMSE for the "bottom-up" approach compared to the "top-down" approach. In most cases, the "top-down" approach outperforms the "bottom-up" approach; however performances are mostly comparable across all models and months of the quarter.

Table A4.2. Relative performances (RMSE) of a "bottom-up" approach compared to our baseline "top-down" approach

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
1st month	0.995	0.979	1.008	1.015	1.015	1.000
2nd month	1.015	1.032	1.090	0.948	0.961	0.988
3rd month	1.021	1.142	1.218	1.189	1.316	0.988

A coefficient below (resp. above) 1 indicates a better (resp. worse) performance relatively to the "top-down" approach

These results support our baseline "top-down" approach of nowcasting directly world GDP – in line with a strand literature supporting "top-down" forecasting such as Burgert and Dées (2008)[15]. However, accuracy gains are limited and nowcasting performances of sub-models for emerging and advanced economies appear decent. Therefore there might be value, from a practitioner perspective, to rely on the "bottom-up" approach when it make sense to disentangle growth dynamics across different country groups.

[15] Burgert M. and Dées S. (2008). "Forecasting world trade: direct versus "bottom-up" approaches", *European Central Bank Working Paper Series*, No 882

Annex 5: Alternative variable selection

In our baseline approach, variable selection for monthly data rely on four criteria: limited publication lags, sufficient timespan, cross-country availability, and correlation with our target variable. Imposing cross-country availability (i.e. if we select a variable, e.g. retail sales, we include these variables across all countries and not only for a few of them) remains an arbitrary choice. This is justified by the fact that we are not only interested in the nowcasting results but also in how the factor evolves over time. Imposing cross-country availability allows us to get a dataset more consistent and balanced in terms of countries and variables covered, therefore allowing us to decompose more clearly the underlying forces behind the factor. In other words, in the interpretability *vs.* complexity trade-off (Alonso *et al.*, 2009[16]), we've favoured the first.

We test an alternative approach in which we relax this cross-country availability constraint. The selection among 500 potential regressors is based on the correlation of a unique regressor with the target variable following the "sure independence screening" approach proposed by Fan and Lv (2008)[17]. Series are ranked by their degree of correlation with the target variable[18]. Then, we include them progressively by group of 5 variables, and compute the monthly factor and the out-of-sample RMSE over 2005-2020. The number of variables grows by 5 at each step: first we run the FA-MIDAS with the top 5 variables, then with the top 10, and so on.

Results appear to validate our baseline approach to a certain extent. First, selected variables are close: out of the top 100 individual series more correlated with the target variable, 92 are selected in the baseline approach. Second, performances are at their peak when including around 150 series. Results show a U-shaped relationship between the number of series included in the factor and performances: adding series yields better performance until a certain point after which performance deteriorates (see **Figure A5.1**). This point falls at 165 series, very close to the number of series included under the baseline approach (151). In addition, the factor obtained through this optimal selection is very close to the one obtained in our baseline selection (see **Figure A5.2**). It should be noted however that the U-shaped relation appears more clearly for the first month of the quarter. This is in line with the intuition: in the beginning of a quarter, there are little chances that only a few series would bring enough information alone so the factor model has to aggregate a large number of sources. This is less and less the case as we advance in the quarter and the optimal number of series to select – in terms of minimizing the out-of-sample RMSE – falls below 150 for the 2nd and 3rd month. We however stick with our selection around 150 variables since the purpose is

[16] Alonso J., Magdalena L., and González-Rodríguez G. (2009). "Looking for a good fuzzy system interpretability index: An experimental approach", *International Journal of Approximate Reasoning*, 51(1), pp. 115-134

[17] Fan J. and Lv J. (2008). "Sure independence screening for ultrahigh dimensional feature space", *Journal of the Royal Statistical Society Series B*, vol. 70(5), pp. 849-911

[18] To avoid potential bias, we exclude 2020 and compute the correlation only over 1998-2019. For each variables, we compute correlations for the 1st and 3rd months of the quarter by taking respectively the value of variables in the 1st month and the average of its values over the full quarter for the 3rd month.

rather to enhance performances in the very first month of the quarter – where other estimates are limited.

Figure A5.1. Relationship between the number of series in the factor and performances (out-of-sample RMSFE – relative to the best performance = 1)

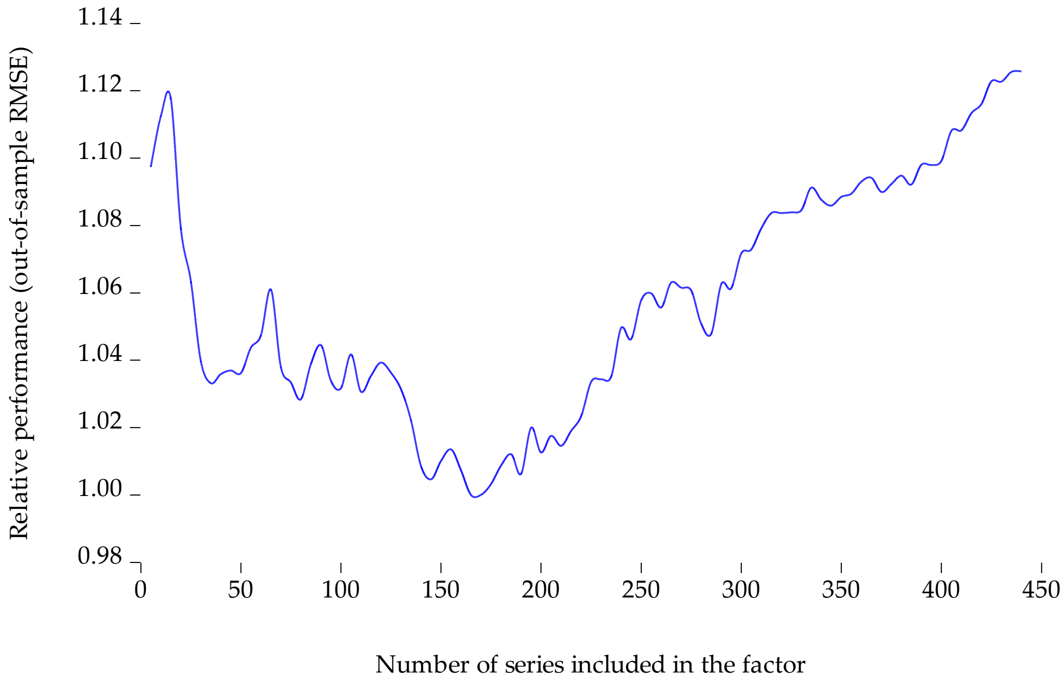


Figure A5.2. Monthly factors: baseline *vs.* alternative approach