Degiannakis
*Swiss Journal of Economics and Statistics* (2023) 159:7
https://doi.org/10.1186/s41937-023-00109-8

Swiss Journal of
Economics and Statistics

ORIGINAL ARTICLE

Open Access

# The D-model for GDP nowcasting

Stavros Degiannakis[1,2*]

**Abstract**

The paper provides a disaggregated mixed-frequency framework for the estimation of GDP. The GDP is disaggregated into components that can be forecasted based on information available at higher sampling frequency, i.e., monthly, weekly, or daily. The model framework is applied for Greek GDP nowcasting. The results provide evidence that the more accurate nowcasting estimations require (i) the disaggregation of GDP, (ii) the use of a multilayer mixed-frequency framework, and (iii) the inclusion of financial information on a daily frequency. The simulation study provides evidence in favor of the disaggregation into components despite the inclusion of multiple sources of forecast errors.

**Keywords** Nowcasting, Forecasting, GDP, Disaggregation, Factors, Multilayer, Mixed frequency

**JEL Classification** E27, C53

## 1 Introduction

We investigate whether the use of a disaggregated multilayer mixed-frequency framework improves the Greek GDP nowcasting performance. In the disaggregated approach, we nowcast each GDP component separately and aggregate them to obtain a GDP nowcast. This is the first paper to introduce the idea to combine a model for mixed-frequency data with a multilayer strategy.

In the first layer, we estimate a MIDAS regression for each GDP component, in which the dependent variable (e.g., growth in private consumption) is observed on a quarterly basis while the explanatory variables are observed on a monthly frequency. However, some of the explanatory monthly variables are published with a lag of several months, resulting in the unavailability of their most recent values. So, in the second layer of the model, we estimate the unavailable values of the monthly variables based on the information that is available at a higher frequency (i.e., daily frequency). For example, asset prices (e.g., stock prices) which are observed on a daily basis could provide information that is not incorporated in a monthly economic index (i.e., consumer confidence). Overall, we apply the proposed novel framework on a variety of economic and financial data (hard and soft data, mostly domestic) to nowcast Greek GDP and evaluate its nowcasting performance over the period 2005Q1–2020Q3.

Moreover, we provide empirical and simulated evidence that more accurate nowcasting estimations require the use of a disaggregated multilayer mixed-frequency framework. First, we show that the nowcasting ability of the AR(1) used as naive model is not better, if and only if a sophisticated model framework is defined. Second, the disaggregation into components reduces the nowcasting error despite the inclusion of multiple sources of nowcasting errors.

The proposed model framework, named D-model,[1] is very relevant for practitioners and policymakers who need to get informed accurately and in real time of the current state of the economy under investigation.

The rest of the paper is structured in a wise manner: Sect. 2 provides a literature review on GDP nowcasting,

*Correspondence:
Stavros Degiannakis
stavros.degiannakis@gmail.com
[1] Economic Research Department, Bank of Greece, 21 E. Venizelos Avenue, 10250 Athens, Greece
[2] Department of Economics and Regional Development, Panteion University of Social and Political Sciences, 136 Syggrou Av., 17671 Athens, Greece

[1] The acronym D-model stands for the Disaggregated model.

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Sect. 3 presents the D-model's construction, and Sect. 4 describes the dataset. Section 5 presents in detail the model specifications for the Greek GDP nowcasting. In Sect. 6, we proceed to a number of additional model extensions for robustness purposes, and in Sect. 7 we estimate nowcasts from naïve models in order to have a reference point in the evaluation of the nowcasting performance, which is illustrated in Sect. 8. Section 9 presents Monte Carlo simulations which provide evidence in favor of the disaggregation into components and, finally, Sect. 10 presents the conclusions.

## 2 Literature review

Dynamic factor models (DFMs) and bridge models (BMs) are the most pop2004ular tools in short-term forecasting on real activity variables, such as the GDP growth. Bridge equations for forecasting GDP have been studied by Baffigi et al. (2004) and Diron (2008), among others. Barhoumi et al. (2008) study factor models for ten European countries and the euro area as a whole, concluding in their interior performance compared to averages of traditional bridge equations. Factor models for forecasting GDP, also have been applied by Marcellino et al. (2003) for euro-area data, Artis et al. (2005) for the UK, Den Reijer (2005) for the Netherlands, Duarte and Rua (2007) for Portugal, Schumacher (2007) for Germany, and Van Nieuwenhuyze (2005) for Belgium.

Both types of models come with their advantages and flaws. BMs are characterized by two empirical limitations. Firstly, the monthly series must be sufficiently long to guarantee the precision of the estimates. Secondly, it is not possible to include a large number of variables, because of the risk of multicollinearity and losses of degrees of freedom. On the other hand, DFMs are presented as a less restrictive alternative tool compared to BMs, especially for short-term forecasting of GDP growth (see, inter alia, Angelini et al., 2008; Bańbura and Rünstler, 2011). A wider set of collinear monthly indicators is parsimoniously summarized with only a few common factors, making the projection possible and the number of parameters limited.

In Appendix A, we provide an overview of selected papers dealing with short-term GDP forecasting techniques. A set of interesting conclusions can be derived:

i. There is a controversy about which of the two competitive frameworks (DFMs vs. BMs) has the best forecasting accuracy. There is a forecasting debate between DFMs and BMs and the evaluation of several forecasting error measures does not provide a clear view in favor of DFMs.

ii. According to empirical findings, the performance of the DFMs compared to a set of benchmarking models (random walk and autoregressive models) is clearly better, as their evaluation with several forecasting error measures, provides evidence in favor of their superiority. The same holds true for Stakénas (2012) who provides evidence of factor models' performance superiority compared to naïve benchmark models.

Interesting conclusions driven from studies that worth mentioning and focus mainly on the euro area countries, in favor of DFM, can be summed up as follows: Angelini et al. (2008) who estimate a DFM for the euro-area economy and find that for GDP and a number of components, factor model forecasts beat the forecasts from alternative models such as quarterly models and bridge equations. Again, in a follow up paper Angelini et al. (2011), using euro-area data, provide evidence that factor model improves upon the pool of bridge equations. Also, Barhoumi et al. (2008) maintain that for the euro-area countries, factor models which exploit a large number of releases, do generally better than averages of bridge equations. Likewise, according to Bańbura and Rünstler (2011) once more for the euro-area economy, highlight the importance of survey data on both forecast weights and forecast precision measures, the moment that real activity data obtain rather low weights, apart perhaps from the backcasts. Financial data provide complementary information to both real activity and survey data for nowcasts and one-quarter-ahead forecasts of GDP.

However, apart from those two popular tools, another framework has made a dynamic appearance, the mixed sampling frequency modeling framework that has recently been incorporated into the GDP nowcasting literature. Ghysels et al. (2006) and Andreou et al., (2010, 2013) propose Midas (mixed-data sampling) model when one desires to relate a dependent variable (i.e., the quarterly GDP) with explanatory variables sampled in higher sampling frequency (i.e., monthly or weekly data).

Chernis and Sekkel (2017) estimate DFM, BM as well as Midas models for nowcasting the Canadian gross domestic product. They compare the average of the Midas predictions against the forecast of the DFM (by underweighting the poor performing variables) and conclude that the DFM outperforms its competitors. Clements and Galvao (2009), on the other hand, forecast US growth with Midas models and provide important findings regarding the outperformance of Midas framework in exploiting information from the leading indicators. Marcellino and

Schumacher (2010) introduce a Factor-Midas approach to nowcast and forecast quarterly German GDP growth. They find that the most parsimonious Midas projection is the best performing overall. Kuzin et al. (2011) compare the forecasting ability of Midas and MFVAR in forecasting euro-area quarterly GDP and find that Midas tends to perform better for shorter horizons, and MFVAR for longer horizons. Jansen et al. (2016) evaluate the predictive ability of almost all the available statistical models (i.e., VAR, Bayesian VAR, mixed-frequency VAR (MFVAR), DFM, BM, Midas) in predicting GDP for the euro area, Germany, France, Italy, Spain, and the Netherlands. They conclude that the dynamic factor model is the best model overall due to its ability to incorporate more information.

Furthermore, Kim and Swanson (2018) apply Factor-Midas models for nowcasting and forecasting the Korean GDP. In their forecasting exercise, models with one or two factors are the best for all forecasting horizons, whereas in backcasting and nowcasting horizons, models with more factors are preferred. They also notice that as forecast horizon gets shorter (i.e., move from forecast → nowcast → backcast), the AR and RW models perform better. Also, Andreou et al. (2013) use Factor-Midas to examine the usefulness of daily financial data to forecast macroeconomic series. Foroni and Marcellino (2014) nowcast the quarterly growth rate of the euro-area GDP and conclude that the Midas model outperforms MFVAR at most forecasting horizons. Additionally, they investigate the potential usefulness of disaggregating the information contained in the components of GDP for nowcasting total GDP growth. In their concluding section they state "…findings for the aggregated nowcasts are promising, meaning that there is scope for forecasting the single components to shed light on the total GDP measure."

Finally, a comparison of Midas and bridge equation models for the euro-area GDP growth is provided by Schumacher (2016). Schumacher estimates Midas models with different specifications for the lag polynomials: exponential Almon, multiplicative, unrestricted, etc. Results favor the most parsimonious specifications, with only a few AR and indicator lags. Midas tends to outperform bridge equations noticing, however, that results depend on the particular dataset and the sample chosen.

## 3 Model description

### 3.1 GDP disaggregation into components

The proposed model aims to estimate GDP in constant prices according to the fixed-based approach, defined as $\left(Y_q^{(0)}\right)$, by nowcasting the components of GDP from the expenditure side. Hence, we define one model for each one of the components: private consumption of goods and services, $\left(Y_q^{(1)}\right)$, government spending on public goods and services, $\left(Y_q^{(2)}\right)$, investment in business capital goods, $\left(Y_q^{(3)}\right)$, exports of goods $\left(Y_q^{(4)}\right)$, exports of services $\left(Y_q^{(5)}\right)$, imports of goods, $\left(Y_q^{(6)}\right)$, imports of services, $\left(Y_q^{(7)}\right)$ and changes in inventories, $\left(Y_q^{(8)}\right)$. Naturally:

$$Y_q^{(0)} = \left(\sum_{k=1}^{5} Y_q^{(k)} - \sum_{k=6}^{7} Y_q^{(k)} + Y_q^{(8)}\right). \tag{1}$$

### 3.2 Mixed sampling frequency framework

Let us denote as $y_q^{(k)} = log\left(Y_q^{(k)}/Y_{q-1}^{(k)}\right)$ the *q-o-q* growth rate. We construct the Midas regression in order to extract the information that is available at higher sampling frequencies. The dependent variable is observed on a quarterly basis, but explanatory variables are available at a higher frequency, i.e., on a monthly basis. However, there is information that is available at an even higher sampling frequency such as the asset prices from financial markets. So, what we suggest is the construction of a multilayer model framework, in which the different sampling frequencies are defined as layers. The rationality behind the multilayer model framework is to estimate the missing information at a lower frequency based on the information that is available at a higher frequency. For example, the price of the stock index which is observed on a daily basis could provide information that is not incorporated in the index of economic climate, which is observed monthly. In turn, the index of economic climate may provide information for the GDP component investment on business capital goods which is observed quarterly. Hence, the proposed model framework is estimated in the form:

$$y_q^{(k)} = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \tag{2}$$

where the error term is defined as $\varepsilon_q \sim N\left(0, \sigma^2_{\varepsilon_q}\right)$.[2] The $\boldsymbol{X}_{(m)}$ denotes the vector of variables observed at a monthly frequency. The $\beta_0$ is a coefficient, $\boldsymbol{\theta}_j$ is a vector of coefficients to be estimated, $p$ is the Almon polynomial order, $\kappa$ is the number of lagged months to employ, and $s = 3$ denotes the number of months of each quarter. The $i$ term determines the time that the information set is available as well as the capacity of the model to estimate predictions without imposing a look ahead bias. For example, if we set $i = 0$, we are able to nowcast the GDP component, i.e., $y^{(k)}_{q\backslash q}$. If we set $i \geq 1$ and $is \geq 3$, then we are able to estimate one-quarter ahead $y^{(k)}_{q+1\backslash q}$, etc.

### 3.3 Multilayer framework

Using the same rational, we create the next layer that represents the estimation of the non-available values of the variables at a monthly frequency. Let us assume that in the 1st layer, where the dependent variable is at a quarterly frequency, the explanatory monthly variable is an index of retail sales which is published with a lag of 2 months. The most recent value of the monthly variable is not available, so it must be estimated based on information that is available. Hence, for the values of the monthly variables that are not available, we have to define the 2nd layer of our model:

$$x_m = \gamma_0 + \boldsymbol{\gamma} \tilde{\boldsymbol{X}}'_m + \sum_{r=0}^{l-1} \mathbf{Z}'_{(d-r-is)} \left( \sum_{j=0}^{q} r^j \boldsymbol{\varphi}_j \right) + \varepsilon_m, \tag{3}$$

where $\varepsilon_m \sim N\left(0, \sigma^2_{\varepsilon_m}\right)$. The $\gamma_0$ is a coefficient, $\boldsymbol{\gamma}$ and $\boldsymbol{\varphi}_j$ are vectors of coefficients to be estimated, $q$ is the polynomial order, $l$ denotes the number of lagged days and $s = 22$ denotes the number of trading days of each quarter. The $\widetilde{\boldsymbol{X}}_m$ denotes the vector of variables observed at monthly frequencies and provide explanatory power for the $x_m$. The $\boldsymbol{Z}_{(d)}$ denotes the vector of variables observed on a daily frequency. Regarding the $i$ term, for $i = 0$, we estimate $x_{m\backslash m}$, while for $i \geq 1$ and $is \geq 22$, we estimate the one-month ahead $x_{m+1\backslash m}$, and for $i \geq 2$ and $is \geq 44$, we estimate the two months ahead $x_{m+2\backslash m}$, and so on.

### 3.4 Nowcasting error correction

We assume that the GDP nowcasts contain a forecast error with an autocorrelated structure. Possible sources of the autocorrelated structure of the forecast error could be (i) the multiple revisions of the figures, (ii) the construction of GDP as a summation of its components which are also revised frequently, (iii) the inclusion of multiple sources of forecast errors; as the computation of the nowcast requires the summation of multiple nowcast values. Thus, we propose a short-term forecast error structure, around the long-term structure, $Y^{(0)}_{q-1} = \mu + \beta_1 Y^{(0)}_{q-1\backslash q}$:

$$\begin{aligned} \Delta Y^{(0)}_q = & \sum_{j=0}^{K} \delta_j \Delta Y^{(0)}_{q-j\backslash q} + \sum_{j=1}^{J} \lambda_j \Delta Y^{(0)}_{q-j} \\ & + \alpha \left( Y^{(0)}_{q-1} - \beta_1 Y^{(0)}_{q-1\backslash q} - \mu \right) + u_q, \end{aligned} \tag{4}$$

for $u_q \sim N\left(0, \sigma^2_u\right)$, where $Y^{(0)}_q$ denotes the published GDP for quarter $q$ and $Y^{(0)}_{q-j\backslash q}$ is the nowcasted value of GDP of quarter $q$-$j$ based on the information that is available up to the most recent quarter, i.e., $q$.

## 4 Data description

The handling of exogenous variables in nowcasting models should be done very carefully. The usual practice of creating a sandbox, which encloses all the variables that we have managed to collect, is not the most appropriate. Schumacher (2010) and Boivin and Ng (2006) note that only a careful preselection of predictors helps in exploiting the additional information from large and heterogeneous data. Thus, more data are not always better for nowcasting or for forecasting. Moreover, Boivin and Ng (2006) show that the sample size of the dataset has only a minor effect on the estimation. In our case, the dataset has been constructed based on the economic intuition, the current state of the literature and the availability of data in a continuous format for the adequate time frame. As first noted by Stock and Watson (2002a), the appropriate transformations of the data must be applied, so natural logarithms were taken for the majority of the variables (except, i.e., for interest rates) and stationarity was obtained by appropriately differentiating time series. When there were evident any scale effects, the variables were standardized to have a zero mean and unit sample variance. The vast majority of our variables, that were

[2] For the estimation of the models, we define a specific conditional mean, but we do not need to specify the distribution of the error term, as long as the independency of residuals over time holds. The assumption of normally distributed errors is required for the computation of the maximum likelihood.

available in levels, were standardized and de-seasonalized, i.e., for $x_{i,t}$ denoting the $i^{th}$ variable for month $t$; the de-seasonalized and standardized variables are: $\tilde{x}_{i,t} = \left(x^*_{i,t} - \overline{x}^*_i\right)/\sqrt{V\left(x^*_i\right)}$, where $\overline{x}^*_i$ and $V\left(x^*_i\right)$ are the mean and variance estimates of the de-seasonalized $i$th variable, $x^*_{i,t}$, respectively.

The sample runs from January 2002 up to December 2020. Regarding the quarterly frequency, the data are available up to the 3rd quarter of 2020. The sample size is dictated by the availability of data. The out-of-sample evaluation period runs from 2005Q1 up to 2020Q3 and, despite its short length, includes both normal times and crisis period, i.e., Greek sovereign crisis. We use the recursive estimation scheme due to the small sample size, as the alternative approach of the rolling scheme, requires the use of a fixed window in order to re-estimate the parameters.

Also, we highlight the ragged-edge data problem. Let us consider that the consumer price index of previous month is released early in the current month, whereas the producer price index is released in the middle of the month. In between these releases, new vintages of GDP may be released. This is called the ragged-edge data problem. Kim and Swanson (2018), among others, have suggested the vertical alignment and the autoregressive interpolation for the missing values. In our proposed model framework, any variable is considered as observed after being published. But in the case that a variable is not available at the time when we want to proceed to nowcasting, then the method of estimating any missing values is defined explicitly.[3]

Regarding the Greek economy, the quarterly datasets are the GDP and its components, private consumption on goods and services, government spending on public goods and services, investment on business capital goods, exports of goods, exports of services, imports of goods, imports of services. Table 1 presents the data that we have used as explanatory variables. In the D-model we did evaluate almost the entire available dataset, but the variables that were finally incorporated, at a monthly frequency, are, the HICP, loans to private sector, loans to firms, financial conditions Index, the economic sentiment indicator, the purchasing managers' index, the interest rate on new loans, capital goods other than transport equipment—CAPG1, capital goods parts and accessories—CAPG2, the retail trade volume index, the retail trade turnover index, services, the confidence indicator, the consumer confidence indicator, the retail confidence indicator, the employment expectation index, the total volume of retail sales, the volume of retail sales excluding fuel, new private passenger car registrations, price expectations over next 3 months, value-added tax, deposits of households (in flows), and credit to households (in flows). Also, from the balance of payments, we collect the importation of goods, importation of fuels, importation of vessels, importation of other services, travel receipts, transportation receipts, and other receipts. On a daily frequency, the incorporated variables are: the Athens stock exchange main general index, and the 10-year Greek government bond yield.

## 5 Model specifications for the Greek GDP

In Sect. 3, we described the proposed disaggregated mixed-frequency framework for the estimation of GDP. In this section, we will present in detail the model framework for the components of GDP.

### 5.1 Private consumption on goods and services

Private consumption of goods and services, $Y_q^{(1)}$, is highly related to the retail trade volume index, $x_m^{(1)}$, and the retail turnover volume index, $x_m^{(2)}$. However, these indices are published by the Hellenic statistical authority with a publication lag of three months. Specifically, the indices for any month $m$ (i.e., June, 2020) are published the last day of month $m+2$ (i.e., 31st of August, 2020); therefore, we are forced to consider a publication lag of three months. Additionally, a wide information set has been constructed that includes variables such as the services confidence indicator, $x_m^{(3)}$, the consumer confidence indicator, $x_m^{(4)}$, the retail confidence indicator, $x_m^{(5)}$, the economic sentiment indicator, $x_m^{(6)}$, the employment expectation index, $x_m^{(7)}$, the total volume of retail sales, $x_m^{(8)}$, the volume of retail sales excluding fuel, $x_m^{(9)}$, new private passenger car registrations, $x_m^{(10)}$, price expectations over the next 3 months, $x_m^{(11)}$, HICP, $x_m^{(12)}$, value-added taxation, $x_m^{(13)}$, deposits of households (flows), $x_m^{(14)}$, credit to households (flows), $x_m^{(15)}$, etc.[4] The publication of the information over time is visualized in Table 2.

But, a preliminary analysis, as presented in Fig. 1, which presents the scatterplot of the aforementioned variables, enhances our belief that $x_m^{(1)}$ and $x_m^{(2)}$ can provide

[3] We do not preselect a specific method, i.e., AR(1), for estimating the missing values. It is case wise.

[4] The variables are seasonally adjusted with the X12 method. For variables with non-positive values, the transformation $x_t^* = x_t - min\left(x_t\right\}_{t=0}^{T}\right) + 1$ is used.

**Table 1** Explanatory variables under investigation

| Economic outlook | Source | Freq | Lag | Financial variables | Source | Freq | Lag |
|---|---|---|---|---|---|---|---|
| Turnover Index in Retail Trade (Overall) | HSA | M | 3 | Oil (Brent) USD | Bloomberg | D | 0 |
| Turnover Index in Retail Trade (except automotive fuel) | HSA | M | 3 | EUR/USD exchange rate | Bloomberg | D | 0 |
| Retail Trade Volume Index (Overall) | HSA | M | 3 | Oil (Brent) Euro | Bloomberg | D | 0 |
| Retail Trade Volume Index (except automotive fuel) | HSA | M | 3 | Athens General Stock Exchange Index | Bloomberg | D | 0 |
| Consumer Credit | BoG | M | 2 | FTSE Athens Stock Exchange Index | Bloomberg | D | 0 |
| New Private Passenger Car Registrations (ELSTAT) | HSA | M | 2 | EURO STOXX 50—SX5E | Bloomberg | D | 0 |
| Private Building Activity (Construction permits) | HSA | M | 3 | NEXT 150—N150 | Bloomberg | D | 0 |
| Construction Activity (Cement production) | HSA | M | 3 | SPEURO | Bloomberg | D | 0 |
| Housing Loans | BoG | M | 2 | Germany Government Bond 10Y—GDBR10 | Bloomberg | D | 0 |
| Credit to the Private Sector (Corporations and Households) | BoG | M | 2 | Greek Government Bond—10YGGGB10 | Bloomberg | D | 0 |
| Credit (Corporations) | BoG | M | 2 | 10 year bond spreads | Bloomberg | D | 0 |
| Credit (Individuals and private non-profit institutions) | BoG | M | 2 | M1 | BoG | M | 3 |
| Deposits by households and enterprises | BoG | M | 2 | M2 | BoG | M | 3 |
| Exports of Goods | BoP-BoG | M | 2 | M3 | BoG | M | 3 |
| Imports of Goods | BoP-BoG | M | 2 | **Prices** | | | |
| Exports of Services | BoP-BoG | M | 2 | PPI Total (Industrial Produrcer Prices) | HSA | M | 1 |
| Imports of Services | BoP-BoG | M | 2 | PPI Total (Industry excluding Energy on the Domestic Market) | HSA | M | 1 |
| Tourism receipts | BoP-BoG | M | 2 | PPI (Intermediate Goods) | HSA | M | 1 |
| Transport receipts | BoP-BoG | M | 2 | PPI (Capital Goods) | HSA | M | 1 |
| Tourism payments | BoP-BoG | M | 2 | PPI (Durable Consumer Goods) | HSA | M | 1 |
| Transport payments | BoP-BoG | M | 2 | PPI(Non-durable Consumer Goods) | HSA | M | 1 |
| **Sentiment Indices** | | | | PPI Total (Consumer Goods) | HSA | M | 1 |
| Industrial Production Index in Manufacturing | HSA | M | 2 | PPI (Energy) | HSA | M | 1 |
| Turnover Index in Industry | HSA | M | 2 | PPI (Energy/Fuel) | HSA | M | 1 |
| New Orders in Industry (Total) | HSA | M | 3 | PPI (Industrial Producer Prices Non-Domestic Market) | HSA | M | 1 |
| New Orders in Industry (Non-Domestic Market) | HSA | M | 3 | PPI (Import prices in Industry) | HSA | M | 1 |
| Unemployment rate | HSA | M | 1 | CPI | HSA | M | 1 |
| Business Expectations in Construction | FEIR | M | 1 | HICP | HSA | M | 1 |
| Business Expectations in Manufacturing | FEIR | M | 1 | HICP Euro Area | HSA | M | 1 |
| Composite Leading Indicator (CLI) Amplitude adjusted | OECD | M | 2 | HICP (Unprocessed Food) | HSA | M | 1 |
| Industry / Business Climate Indicator (BCI) | OECD | M | 2 | HICP (Processed Food) | HSA | M | 1 |
| Economic Sentiment Indicator (Greece) | FEIR | M | 1 | HICP (Industrial Goods excl Energy) | HSA | M | 1 |
| Industrial Confidence Indicator | FEIR | M | 1 | HICP (Energy) | HSA | M | 1 |
| Services Confidence Indicator | FEIR | M | 1 | HICP (Services) | HSA | M | 1 |
| Consumer Confidence Indicator | FEIR | M | 1 | HICP (excl. energy) | HSA | M | 1 |
| Retail Trade Confidence Indicator | FEIR | M | 1 | HICP Core (Goods excl energy and unprocessed food) | HSA | M | 1 |
| Construction Confidence Indicator | FEIR | M | 1 | | | | |
| Economic Sentiment Indicator (Euro Area) | European Commission | M | 1 | | | | |
| PMI | Markit Economics | M | 1 | | | | |

*ELSTAT* Hellenic Statistical Authority, *FEIR* Foundation for Economic and Industrial Research, *BoP-BoG* balance of payments—Bank of Greece, *OECD* Organization for Economic Cooperation and Development, *Lag* publication lag

**Table 2** Publication of information across time for private consumption on goods and services

| | Previous quarter | Quarter for nowcasting | | | Next quarter |
|---|---|---|---|---|---|
| | 3rd month ($m-3$) | 1st month ($m-2$) | 2nd month ($m-1$) | 3rd month ($m$) | |
| Retail trade volume index | $x_{m-3}$ | | | | |
| Confidence and sentiment indicators | | | $CSI_{m-1}^{(exp)}$ | | |

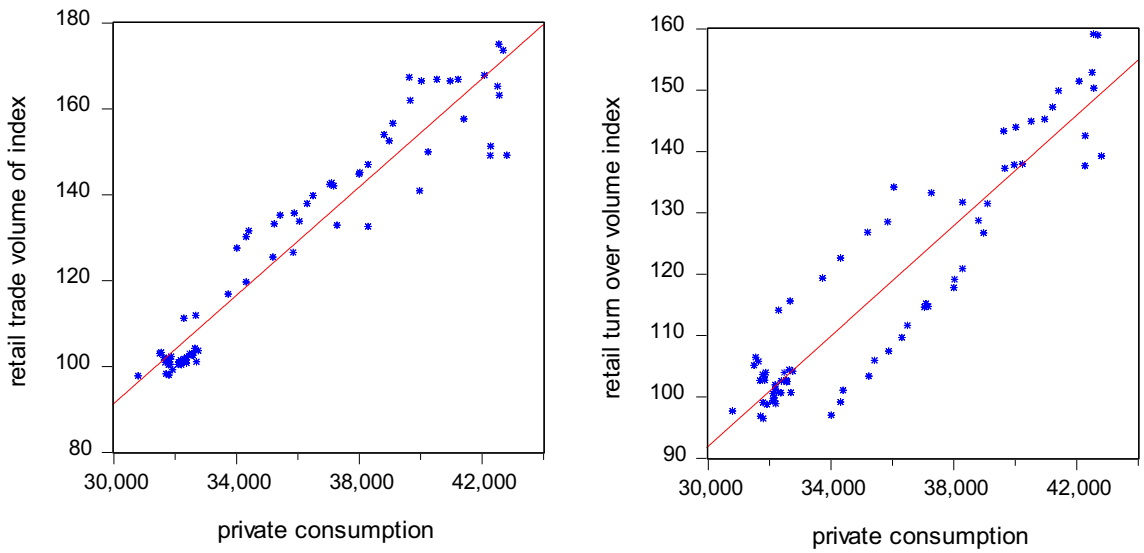



**Fig. 1** Relationship between private consumption on goods and services and **a** retail trade volume of index, **b** retail turnover volume index. The variables are seasonally adjusted. Period: 2002–2020. Source: Hellenic statistical authority

**Table 3** Publication of retail trade volume index and economic sentiment indicator across time

| Q | M | Time to nowcast | X1 published | X3 published | Time to nowcast | X1 published | X3 published |
|---|---|---|---|---|---|---|---|
| Previous quarter | M1 | | V | V | | V | V |
| | M2 | | V | V | | V | V |
| | M3 | | V | V | | | V |
| Current quarter | M1 | | V | V | | | V |
| | M2 | | | V | V | | |
| | M3 | | | V | | | |
| Next quarter | M1 | V | | | | | |
| | M2 | | | | | | |
| | M3 | | | | | | |

accurate information for the nowcast values of $Y_q^{(1)}$. Finally, due to the publication lag of three months, we define a 2nd layer according to Eq. 3 where the non-published values of $x_m^{(1)}$ and $x_m^{(2)}$ variables are estimated based on the vector $\tilde{X}_m' = \left[x_m^{(3)} \ldots x_m^{(7)}\right]'$, published with a lag of one month.

Hence, the nowcasting of private consumption on goods and services, $y_q^{(1)}$, is based on the explanatory power of the retail trade volume index and retail turnover volume index that are published monthly, i.e., $\boldsymbol{X}_{(m)} = \left[x_m^{(1)} x_m^{(2)}\right]'$, and at the same time, the nowcasting of non-published values of $\boldsymbol{X}_{(m)}$ is based on the information available from the confidence and sentiment indicators, i.e., $\tilde{\boldsymbol{X}}_m' = \left[x_m^{(3)} \ldots x_m^{(7)}\right]$.

Table 3 helps us visualize the complexity of nowcasting. Let us assume that we are interested in estimating private consumption for the current quarter. We also assume that the present time (the time when we proceed to the estimations) is the first month of the next quarter. The $x_m^{(1)}$ variable is published up to the 1st month of current quarter, whereas the $x_m^{(3)}$ variable is published up to 3rd month of current quarter. We need to relate private consumption with the $x_m^{(1)}$. The $x_m^{(1)}$ is observed for the 1st month of current quarter, so we do not need any estimation. Regarding the 2nd month, we can estimate the model $x_m^{(1)} = f\left(x_m^{(3)}\right)$ based on the data up to the 1st month of current quarter, and then predict the $x_m^{(1)}$ for the 2nd quarter since we know the value of $x_m^{(3)}$ for the 2nd quarter. Regarding the 3rd month, we estimate the model $x_m^{(1)} = f\left(x_m^{(3)}\right)$ based on the data up to the 1st month of the current quarter and then predict the $x_m^{(1)}$ for the 3rd quarter as a 2-step-ahead forecast (we know the values of $x_m^{(3)}$ for the 2nd and the 3rd quarters).

Let us now assume that the present time is the 2nd month of current quarter. The $x_m^{(1)}$ variable is published up to the 2nd month of previous quarter, whereas the $x_m^{(3)}$ variable is published up to the 1st month of current quarter. Regarding the 1st month, we have to estimate the model $x_m^{(1)} = f\left(x_m^{(3)}\right)$ based on the data up to the 2nd month of the previous quarter and then predict the $x_m^{(1)}$ (as a 2-step-ahead forecast) for the 1st quarter, since we know the value of $x_m^{(3)}$ for the 1st quarter. Regarding the 2nd month, the estimation of a $x_m^{(1)} = f\left(x_m^{(3)}\right)$ model is not helpful as the values of $x_m^{(3)}$ for the 2nd month are not published. Hence, we have to rely on another type of model such as a $x_m^{(1)} = f\left(x_{m-1}^{(3)}\right)$. And so on, for the 3rd month, a $x_m^{(1)} = f\left(x_{m-2}^{(3)}\right)$ model may be employed. As Schumacher and Breitungth (2008) accurately highlight in footnote 6 of page 392: "*Note that due to the publication lags of GDP, however, the effective forecast horizon needed for computing the forecasts has to be longer. For example, the data of vintage October 2004 (2004M10) contains GDP data up to 2004Q2 and monthly information up to 2004M9. For a forecast of the value in 2005Q1, we effectively need a three-quarter-ahead forecast from the end of the GDP sample.*"

Summing up, the framework for the private consumption is:

$$y_q^{(1)} = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}_{(m-\tau-is)}' \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \tag{5}$$

$$\begin{bmatrix} x_m^{(1)} \\ x_m^{(2)} \end{bmatrix} = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \tilde{\boldsymbol{X}}_m' + \begin{bmatrix} \varepsilon_{1,m} \\ \varepsilon_{2,m} \end{bmatrix}, \tag{6}$$

where $\beta_0$ is a scalar, $\boldsymbol{\theta}_j$ is a vector of coefficients, $\boldsymbol{X}_{(m)}' = \left[x_m^{(1)}, x_m^{(2)}\right]$, $\varepsilon_q \sim N\left(0, \sigma_{\varepsilon_q}^2\right)$, $\boldsymbol{\gamma_i} = [\gamma_{i,3} \ldots \gamma_{i,7}]$, $\tilde{\boldsymbol{X}}_m' = \left[x_m^{(3)} \ldots x_m^{(7)}\right]'$, $\varepsilon_{i,m} \sim N\left(0, \sigma_{\varepsilon_{i,m}}^2\right)$, $Cov\left(\varepsilon_{i,m}, \varepsilon_{i',m}\right) = 0$. The evaluation of the nowcasting accuracy showed that only the nested model with the retail trade volume index provides a better performance. The simultaneous inclusion of $x_m^{(1)}$ and $x_m^{(2)}$ creates multicollinearity issues; therefore, the outcome has a worse forecasting performance.[5]

The variables are seasonally adjusted with the X12 method and any variable with non-positive values is transformed as $x_m^* = x_m - min(x_m) + 1$. When we are in the 3rd month of the nowcasted quarter, the required values for the estimation of the proposed model framework are available at the time when we proceed with the estimations, i.e., do belong to the information set.

[5] The existence of multicollinearity deteriorates the forecasting accuracy.

When we are in the 2nd month of the current quarter and we want to nowcast the $\boldsymbol{X}'_{(m)}$ for the 1st month of the quarter, the $\widetilde{\boldsymbol{X}}'_m$ is available up to the 1st month of the current quarter. Hence, we estimate Eq. (5) with data up to the 2nd month of the previous quarter (because the values of $\boldsymbol{X}'_{(m)}$ are available up to the 2nd month of previous quarter) and predict the values of $\boldsymbol{X}'_{(m)}$ for the 1st quarter of current month as a 2-step-ahead forecast, i.e.,

$$\begin{bmatrix} x^{(1)}_{m\backslash m-2} \\ x^{(2)}_{m\backslash m-2} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-2)}_{1,0} \\ \gamma^{(m-2)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-2)}_1 \\ \boldsymbol{\gamma}^{(m-2)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_m. \tag{7}$$

We are still in the 2nd month of the current quarter and we want to nowcast the $\boldsymbol{X}'_{(m)}$ for the 2nd month of the quarter. Keeping in mind the publication lags, Eq. (5) is not usable for nowcasting 2nd month's values. So, we propose the estimation of a structure that provides nowcast values based on the available information set:

$$y^{(1)}_q = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \tag{8}$$

$$\begin{bmatrix} x^{(1)}_m \\ x^{(2)}_m \end{bmatrix} = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-1} + \begin{bmatrix} \varepsilon_{1,m} \\ \varepsilon_{2,m} \end{bmatrix}. \tag{9}$$

In this case, the nowcasting of $\boldsymbol{X}'_{(m)}$ is a 3-step-ahead forecast. We estimate Eq. (8) with the data up to the 2nd month of previous quarter (i.e., the values of $\boldsymbol{X}'_{(m)}$ are available up to the 2nd month of previous quarter) and then predict the values of $\boldsymbol{X}'_{(m)}$ for the 2nd quarter of current month as a 3-step-ahead forecast:

$$\begin{bmatrix} x^{(1)}_{m\backslash m-3} \\ x^{(2)}_{m\backslash m-3} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-3)}_{1,0} \\ \gamma^{(m-3)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-3)}_1 \\ \boldsymbol{\gamma}^{(m-3)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-1}. \tag{10}$$

Finally, we are in the 2nd month of the current quarter and we want to nowcast the $\boldsymbol{X}'_{(m)}$ for the 3rd month of the quarter. Hence, the real-time nowcast for the third month of the quarter is computed as a 4-step-ahead forecast:

$$\begin{bmatrix} x^{(1)}_{m\backslash m-4} \\ x^{(2)}_{m\backslash m-4} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-4)}_{1,0} \\ \gamma^{(m-4)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-4)}_1 \\ \boldsymbol{\gamma}^{(m-4)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-2}, \tag{11}$$

which is based on the model, $\begin{bmatrix} x^{(1)}_m & x^{(2)}_m \end{bmatrix}' = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \widetilde{\boldsymbol{X}}'_{m-2} + \begin{bmatrix} \varepsilon_{1,m} & \varepsilon_{2,m} \end{bmatrix}$.

Maintaining the same rationale, when we are in the 1st month of the current quarter, and we want to nowcast the $\boldsymbol{X}'_{(m)}$ for:

a. the 1st month of the quarter, based on the model $\begin{bmatrix} x^{(1)}_m & x^{(2)}_m \end{bmatrix}' = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \widetilde{\boldsymbol{X}}'_{m-1} + \begin{bmatrix} \varepsilon_{1,m} & \varepsilon_{2,m} \end{bmatrix}$, then real-time nowcast is computed as a 3-step-ahead forecast:

$$\begin{bmatrix} x^{(1)}_{m\backslash m-3} \\ x^{(2)}_{m\backslash m-3} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-3)}_{1,0} \\ \gamma^{(m-3)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-3)}_1 \\ \boldsymbol{\gamma}^{(m-3)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-1}, \tag{12}$$

b. the 2nd month of the quarter, based on the model $\begin{bmatrix} x^{(1)}_m & x^{(2)}_m \end{bmatrix}' = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \widetilde{\boldsymbol{X}}'_{m-2} + \begin{bmatrix} \varepsilon_{1,m} & \varepsilon_{2,m} \end{bmatrix}$, then the real-time nowcast is computed as a 4-step-ahead forecast:

$$\begin{bmatrix} x^{(1)}_{m\backslash m-4} \\ x^{(2)}_{m\backslash m-4} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-4)}_{1,0} \\ \gamma^{(m-4)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-4)}_1 \\ \boldsymbol{\gamma}^{(m-4)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-2}, \tag{13}$$

c. the 3rd month of the quarter, based on the model $\begin{bmatrix} x^{(1)}_m & x^{(2)}_m \end{bmatrix}' = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \widetilde{\boldsymbol{X}}'_{m-3} + \begin{bmatrix} \varepsilon_{1,m} & \varepsilon_{2,m} \end{bmatrix}$, then the real-time nowcast is computed as a 5-step-ahead forecast:

$$\begin{bmatrix} x^{(1)}_{m\backslash m-5} \\ x^{(2)}_{m\backslash m-5} \end{bmatrix} = \begin{bmatrix} \gamma^{(m-5)}_{1,0} \\ \gamma^{(m-5)}_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}^{(m-5)}_1 \\ \boldsymbol{\gamma}^{(m-5)}_2 \end{bmatrix} \tilde{\boldsymbol{X}}'_{m-3}. \tag{14}$$

### 5.2 Government spending on public goods and services

An annual estimate of government spending is published in the state budget. Usually, the state budget report is submitted between the last days of October and the first days of November. The figures are presented on an annual basis. Thus, we can infer the estimate of government spending for the last quarter of the current year. If $\widehat{y}^{(2)}_a$ is the estimated annual growth of government spending for year $a$, we can nowcast the

**Table 4** Publication of information across time for export of goods

| | Quarter for nowcasting | | | Next quarter |
|---|---|---|---|---|
| | 1st month ($m-2$) | 2nd month ($m-1$) | 3rd month ($m$) | |
| Balance of payments | $x_{m-2}$ | | | |
| PMI index | | $pmi_{m-1}^{(exp)}$ | | |

public consumption for the last quarter of the year as $\widehat{Y}_q^{(2)} = \left(Y_{a-1}^{(2)}\left(1+\widehat{y}_a^{(2)}\right)\right) - \left(\sum_{i=1}^{3} Y_{q-i}^{(2)}\right)$.

The $\widehat{y}_a^{(2)}$ is the official nowcast that incorporates all the available information for government spending and is published between the 1st and 2nd month of last quarter. Hence, the $\widehat{Y}_q^{(2)}$ has those characteristics necessary to be considered a landmark estimator. But, if we measure the nowcasting accuracy of $\widehat{y}_a^{(2)}$ based on the mean absolute percentage error of the last quarter of each year, we reach a value of $(Q/4)^{-1}\sum_{q=1(4)}^{Q}\left|\widehat{Y}_q^{(2)} - Y_q^{(2)}\right|/Y_q^{(2)} = 10.31\%$.

In order to proceed with an evaluation of the official nowcasting of government spending, we estimate the forecast values of $Y_q^{(2)}$ from the random walk model. The mean absolute percentage error for the whole period equals 10.76%, which is very close to that of official nowcast. On the other hand, a naïve first-order autoregressive model of the form:

$$\begin{aligned} &(1-L)logY_q^{(2)} = \beta_0 + e_q, \\ &e_q = \beta_1 e_{q-1} + \varepsilon_q, \end{aligned} \tag{15}$$

for $\varepsilon_q \sim N\left(0, \sigma_\varepsilon^2\right)$, leads to $\mathrm{Q}^{-1}\sum_{\mathrm{q}=1}^{\mathrm{Q}}\left|\mathrm{Y}_{\mathrm{q}+1\backslash q}^{(2)} - \mathrm{Y}_{\mathrm{q}+1}^{(2)}\right| / \mathrm{Y}_{\mathrm{q}+1}^{(2)} = 2.33\%$, where $\mathrm{Y}_{\mathrm{q}+1\backslash q}^{(2)} = \beta_0^{(q)}\left(1-\beta_1^{(q)}\right) + \beta_1^{(q)}\mathrm{Y}_{\mathrm{q}}^{(2)}$ is the one-quarter ahead forecast based on the information set of the previous quarter. Hence, we select the AR(1) model, as it leads to much lower forecast errors.[6]

### 5.3 Investment on business capital goods

For the nowcasting of investments, the preliminary analysis provides strong evidence for the usability of the daily financial data. More specifically, we observe that Athens stock exchange main general index, $Z_d^{(1)}$, and Greek 10-year government bond yield, $Z_d^{(2)}$, are adequate explanatory variables for $\left(Y_q^{(3)}\right)$. Hence, the proposed model framework is estimated in the form:

$$y_q^{(3)} = \beta_0 + \sum_{r=0}^{l-1} \mathbf{Z}'_{(d-r-is)}\left(\sum_{j=0}^{q} r^j \boldsymbol{\varphi}_j\right) + \varepsilon_q, \tag{16}$$

where $\beta_0$ is a scalar coefficient, $\boldsymbol{\varphi}_j$ is a vector of coefficients, $\boldsymbol{Z}'_{(d)} = \left[z_d^{(1)}, z_d^{(2)}\right]$ denotes the vector of variables observed at a daily frequency, i.e., $z_d^{(1)} = log\left(Z_d^{(1)}/Z_{d-1}^{(1)}\right)$ and $z_d^{(2)} = log\left(Z_d^{(2)}/Z_{d-1}^{(2)}\right)$, and $s = 66$ denotes the number of trading days of each quarter and $\varepsilon_q \sim N\left(0, \sigma_{\varepsilon_q}^2\right)$. Regarding the $i$ term, for $i = 0$, we estimate $y_{q\backslash q}^{(3)}$, and for $i \geq 1$ and $is \geq 66$, we estimate the one-month ahead $y_{q+1\backslash q}^{(3)}$, and when $i \geq 2$ and $is \geq 132$, we estimate the two-month ahead $y_{q+2\backslash q}^{(3)}$, and so on.

### 5.4 Exports of goods

The quarterly export of goods and services is related to the export of fuels, vessels, other services, travel receipts, transportation receipts, and other receipts, which are available on a monthly basis from the balance of payments:

$$y_q^{(4)} = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}'_{(m-\tau-is)}\left(\sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j\right) + \varepsilon_q, \tag{17}$$

Thus, for $y_q^{(4)}$ the $\boldsymbol{X}_{(m)}$ includes information available from the balance of payments on a monthly frequency. More specifically, we define $\boldsymbol{X}_{(m)} \equiv x_{(m)} = \sum_{k=1}^{3} x_m^{(k)} + 0.2x_m^{(4)}$, for export of fuels $x_m^{(1)}$, export of vessels $x_m^{(2)}$, other exports $x_m^{(3)}$ and travel receipts $x_m^{(4)}$. These variables are in nominal values and not seasonally adjusted; thus, the $x_{(m)}$ is seasonally adjusted with the X12 method.

With $x_{(m)}$ we have reached to a model that nowcasts the values that have not been published based on information available for the seasonally adjusted Purchasing Managers' sub index New Export Orders:

$$(1-L)log(x_m) = \gamma_0 + \gamma_1(1-L)log\left(pmi_{m-1}^{(exp)}\right) + \varepsilon_m. \tag{18}$$

[6] For comparability, the mean absolute percentage forecast error taking only the last quarter of each year is 2.28%.

The balance of payments is published with a lag of 2 months. Thus, we estimate the model based on the most recently available information set, $\boldsymbol{I}_m = \left\{x_{m-2}, pmi_{m-1}^{(exp)}\right\}$. When we are in the 3rd month of the quarter, we can estimate the coefficients of the model $\gamma_0^{(m-2)}, \gamma_1^{(m-2)}$. The real-time nowcast of $x_m$ for the 3rd month of the quarter is computed as:

$$x_{m\backslash m-2} = e^{log(x_{m-1\backslash m-2})+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-1}^{(exp)}\right)}. \tag{19}$$

Table 4 visualizes the publication of information across time. We are interested to nowcast for the current quarter, and at the present time we are in the 3rd month of the quarter, or $(m)$. The PMI is published with a lag of 1 month, or $(m-1)$, whereas the balance of payments is published with a lag of 2 months, or $(m-2)$. So, the model can be estimated with the most recent information available at $(m-2)$. Afterwards, we are able to compute the 1-step-ahead forecast $x_{m-1\backslash m-2}$ as the $pmi_{m-2}^{(exp)}$ is available. Additionally, we can compute the 2-step-ahead forecast $x_{m\backslash m-2}$ as the $pmi_{m-1}^{(exp)}$ is also available.

$$log(x_m) = (1+\gamma_1)log(x_{m-1}) - \gamma_1 log(x_{m-2}) + \gamma_0(1-\gamma_1) + \varepsilon_m. \tag{21}$$

Keeping in mind the publication lag of 2 months, the adequate prediction scheme is $x_{m\backslash m-3} = e^{\left(1+\gamma_1^{(m-3)}\right)} log\left(x_{m-1\backslash m-3}\right) - \gamma_1^{(m-3)} log\left(x_{m-2\backslash m-3}\right) + \gamma_0^{(m-3)}\left(1-\gamma_1^{(m-3)}\right)$, where $x_{m-1\backslash m-3}$ and $x_{m-2\backslash m-3}$ should also be computed iteratively. Also, the real-time nowcast of $x_m$ for the 2nd month of the quarter is computed as:

$$x_{m\backslash m-2} = e^{log(x_{m-1\backslash m-2})+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-1}^{(exp)}\right)}. \tag{22}$$

Finally, for the 1st month of the quarter, the $x_m$ is estimated as:

$$x_{m-1\backslash m-2} = e^{log(x_{m-2})+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{23}$$

When we are in the 1st month of the quarter, the real-time nowcast of $x_m$ for the 3rd month of the quarter is estimated by the autoregressive pattern previously defined. Hence, the adequate prediction scheme is:

$$x_{m\backslash m-4} = e^{(1+\gamma_1(m-4))\log\left(x_{m-1\backslash m-4}\right)-\gamma_1^{(m-4)}\log\left(x_{m-2\backslash m-4}\right)+\gamma_0^{(m-4)}\left(1-\gamma_1^{(m-4)}\right)}. \tag{24}$$

Similarly, for the 2nd month of the quarter:

$$x_{m-1\backslash m-4} = e^{(1+\gamma_1(m-4))\log\left(x_{m-2\backslash m-4}\right)-\gamma_1^{(m-4)}\log\left(x_{m-3\backslash m-4}\right)+\gamma_0^{(m-4)}\left(1-\gamma_1^{(m-4)}\right)} \tag{25}$$

Also, the real-time nowcast of $x_m$ for the 2nd month of the quarter is computed as:

$$x_{m-1\backslash m-2} = e^{log(x_{m-2})+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{20}$$

Finally, for the 1st month of the quarter, the $x_m$ has been published.

When we are in the 2nd month of the quarter, the real-time nowcast of $x_m$ for the 3rd month of the quarter cannot be nowcasted as $pmi^{(exp)}$ the 2nd month is not published. So, we estimate a time series model that captures the autoregressive pattern:

Regarding the 1st month of the quarter, the available infmation set is $\boldsymbol{I}_m = \left\{x_{m-2}, pmi_{m-1}^{(exp)}\right\}$, so the real-time nowcast is computed as:

$$x_{m\backslash m-2} = e^{log(x_{m-1\backslash m-2})+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-1}^{(exp)}\right)}. \tag{26}$$

### 5.5 Exports of services

For the export of services, $y_q^{(5)}$, we define $\boldsymbol{X}_{(m)} \equiv x_{(m)} = 0.8x_m^{(4)} + x_m^{(5)} + x_m^{(6)}$, for travel receipts $x_m^{(4)}$, transportation receipts $x_m^{(5)}$ and other receipts $x_m^{(6)}$:

$$y_q^{(5)} = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \tag{27}$$

For the seasonally adjusted $x_{(m)}$ we have reached to a model that nowcasts the values that have not been published based on the information that is available for the seasonally adjusted Purchasing Managers' sub index New Export Orders:

$$(1-L)log(x_m) = \gamma_0 + \gamma_1(1-L)log\left(pmi_{m-2}^{(exp)}\right) + \varepsilon_m. \tag{28}$$

The rationale behind the computations is in line with the approach followed for the export of goods and is available in Appendix B.

### 5.6 Imports of goods

For the quarterly import of goods $y_q^{(6)}$ the $\boldsymbol{X}_{(m)}$ is expressed with the information available from the balance of payments at a monthly frequency; $\boldsymbol{X}_{(m)} \equiv x_{(m)} = \sum_{k=7}^{9} x_m^{(k)}$, where $k = 7, 8, 9$ denotes importation of fuels, importation of vessels and importation of other goods, respectively. These variables are in nominal values and not seasonally adjusted; thus, the $x_{(m)}$ is seasonally adjusted with the X12 method. For $x_{(m)}$ we have reached to a model that nowcasts the monthly non-published values based on the first-order autoregressive pattern of *q-o-q* log returns:

$$\begin{aligned} y_q^{(6)} &= \beta_0 + \sum_{\tau=0}^{\kappa-1} (1-L)\boldsymbol{lX}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \\ log(x_m) &= (1+\gamma_1)log(x_{m-1}) - \gamma_1 log(x_{m-2}) + \gamma_0(1-\gamma_1) + \varepsilon_m, \\ \varepsilon_q &\sim N\left(0, \sigma_{\varepsilon_q}^2\right) \text{ and } \varepsilon_m \sim N\left(0, \sigma_{\varepsilon_m}^2\right). \end{aligned} \tag{29}$$

The $\boldsymbol{lX}_{(m)}$ denotes the vector of log-transformation of the variables $x_{(m)}$.

### 5.7 Imports of services

For the quarterly import of services, $y_q^{(7)}$, the balance of payments provides all the necessary information for the estimation of non-published values. We define $\boldsymbol{X}_{(m)} \equiv x_{(m)} = \sum_{k=10}^{12} x_m^{(k)}$, where $k = 10, 11, 12$ denotes travel receipts, transportation receipts and other receipts, respectively. As in the previous sections, the $x_{(m)}$ is seasonally adjusted with the X12 method. For $x_{(m)}$ we have reached to a model that nowcasts the monthly non-published values based on the second-order autoregressive pattern of *q-o-q* log returns:

$$\begin{aligned} y_q^{(7)} &= \beta_0 + \sum_{\tau=0}^{\kappa-1} (1-L)log\boldsymbol{X}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \\ log(x_m) &= (1+\gamma_1)log(x_{m-1}) - (\gamma_1-\gamma_2)log(x_{m-2}) \\ &\quad - \gamma_2 log(x_{m-3}) + \gamma_0(1-\gamma_1-\gamma_2) + \varepsilon_m, \varepsilon_q \\ &\sim N\left(0, \sigma_{\varepsilon_q}^2\right) \text{ and } \varepsilon_m \sim N\left(0, \sigma_{\varepsilon_m}^2\right). \end{aligned} \tag{30}$$

### 5.8 Changes in inventories

The changes in inventories are fully unpredictable, but with an autocorrelated structure across quarters. Hence, we assume, a priori, that for the future value of $Y_q^{(8)}$, the only available information is its first-order autocorrelated structure:

$$\begin{aligned} Y_q^{(8)} &= \beta_0 + e_q. \\ e_q &= \beta_1 e_{q-1} + \varepsilon_q, \end{aligned} \tag{31}$$

for $\varepsilon_q \sim N\left(0, \sigma_\varepsilon^2\right)$. So, the nowcast value for the changes in inventories is the one-quarter ahead forecast based on the information set of the previous quarter: $\mathrm{Y}_{\mathrm{q+1\backslash q}}^{(8)} = \beta_0^{(q)}\left(1-\beta_1^{(q)}\right) + \beta_1^{(q)}\mathrm{Y}_{\mathrm{q}}^{(8)}$.[7]

## 6 Robustness tests

A number of additional model extensions for robustness purposes are discussed in the paragraphs which follow:

1. The Midas models have been replaced by regression models aggregating the data from a higher sampling frequency to a quarterly frequency. According to the findings presented in Sect. 8, the use of mixed-data sampling frequency estimators is definitely necessary for returning accurate nowcasts.
2. In Sect. 5.1, the variable selection is plausible, but how do we know that the omitted variables do not help? If that was the case then we may prefer a data driven way that would have examined all the available explanatory variables. Of course, the use of explanatory variables that are highly linearly related induces the problem of multicollinearity. A common strategy to reduce the risk of multicollinearity is the estimation of factors that express the majority of the variability of the original variables. Principal component analysis has been applied for the estimation of the factors. Illustratively, in the case of private

[7] The random walk, 1st difference transformation, distributed lagged models have also been tested, but the AR(1) performs better.

consumption on goods and services, we present the replacement of monthly confidence and sentiment indicators with factors estimated from the PCA. Let us define as $\widetilde{\boldsymbol{X}}_m$ the matrix with the $M$ selected variables for $m^{th}$ month. The factors are estimated as:

$$\tilde{\boldsymbol{X}}_m = \boldsymbol{\Lambda}\boldsymbol{X}_m^{(f)} + \boldsymbol{e}_m, \tag{32}$$

where $\boldsymbol{\Lambda}$ is the matrix of factor loadings, $\boldsymbol{X}_m^{(f)} = \left[f_m^{(1)}, \ldots, f_m^{(M)}\right]'$ is the vector with the common factors, and $\boldsymbol{e}_m$ is the vector of the idiosyncratic component. Summing up, the private consumption model is estimated as:

$$y_q^{(1)} = \beta_0 + \sum_{\tau=0}^{\kappa-1} \boldsymbol{X}'_{(m-\tau-is)} \left( \sum_{j=0}^{p} \tau^j \boldsymbol{\theta}_j \right) + \varepsilon_q, \tag{33}$$

$$\begin{bmatrix} x_m^{(1)} \\ x_m^{(2)} \end{bmatrix} = \begin{bmatrix} \gamma_{1,0} \\ \gamma_{2,0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix} \boldsymbol{X}_m^{(f)} + \begin{bmatrix} \varepsilon_{1,m} \\ \varepsilon_{2,m} \end{bmatrix}, \tag{34}$$

where $\beta_0$ is a scalar coefficient, $\boldsymbol{\theta}_j$ is a vector of coefficients, $\boldsymbol{X}'_{(m)} = \left[x_m^{(1)} x_m^{(2)}\right]$, $\varepsilon_q \sim N\left(0, \sigma^2_{\varepsilon_q}\right)$, $\boldsymbol{\gamma_i} = [\gamma_{i,3} \ldots \gamma_{i,7}]$, $\boldsymbol{X}_m^{(f)} = \left[f_m^{(1)}, \ldots, f_m^{(4)}\right]$, $N\left(0, \sigma^2_{\varepsilon_{i,m}}\right)$, $Cov\left(\varepsilon_{i,m}, \varepsilon_{i',m}\right) = 0$. The model has been estimated with 4 as well as with 2 factors and the forecasting accuracy was statistically indistinguishable.[8]

3. For the investment nowcasting, we have estimated models by adding explanatory variables available at a monthly frequency. The most informative model specifications are still those based on the Athens stock exchange main general index, $Z_d^{(1)}$, and the Greek 10-year government bond yield, $Z_d^{(2)}$. The only monthly variable that has satisfactory nowcasting accuracy is the financial conditions index, $x_m^{(6)}$. However, none of the additional models are able to perform better when compared to those based on the daily dataset. In Sect. 8, the additional models are also presented.

**Table 5** D-model's estimated coefficients with their $p$ values based on the latest available information set

| | Estimated coefficient | *p* value | | Estimated coefficient | *p* value |
|---|---|---|---|---|---|
| Private consumption | | | | | |
| $\beta_0$ | 13,007.74 | 0.0000 | $\theta_1$ | 130.5150 | 0.0002 |
| $\theta_2$ | −55.78792 | 0.0020 | $\theta_3$ | 5.288621 | 0.0031 |
| $\gamma_{1,0}$ | 102.7245 | 0.0000 | $\gamma_{2,0}$ | 107.7899 | 0.0000 |
| $\gamma_{1,3}$ | 1.126139 | 0.0000 | $\gamma_{2,3}$ | 1.138982 | 0.0000 |
| $\gamma_{1,4}$ | 0.909918 | 0.0000 | $\gamma_{2,4}$ | 0.480416 | 0.0000 |
| $\gamma_{1,5}$ | 0.217591 | 0.1828 | $\gamma_{2,5}$ | 0.219026 | 0.0806 |
| $\gamma_{1,6}$ | −2.763324 | 0.0000 | $\gamma_{2,6}$ | −3.336665 | 0.0000 |
| $\gamma_{1,7}$ | −0.816437 | 0.0045 | $\gamma_{2,7}$ | −0.329955 | 0.1316 |
| Government spending | | | | | |
| $\beta_0$ | −0.002027 | 0.6184 | $\beta_1$ | 0.021532 | 0.8335 |
| Investment | | | | | |
| $\beta_0$ | −0.007570 | 0.3269 | | | |
| $\varphi_{1,1}$ | 5.256264 | 0.0008 | $\varphi_{2,1}$ | −0.132306 | 0.0326 |
| $\varphi_{1,2}$ | −4.307391 | 0.0022 | $\varphi_{2,2}$ | 0.013762 | 0.2169 |
| $\varphi_{1,3}$ | 0.727777 | 0.0051 | $\varphi_{2,3}$ | −0.000349 | 0.0982 |
| Exports of goods | | | | | |
| $\beta_0$ | 843.1290 | 0.0000 | $\theta_1$ | 1.276792 | 0.0000 |
| $\theta_2$ | −0.415189 | 0.0000 | $\theta_3$ | 0.033869 | 0.0004 |
| $\gamma_0$ | 0.003714 | 0.4766 | $\gamma_1$ | 0.131338 | 0.0115 |
| Exports of services | | | | | |
| $\beta_0$ | −1550.526 | 0.0001 | $\theta_1$ | 2.480880 | 0.0000 |
| $\theta_2$ | −0.727720 | 0.0000 | $\theta_3$ | 0.048568 | 0.0000 |
| $\gamma_0$ | 0.001441 | 0.7002 | $\gamma_1$ | 0.069996 | 0.0596 |
| Imports of goods | | | | | |
| $\beta_0$ | −7.22E−05 | 0.9852 | $\theta_1$ | −0.437440 | 0.0000 |
| $\theta_2$ | 0.726357 | 0.0000 | $\theta_3$ | −0.117135 | 0.0000 |
| $\gamma_0$ | 0.001164 | 0.6943 | $\gamma_1$ | −0.349545 | 0.0000 |
| Imports of services | | | | | |
| $\beta_0$ | 0.002177 | 0.6584 | $\theta_1$ | −0.583303 | 0.0000 |
| $\theta_2$ | 1.047051 | 0.0000 | $\theta_3$ | −0.173271 | 0.0000 |
| $\gamma_0$ | 0.001775 | 0.5597 | $\gamma_1$ | −0.582359 | 0.0000 |
| $\gamma_2$ | −0.254289 | 0.0000 | | | |
| Changes in inventories | | | | | |
| $\beta_0$ | −0.002027 | 0.6184 | $\beta_1$ | 0.021532 | 0.8335 |

On the quarterly (monthly) frequency, the estimated parameters refer to current quarter (2nd month of current quarter), based on information available on the 3rd month of current quartet

## 7 Nowcasting with naive models

For benchmark purposes, a random walk (the projected growth rate is the most recently available plus the average log-growth), a first-order autoregressive model, and a regression model on a quarterly frequency using the same information as in the disaggregate framework are estimated for the quarterly data. The models are considered in the forms:

[8] Schumacher and Breitungth (2008) noted that the forecast performance declines as the number of factors increased from one to three. On the other hand, Stock and Watson (2002a) found that the large number of factors do not affect the forecasting accuracy.

**Table 6** MAPE and RMSE loss functions for the consumption models

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current Quarter | M1 | 3.51 | 1431.1 | 9.70 | 3828.7 | | | | |
| | M2 | 3.26 | 1352.4 | 9.69 | 3838.4 | | | | |
| | M3 | 3.30 | 1364.7 | 9.44 | 3728.1 | 2.34 | 1363.0 | 2.43 | 1363.1 |
| Next Quarter | M1 | 2.59 | 1088.8 | 6.18 | 2410.0 | | | | |
| | M2 | 1.74 | 899.9 | 3.24 | 1405.6 | | | | |
| | M3 | | | | | 1.79 | 1158.0 | 1.73 | 1075.9 |

**Table 7** MAPE and RMSE loss functions for government spending on public goods and services

| Q | AR | | RW | |
|---|---|---|---|---|
| | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | | | | |
| | 3.47 | 533.5 | 17.79 | 1803.4 |
| Next quarter | | | | |
| | 2.56 | 363.7 | 17.37 | 1761.8 |

### 7.1 Random walk

$$y_q^{(0)} = \beta_0 + \varepsilon_q,\ \varepsilon_q \sim N\left(0, \sigma_q^2\right), \tag{35}$$

where $y_q^{(0)} = log\left(Y_q^{(0)}/Y_{q-1}^{(0)}\right)$. is the *q-o-q* GDP growth rate.

### 7.2 First-order autoregressive

$$y_q^{(0)} = \beta_0 + \beta_1\left(y_{q-1}^{(0)} - \beta_0\right) + \varepsilon_q,\ \varepsilon_q \sim N\left(0, \sigma_q^2\right). \tag{36}$$

### 7.3 Regression model

$$y_q^{(0)} = \beta_0 + \boldsymbol{B}\boldsymbol{X}_{q-1}^{(f)} + \varepsilon_q,\ \varepsilon_q \sim N\left(0, \sigma_q^2\right), \tag{37}$$

where $\boldsymbol{X}_q^{(f)}$ is the vector with the common factors from the PCA dimension reduction method: $\boldsymbol{X}_q = \boldsymbol{\Lambda}\boldsymbol{X}_q^{(f)} + \boldsymbol{e}_q$. The $\boldsymbol{X}_q$ includes all the explanatory variables on a quarterly frequency.

## 8 Nowcasting evaluation

The nowcast evaluation focuses on answering the research question: If we proceed to a more complicated prediction task for GDP nowcasting (as the proposed framework), what is the forecast accuracy gain, compared to simpler nowcasting techniques? We answer this question comparing the forecasting accuracy of the disaggregated Midas model against simpler nowcasting approaches as (i) the disaggregated regression model (i.e., not mixed-frequency modeling based solely on quarterly data), (ii) the aggregated regression model (i.e., neither mixed-frequency modeling nor disaggregated dataset), and (iii) the naïve model techniques (no pain at all!).

As Barhoumi et al. (2008) noted, the nowcasting evaluation exercise must replicate the data availability situation that is faced in the real-time application of the models. As Diebold (2020) noted there are four approaches in forecasting evaluation. (1) Approach based on full-sample estimation and final revised data; (2) approach based on expanding sample estimation and final revised data; (3) approach based on expanding sample estimation and vintage data; and (4) approach based on expanding sample estimation and vintage information. Our nowcasting exercise is based on a sequence of pseudo out-of-sample nowcasts over the

**Table 8** MAPE and RMSE loss functions for investments models

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 3.28 | 357.1 | 5.06 | 763.3 | | | | |
| | M2 | 3.29 | 360.9 | 4.96 | 755.6 | | | | |
| | M3 | 3.61 | 345.5 | 4.95 | 750.8 | 8.22 | 1181.7 | 8.10 | 1172.7 |
| Next quarter | M1 | 3.77 | 389.1 | 4.95 | 750.8 | | | | |
| | M2 | 3.77 | 389.1 | 4.95 | 750.8 | | | | |
| | M3 | | | | | 5.47 | 843.4 | 5.32 | 825.9 |

**Table 9** MAPE and RMSE loss functions for export of goods models (in nominal values)

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 4.66 | 413.9 | 5.61 | 496.0 | | | | |
| | M2 | 4.52 | 360.9 | 5.65 | 432.8 | | | | |
| | M3 | 3.63 | 301.7 | 4.82 | 381.4 | 6.68 | 542.3 | 6.67 | 537.6 |
| Next quarter | M1 | 2.80 | 234.0 | 3.71 | 306.0 | | | | |
| | M2 | 2.62 | 215.5 | 3.31 | 273.0 | | | | |
| | M3 | | | | | 4.85 | 427.2 | 4.82 | 429.7 |

**Table 10** MAPE and RMSE loss functions for the export of services models (in nominal values)

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 7.12 | 865.2 | 7.98 | 913.0 | | | | |
| | M2 | 5.60 | 598.5 | 6.65 | 739.2 | | | | |
| | M3 | 5.27 | 519.6 | 6.53 | 664.3 | 7.98 | 6.65 | 6.53 | 5.48 |
| Next quarter | M1 | 4.59 | 443.0 | 5.48 | 509.2 | | | | |
| | M2 | 4.13 | 404.6 | 4.93 | 458.0 | | | | |
| | M3 | | | | | 7.12 | 5.60 | 5.27 | 4.59 |

**Table 11** MAPE and RMSE loss functions for the import of goods models (in nominal values)

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 4.01 | 787.4 | 4.35 | 861.2 | | | | |
| | M2 | 3.94 | 743.1 | 4.31 | 838.7 | | | | |
| | M3 | 3.16 | 558.1 | 3.49 | 605.1 | 6.12 | 1329.9 | 6.14 | 1308.9 |
| Next quarter | M1 | 2.67 | 429.3 | 2.67 | 439.9 | | | | |
| | M2 | 2.35 | 380.0 | 2.40 | 385.3 | | | | |
| | M3 | | | | | 4.65 | 987.7 | 4.58 | 914.9 |

**Table 12** MAPE and RMSE loss functions for the import of services models (in nominal values)

| Q | M | Midas | | Regression | | AR | | RW | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAP (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 4.31 | 253.2 | 4.79 | 280.8 | | | | |
| | M2 | 4.38 | 232.0 | 4.79 | 261.7 | | | | |
| | M3 | 3.40 | 175.7 | 3.83 | 188.2 | 8.44 | 426.2 | 8.43 | 425.9 |
| Next quarter | M1 | 3.17 | 161.2 | 3.19 | 163.4 | | | | |
| | M2 | 2.95 | 147.6 | 2.90 | 151.2 | | | | |
| | M3 | | | | | 5.26 | 292.2 | 5.20 | 287.7 |

**Table 13** MAE and RMSE loss functions for changes in inventories

| Q | AR | | RW | |
|---|---|---|---|---|
| | MAE | RMSE | MAE | RMSE |
| Current quarter | 673.4 | 953.0 | 2.1 | 1002.0 |
| Next quarter | 586.9 | 821.9 | 2.0 | 980.9 |

Due to the negative values of changes in inventories, the mean absolute error is computed instead of MAPE

evaluation sample based on the final revised data, as vintage data are not available for the Greek economy. A real-time evaluation is truly credible if it is based on vintage information and it is obtained by using nowcasts produced and permanently recorded in real time. Unfortunately, we are not able to provide an evaluation based on vintage information. But, given the data availability, we have produced nowcasts based on final revised data that were available at the time the model was to be estimated. For example, let us assume that we estimate the model that produces the nowcast of private consumption,$Y_q^{(1)}$, and the explanatory variable is the retail trade volume index, $x_m^{(1)}$. The $x_m^{(1)}$ for June is published on 31st of August. If we nowcast $Y_q^{(1)}$ for the Q3 on June, then the $x_m^{(1)}$ of June will not be imported in the information set. But, if we nowcast $Y_q^{(1)}$ for the Q3 on September, then the $x_m^{(1)}$ of June will be imported in the information set.

The coefficients of the proposed model framework are estimated recursively each month. It is very difficult to present the coefficient estimates for all the components of GDP, all the layers and all the possible combinations of nowcasted-month and publication-month. What is relevant for practitioners would be to know the estimated coefficients of the model using the latest data. Thus, in Table 5, we present the estimated coefficients with their $p$ values based on the latest available information set. Moreover, we present in appendix C for the private consumption only, line plots of the estimated coefficients across time and the relative $p$ values. On a quarterly frequency, the estimated parameters refer to the recursive estimation of private consumption for the current quarter, based on information available on the 3rd month of current quartet (see Eq. 5). On a monthly frequency, we present the estimated parameters for the 2nd month of current quarter, based on information available on the 3rd month of current quartet (see Eq. 6). We infer that the values of coefficients change over time, mainly gradually, reflecting the updates of the information set. The parameters are not statistically significant across the total period under evaluation, replicating the changes in the relationships among the various variables.

**Table 14** Robustness purposes: investment model MAPE and RMSE loss functions

| Q | M | Midas (with $z_d^{(2)}, x_m^{(6)}$)) | | Midas (with $z_d^{(1)}, x_m^{(6)}$)) | | Regression (with $z_q^{(2)}$) | | Regression (with $x_q^{(6)}$) | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 3.68 | 404.7 | 3.42 | 407.1 | 5.06 | 763.3 | 4.95 | 776.8 |
| | M2 | 3.58 | 394.6 | 3.35 | 454.0 | 4.96 | 755.6 | 5.01 | 771.8 |
| | M3 | 3.90 | 425.5 | 3.60 | 405.9 | 4.95 | 750.8 | 4.91 | 765.8 |
| Next quarter | M1 | 3.22 | 364.6 | 3.45 | 353.2 | 4.95 | 750.8 | 4.96 | 768.2 |
| | M2 | 3.22 | 364.6 | 3.45 | 353.2 | 4.95 | 750.8 | 4.96 | 768.2 |

**Table 15** MAPE and RMSE loss functions for the disaggregated GDP models

| Q | M | Disaggregated Midas | | Disaggregated regression | | AR | | RW | | Regression model | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 3.09 | 1831.0 | 6.65 | 3816.9 | | | | | | |
| | M2 | 2.75 | 1653.6 | 6.52 | 3665.2 | | | | | | |
| | M3 | 2.90 | 1718.8 | 6.54 | 3650.9 | 2.44 | 1938.4 | 2.54 | 1960.5 | | |
| Next quarter | M1 | 2.19 | 1378.3 | 4.23 | 2317.5 | | | | | | |
| | M2 | 1.77 | 1255.2 | 2.14 | 1322.2 | | | | | | |
| | M3 | | | | | 1.59 | 1335.1 | 1.55 | 1285.9 | 2.01 | 1563.8 |

The $Y_q^{(0)}$ nowcasts are estimated for the quarters 2005Q1 to 2020Q3. For each quarter, we provide 5 different nowcasts of GDP, depending on the time we proceed to the estimation of the nowcasting. So, we estimate the GDP assuming that we are in the 1st month of the current quarter, the 2nd month of current quarter and so on, up to the 2nd month of the next quarter. The loss functions on which the forecasting evaluation is based on are *i*) the mean absolute percentage distance, MAPE, between actual and estimated GPD and *ii*) the root mean squared error, RMSE. So, we evaluate the nowcasting accuracy based on the GDP in billions, for quarter *q:*

$$MAPE = Q^{-1} \sum_{q=1}^{Q} \frac{\left| Y_q^{(0)} - \hat{Y}_q^{(0)} \right|}{Y_q^{(0)}}, \tag{38}$$

and

$$RMSE = \sqrt{Q^{-1} \sum_{q=1}^{Q} \left( Y_q^{(0)} - \hat{Y}_q^{(0)} \right)^2}, \tag{39}$$

where $\widehat{Y}_q^{(0)}$ is the GDP nowcast. The Hansen et al. (2011) Model Confidence Set is utilized in order to define the set of models that consists of the best nowcasting models, according to our predefined MAPE and RMSE loss functions. The null hypothesis $H_{0,M} : E\left(d_{(j),(j^*),q}\right) = 0$, for $\forall\ j, j^* \in M$ $M \subset M^0$, is tested against the alternative one $H_{1,M} : E\left(d_{(j),(j^*),q}\right) \neq 0$. The test at each iteration, for $\forall$ $M \subset M^0$, identifies the model that should be rejected under the $H_{0,M}$. If $\Psi_{q,(j)}$ denotes the value of the predicted squared error of model $j$ at quarter $q$, or $\Psi_{q,(j)} = \left(Y_q^{(0)} - \widehat{Y}_{q,(j)}^{(0)}\right)^2$, then $d_{(j),(j^*),q} = \Psi_{q,(j)} - \Psi_{q,(j^*)}$ is the evaluation differential for $\forall\ j, j^* \in M^0$. A high $p$ value provides evidence supporting the hypothesis that the model does belong to the model confidence set.

The most widely used tests for evaluating the statistical difference among competing forecasting models are the Diebold and Mariano (1995) test, the Equal Predictive Accuracy test of Clark and West (2007), the Reality Check for Data Snooping of White (2000), the Superior Predictive Ability of Hansen (2005) and the Model Confidence Set of Hansen et al. (2011). Each method has its pros and cons, and the Diebold and Mariano test is best suited for pairwise comparisons, while Model Confidence Set is more appropriate for simultaneously evaluating the forecasting performance of competing models, without predefining a benchmark model.

Tables 6, 7, 8, 9, 10, 11, 12, and 13 present the mean absolute percentage error and the root mean squared error for private consumption on goods and services (Table 6), government spending on public goods and services (Table 7), investment in business capital goods (Table 8), exports of goods (Table 9), exports of services (Table 10), imports of goods (Table 11), imports of services (Table 12), and changes in inventories (Table 13), respectively.

Indicatively, in Table 6, the MAPE loss function of nowcasting the consumption on goods and services when we have information published up to the 1st month of current quarter is 3.51% based on the Midas model and 9.70% based on the regression model. The regression model aggregates the data from the higher sampling frequency to the quarterly frequency as described in the robustness section. Overall, as we move from the 1st month of the current quarter to the 2nd month of next quarter, the error decreases for both model specifications (i.e., Midas and regression) and both loss functions (i.e., MAPE and RMSE). The AR(1) and RW are the first-order autoregressive and the random walk models, respectively, which used as naïve benchmarks. The naïve models are estimated for the 3rd month of the current quarter because of the 3-month publication lag of quarterly data. The naïve models have inferior performance in all the cases except in the case of the consumption on goods and services. Regarding consumption, the naïve models are beaten, in terms of nowcasting accuracy, by the Midas model only when the information for the 2nd month of next quarter is available.

The analysis in Tables 8, 9, 10, 11, and 12 reaches similar findings. In the vast majority of the cases, the Midas model outperforms the regression and the naïve models. Also, the Midas model has always better performance compared to the naïve models, even with the information available two months ago. Overall, the Midas models have better performance than the naïve models. The worst performance of the Midas model is in the case of consumption, where the information of the 2nd month of next quarter is required in order to beat the performance of the naïve models.

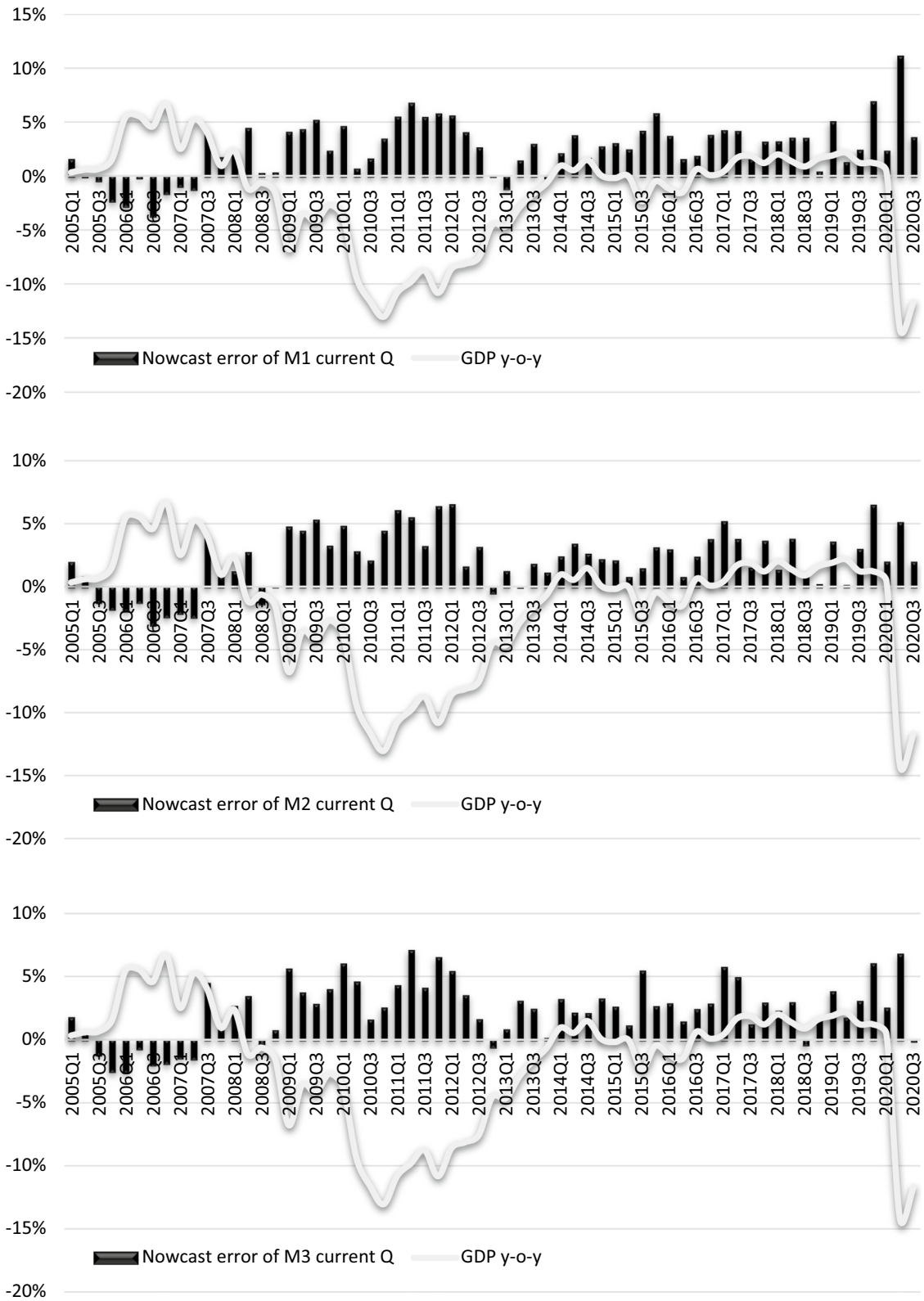



**Fig. 2** GDP *y–o-y* growth rate and the relative *y–o-y* nowcasting errors, $\left(\widehat{Y}_q^{(0)} - Y_q^{(0)}\right)/Y_{q-4}^{(0)}$ of the disaggregated Midas model. (i) The nowcasting error based on the data that are available up to 1st month of current quarter. (ii) The nowcasting error based on the data that are available up to 2nd month of the current quarter. (iii) The nowcasting error based on the data that are available up to 3rd month of the current quarter. (iv) The nowcasting error based on the data that are availableup to 1st month of the next quarter.(v) The nowcasting error based on the data that are available up to 2nd month of the next quarter

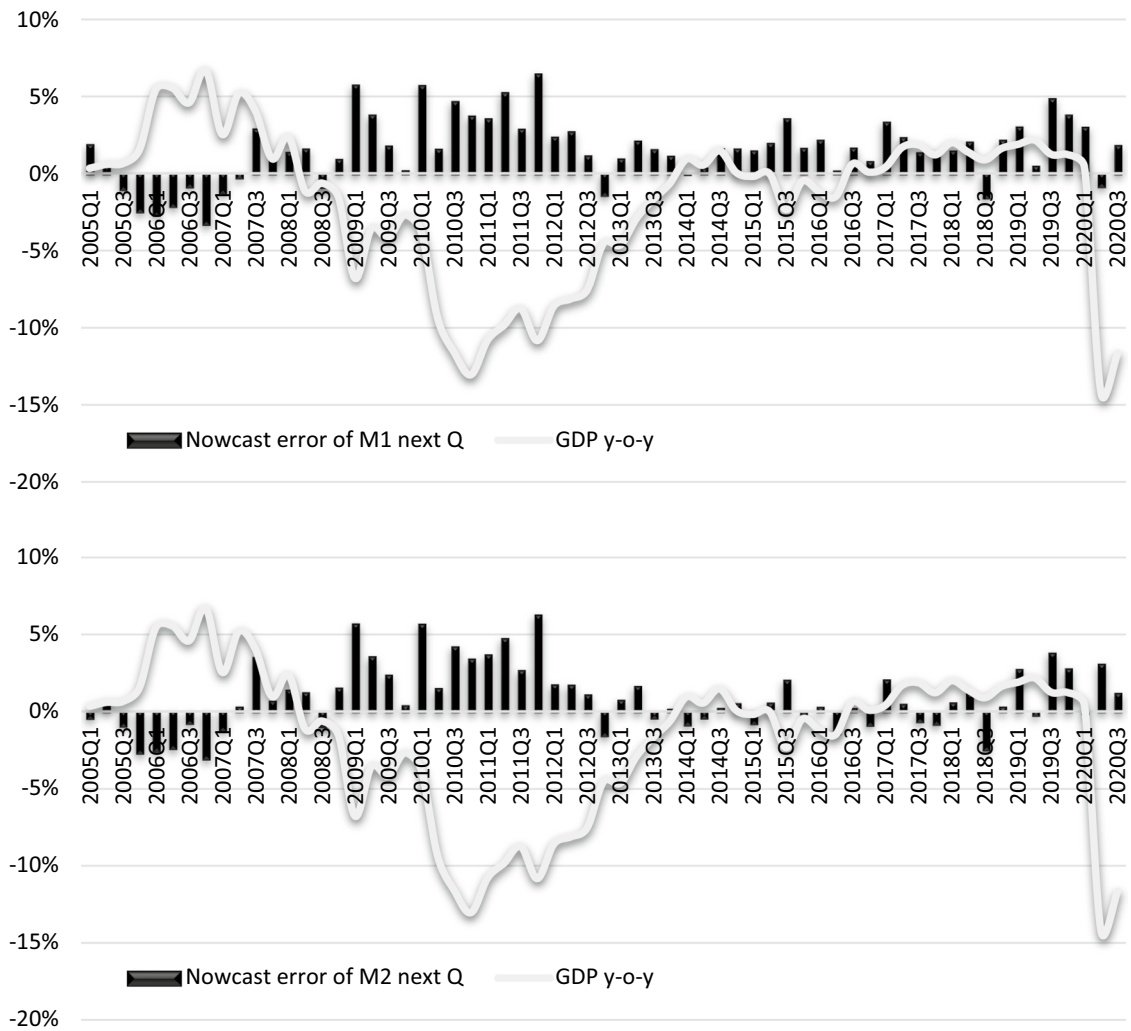



**Fig. 2** continued

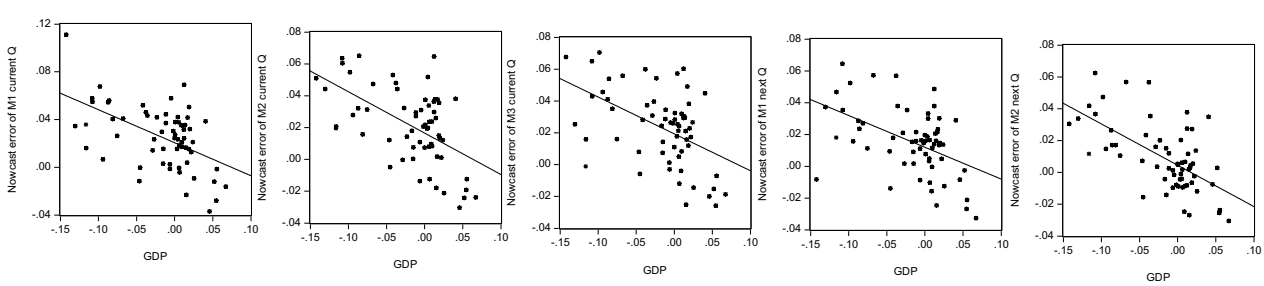



**Fig. 3** Scatterplots between *y–o-y* GDP growth and the *y–o-y* nowcasting errors of the disaggregated Midas model

As the nowcasting of government spending (Table 7) and the changes in inventories (Table 13) do not use a mixed-frequency framework, the nowcasting is conducted once the quarterly data are published.

As discussed in the robustness section, we run a series of models in order to investigate the usability of data sampled at higher frequencies. Table 14 presents the MAPE and RMSE loss functions for the best performing Midas and regression models which include additional variables. Indeed, only one monthly variable has satisfactory nowcasting accuracy; the financial conditions index, $x_m^{(6)}$. None of the additional models is able to perform better compared to those based on the daily dataset.

**Table 16** MAPE and RMSE loss functions for the disaggregated GDP models with error correction forecast

| Q | M | Disaggregated midas | | Disaggregated regression | | AR | | RW | | Regression model | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE | MAPE (%) | RMSE |
| Current quarter | M1 | 1.64 | 1269.1 | 1.82 | 1520.7 | | | | | | |
| | M2 | 1.66 | 1249.4 | 1.85 | 1500.7 | | | | | | |
| | M3 | 1.75 | 1282.6 | 1.84 | 1493.7 | 2.44 | 1938.4 | 2.54 | 1960.5 | | |
| Next quarter | M1 | 1.46 | 1138.3 | 1.97 | 1555.5 | | | | | | |
| | M2 | 1.56 | 1197.2 | 1.82 | 1414.4 | | | | | | |
| | M3 | | | | | 1.59 | 1335.1 | 1.55 | 1285.9 | 2.01 | 1563.8 |

Table 15 presents the nowcasting error when we estimate the GDP as a summation of the nowcasting of its components. For example, the MAPE loss function is 1.77% based on the Midas specifications when we take into consideration the data that are available up to the 2nd month of next quarter. When we use the regression model then the MAPE loss function becomes 2.14%. So, we reach at a very important finding. The Midas nowcasting based on the disaggregated data is by far better than the regression nowcasting. But a naïve model is able to provide a better nowcasting accuracy for the 3rd month of next quarter. The lower values of MAPE and RMSE for the naïve models compared to the Midas model are somehow in contradiction with the results presented for the nowcasting of GDP components. This is because, when we nowcast the GDP components separately, the Midas model has always a better nowcasting performance compared to the naïve models, except for the private consumption (where the AR(1) model performs slightly better). But if we aggregate the nowcasts of the components, then the GDP nowcasting has a higher MAPE compared to the MAPE of the naïve AR(1) model. The observed performance of the AR(1) model on the GDP nowcasting, leads us to model the forecast error in GDP nowcasting with an additional autocorrelated structure on the nowcasts of GDP components. The possible sources of the autocorrelated structure of the forecast error have been discussed in Sect. 3 (see the nowcasting error correction). Figure 2 plots the *y–o-y* growth rate of GDP against the *y–o-y* nowcasting error, which is defined as: $\frac{\left(\widehat{Y}_q^{(0)}-Y_{q-4}^{(0)}\right)}{Y_{q-4}^{(0)}}-\frac{\left(Y_q^{(0)}-Y_{q-4}^{(0)}\right)}{Y_{q-4}^{(0)}}\equiv\frac{\left(\widehat{Y}_q^{(0)}-Y_q^{(0)}\right)}{Y_{q-4}^{(0)}}$. Naturally, there is a positive relationship between the magnitude of the growth rate and the nowcasting error. Moreover, we observe that the majority of the nowcasting errors are positive (mainly in the estimation based on the data available in the 1st month of the current quarter). This positive bias of the nowcasting errors indicates an autocorrelated error structure, which justifies the use of the nowcasting error correction. The unconditional correlation between *y–o-y* GDP growth and the *y–o-y* nowcasting error ranges from -48% (for M1 of next quarter) up to -62% (for M2 of next quarter). Indicatively, Fig. 3 presents the scatterplots between *y–o-y* GDP growth and the *y–o-y* nowcasting errors, which confirms the autocorrelated error structure.

Table 16 presents the nowcasting error from the model that accounts for the error forecast correction. For example, the MAPE loss function is 1.75% based on the Midas specifications when we take into consideration the data that are available up to the 1st month of the current quarter. When we use the regression model, then the MAPE loss function is 1.84%, whereas the MAPE value of naïve AR(1) model is 2.44%. So, we conclude that the modeling of the nowcasting error structure, as proposed in Sect. 3, reduces the nowcasting error. Figure 4 plots the *y–o-y* growth rate of GDP and the relative *y–o-y* nowcasting errors of the disaggregated Midas model with the nowcasting error correction. We observe that the aforementioned positive bias of the nowcasting errors has decreased significantly, i.e., that is why the nowcasting accuracy has increased and is statistically significant.

This error reduction is statistically significant according to the $p$ values of the MCS test, which are presented in Table 17. A high $p$ value denotes that the model is included in the confidence set of the models with the lowest value in loss function, according to the MCS test. For example, if we define a 0.7 significance level, the disaggregated Midas with an error correction forecast becomes the only model to be included in the confidence set in all cases except for the nowcasting in the 2nd month of the next quarter. So, we conclude that *i*) this is a superior model for nowcasting the GDP at any nowcasting month and *ii*) only when

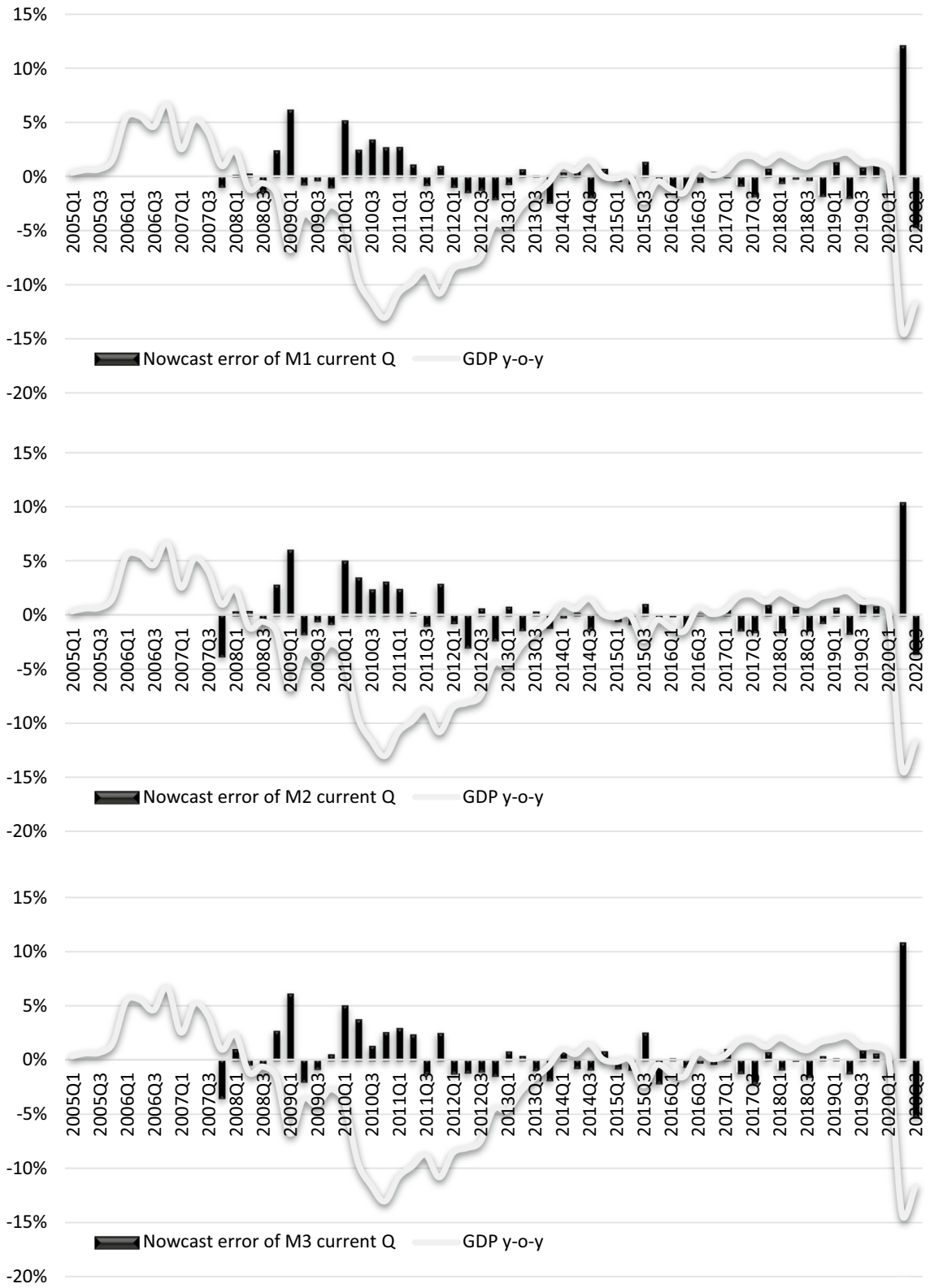



**Fig. 4** GDP *y–o-y* growth rate and the relative *y–o-y* nowcasting errors, $\left(\widehat{Y}_q^{(0)} - Y_q^{(0)}\right)/Y_{q-4}^{(0)}$ of the disaggregated Midas model with the nowcasting error correction. (i) The nowcasting error based on the data that are available up to the 1st month of the current quarter. (ii) The nowcasting error based on the data that are available up to the 2nd month of current quarter. (iii) The nowcasting error based on the data that are available up to the 3rd month of current quarter. (iv) The nowcasting error based on the data that are available up to the 1st month of next quarter. (v) The nowcasting error based on the data that are available up to the 2nd month of next quarter

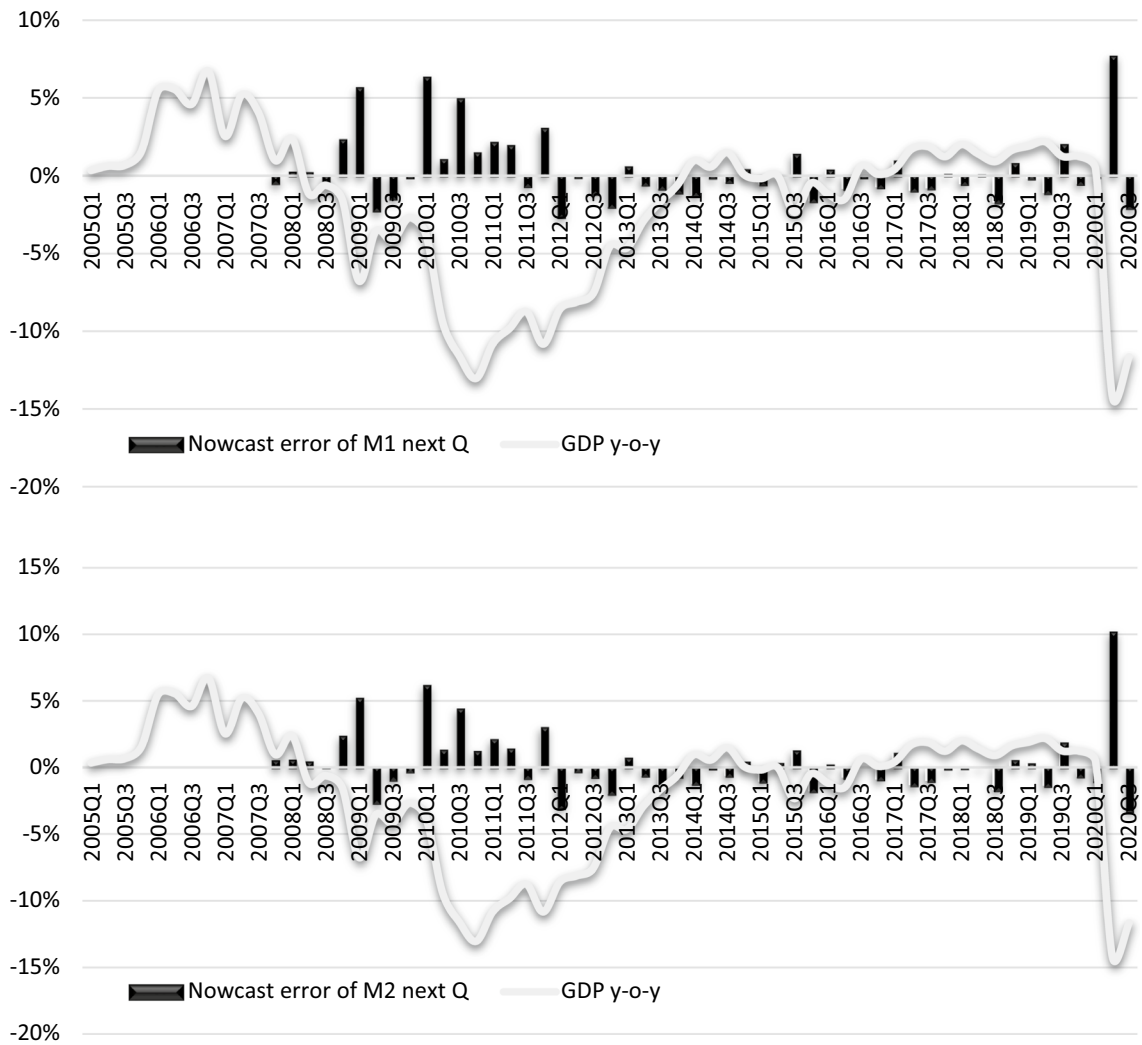



**Fig. 4** continued

the information of the current quarter is fully available does an AR(1) model provide equal nowcasting ability. Please keep in mind that we have presented the AR(1) model in the 2nd month of the next quarter but actually the AR(1) model can be estimated with a 3 month publication lag, in other words during the 3rd month of the next quarter.

## 9 Simulations

The disaggregation of GDP nowcasting has provided us with more accurate nowcasts of the components of GDP; in terms of MAPE and RMSE loss measures, but a naïve AR(1) model was able to provide nowcasts of equal forecasting accuracy for at least the 3rd month of next quarter.

In the paragraphs that follow we examine whether the inclusion of multiple sources of forecast errors is the key determinant of accuracy loss in the GDP nowcasting. As already mentioned, the computation of the GDP nowcasting requires the summation of multiple nowcast values. As GDP is the summation of its $k$ components; $Y_q^{(0)} = \left(\sum_{k=1}^{5} Y_q^{(k)} - \sum_{k=6}^{7} Y_q^{(k)} + Y_q^{(8)}\right)$, naturally, the nowcasting is computed as; $Y_{q\backslash q}^{(0)} = \left(\sum_{k=1}^{5} Y_{q\backslash q}^{(k)} - \sum_{k=6}^{7} Y_{q\backslash q}^{(k)} + Y_{q\backslash q}^{(8)}\right)$. As $Y_q^{(k)} = Y_{q\backslash q}^{(k)} + \varepsilon_{q\backslash q}^{(k)}$, the estimation of GDP nowcasting, $Y_{q\backslash q}^{(0)}$, hides diligently $k$ nowcasting errors, $\varepsilon_{q\backslash q}^{(k)}$. Thus, we run a series of simulations in order to unmask any possible impact of the multiple sources of forecasting errors.

**Table 17** *p* values from model the confidence test

| Q | M | Disaggregated midas | Disaggregated regression | Disaggregated midas with error correction forecast | Disaggregated midas with error correction forecast | AR(1) | RW | Regression model |
|---|---|---|---|---|---|---|---|---|
| *RMSE loss function* | | | | | | | | |
| Current quarter | M1 | 0.0004 | 0.0000 | **1.00** | 0.0256 | | | |
| | M2 | 0.0418 | 0.0000 | **1.00** | 0.0592 | | | |
| | M3 | 0.0105 | 0.0000 | **1.00** | 0.0450 | 0.0105 | 0.0076 | |
| Next quarter | M1 | 0.0732 | 0.0001 | **1.00** | 0.0732 | | | |
| | M2 | 0.0029 | 0.0001 | **1.00** | 0.0988 | 0.2905 | 0.3725 | 0.000 |
| *MAPE loss function* | | | | | | | | |
| Current quarter | M1 | 0.0000 | 0.0000 | **1.00** | 0.2246 | | | |
| | M2 | 0.0060 | 0.0000 | **1.00** | 0.1825 | | | |
| | M3 | 0.0001 | 0.0000 | **1.00** | 0.4604 | 0.0001 | 0.0001 | |
| Next quarter | M1 | 0.0013 | 0.0000 | **1.00** | 0.0107 | | | |
| | M2 | 0.0001 | 0.0000 | **0.9860** | 0.4819 | **1.00** | **0.8415** | 0000 |

A high *p* value denotes that the model is included in the confidence set of the models with the lowest value in loss function, according to the MCS test. For example, if we define a 0.7 significance level, the bold-faced models are included in the confidence set

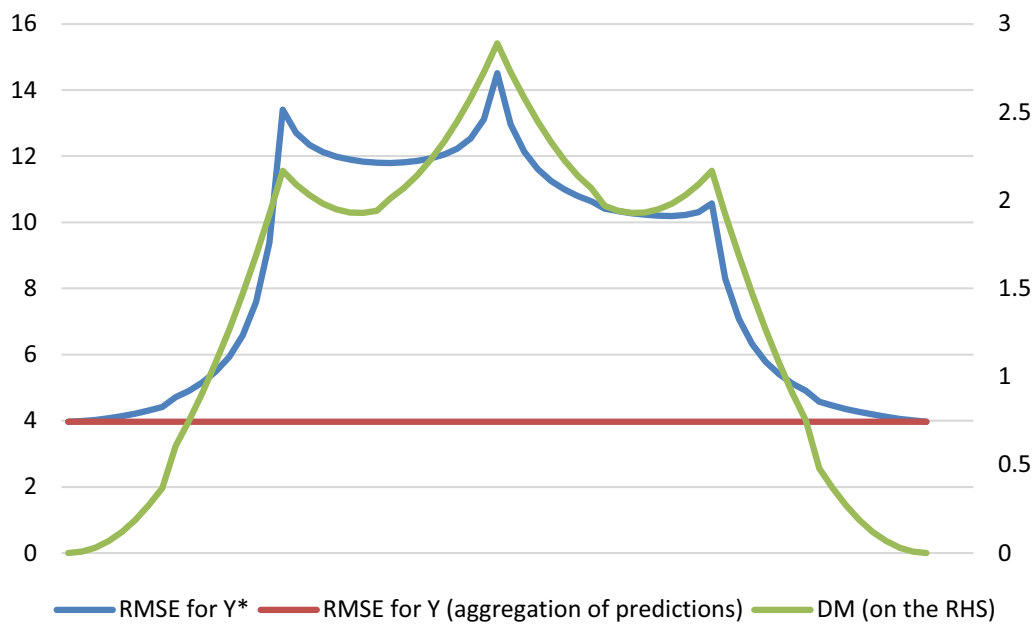



**Fig. 5** Dispersion measure, *DM*, along with the RMSE loss function for $Y^{(0)}_{q+1\backslash q}$ and $Y^{*(0)}_{q+1\backslash q}$. The RMSE for the one-step-ahead forecasts of $Y^{(0)}_{q}$ from an estimated AR(1) model: $Y^{*(0)}_{q+1\backslash q}$ is $RMSE = \sqrt{10.000^{-1}\sum_{q=1}^{10.000}\left(Y^{(0)}_{q+1} - Y^{*(0)}_{q+1\backslash q}\right)^2}$. The RMSE for the one-step-ahead forecasts of $Y^{(0)}_{q}$ from the aggregation of the forecasts: $Y^{(0)}_{q+1\backslash q}$ is $RMSE = \sqrt{10.000^{-1}\sum_{q=1}^{10.000}\left(Y^{(0)}_{q+1} - \sum_{k=1}^{4} Y^{(k)}_{q+1\backslash q}\right)^2}$. The $DM = \sum_{k=1}^{4}\left(\beta_1^{(k)} - \overline{\beta_1}\right)^2$, for $\overline{\beta_1} = \sum_{k=1}^{4}\beta_1^{(k)}/4$. The DM is presented on the right-hand-side axis

### 9.1 Autoregressive framework

We assume an aggregated series $Y^{(0)}_q = \left(\sum_{k=1}^{4} Y^{(k)}_q\right)$, where the *q-o-q* growth rate of each $Y^{(k)}_q$ follows an AR(1) process:

$$\begin{aligned} y^{(k)}_q &= log\left(Y^{(k)}_q / Y^{(k)}_{q-1}\right), \\ y^{(k)}_q &= \beta^{(k)}_0 + \beta^{(k)}_1\left(y^{(k)}_{q-1} - \beta^{(k)}_0\right) + \varepsilon^{(k)}_q, \\ \varepsilon^{(k)}_q &\sim N\left(0, \sigma^{2(k)}_q\right). \end{aligned} \tag{40}$$

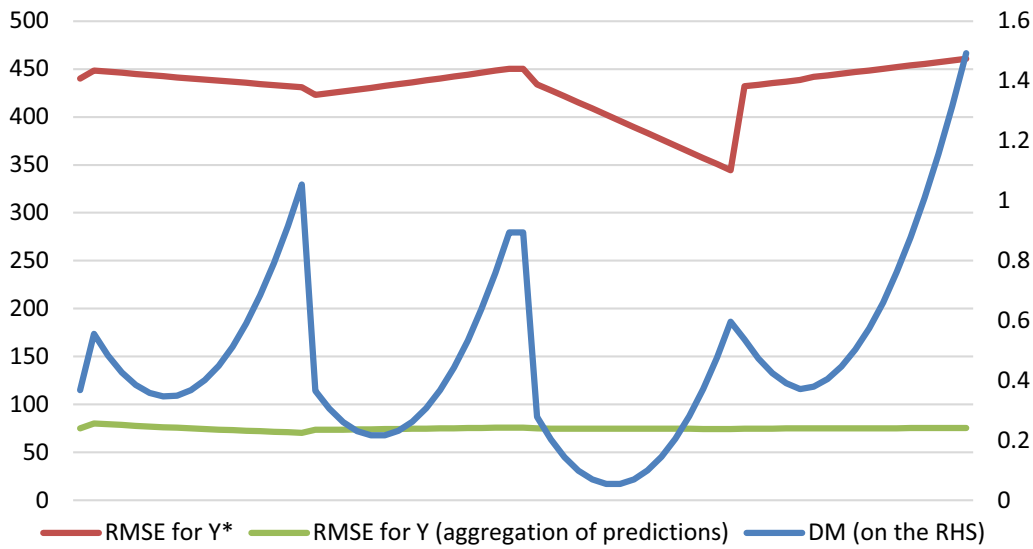



**Fig. 6** RMSE loss function for $Y^{(0)}_{q+1\backslash q}$ and $Y^{*(0)}_{q+1\backslash q}$ along with the dispersion measure, *DM*. The RMSE for the one-step-ahead forecasts of $Y^{(0)}_{q}$ from the estimated regression model: $Y^{*(0)}_{q+1\backslash q}$ is $RMSE = \sqrt{10.000^{-1}\sum_{q=1}^{10.000}\left(Y^{(0)}_{q+1} - Y^{*(0)}_{q+1\backslash q}\right)^2}$. The RMSE for the one-step-ahead forecasts of $Y^{(0)}_{q}$ from the aggregation of the forecasts: $Y^{(0)}_{q+1\backslash q}$ is $RMSE = \sqrt{10.000^{-1}\sum_{q=1}^{10.000}\left(Y^{(0)}_{q+1} - \sum_{k=1}^{4} Y^{(k)}_{q+1\backslash q}\right)^2}$. The $DM = \sum_{k=1}^{4}\left(\beta_1^{(k)} - \overline{\beta_1}\right)^2$, for $\overline{\beta_1} = \sum_{k=1}^{4}\beta_1^{(k)}/4$. The DM is presented on the right-hand-side axis

Then, we compute the one-step-ahead forecasts of $Y^{(k)}_{q}$ as $Y^{(k)}_{q+1\backslash q}$, for $k = 1, .., 4$ as well as the one-step-ahead forecasts of $Y^{(0)}_{q}$ as $Y^{(0)}_{q+1\backslash q} = \sum_{k=1}^{4} Y^{(k)}_{q+1\backslash q}$. Moreover, we assume for the simulated process $Y^{(0)}_{q}$ that it can be estimated as an AR(1) process, thus, we compute one-step-ahead forecasts of $Y^{(0)}_{q}$ from an estimated AR(1) model: $Y^{*(0)}_{q+1\backslash q}$.

By design, the true data generated process of $Y^{(0)}_{q}$ is the aggregation of the components whose *q-o-q* growth rate has a first-order autoregressive structure. So, in terms of the statistical evaluation of forecasting accuracy, the $Y^{(0)}_{q+1\backslash q}$ forecasts must be more accurate compared to $Y^{*(0)}_{q+1\backslash q}$ forecasts according to the classical loss functions, despite the fact that we have imposed $k$ forecasting errors, $\varepsilon^{(k)}_{q+1\backslash q}$.

The simulations have been conducted for various values of parameters $\beta_0^{(k)}, \beta_1^{(k)}$ and the magnitude of the error term, $\sigma_q^{2(k)}$. Indicatively, for $\beta_0^{(k)} = 0.1$ and $-0.8 \leq \beta_1^{(k)} \leq 0.9$, various combinations of the four AR(1) models of Eq. (39) have been simulated. For illustration purposes, we have constructed a measure that represents the dispersion among the values of parameters.[9] The dispersion measure is computed as:

$$DM = \sum_{k=1}^{4}\left(\beta_1^{(k)} - \overline{\beta_1}\right)^2, \tag{41}$$

where $\overline{\beta_1} = \sum_{k=1}^{4}\beta_1^{(k)}/4$. Figure 5 presents the dispersion measure, *DM*, along with the RMSE loss function for $Y^{(0)}_{q+1\backslash q}$ and $Y^{*(0)}_{q+1\backslash q}$. The value of the RMSE loss function for the aggregated forecast $\sum_{k=1}^{4} Y^{(k)}_{q+1\backslash q}$ is stable across the various values of the dispersion measure. On the other hand, the values of $RMSE = \sqrt{10.000^{-1}\sum_{q=1}^{10.000}\left(Y^{(0)}_{q+1} - Y^{*(0)}_{q+1\backslash q}\right)^2}$ are highly related to the values of the dispersion measure. Therefore, we reach the finding that the aggregation of the predictions provides more accurate one-step-ahead predictions despite the inclusion of multiple sources of forecast errors. Moreover, when we ignore the disaggregation (and we compute the $Y^{*(0)}_{q+1\backslash q}$), the loss of forecasting accuracy increases proportionally to the dispersion among the values of the parameters.

### 9.2 Regression framework

For robustness, we create another simulated framework assuming an aggregated series $Y^{(0)}_{q} = \left(\sum_{k=1}^{4} Y^{(k)}_{q}\right)$, where the *q-o-q* growth rate of each $Y^{(k)}_{q}$ follows a regression model:

[9] The simulations and their outputs are available upon request.

$$
\begin{aligned}
y_q^{(k)} &= log\left(Y_q^{(k)}/Y_{q-1}^{(k)}\right),\\
y_q^{(k)} &= \beta_0^{(k)} + \beta_1^{(k)}(1-L)x_q^{(k)} + \varepsilon_q^{(k)},\\
\varepsilon_q^{(k)} &\sim N\left(0, \sigma_q^{2(k)}\right).
\end{aligned} \tag{42}
$$

The initial values of the coefficients in the simulated regressions have been computed from similar regressions based on the actual dataset. Thus, we have assumed as $y_q^{(1)}$ the private consumption of goods and services, $x_q^{(1)}$ the retail trade volume index, $y_q^{(2)}$ the investment in business capital goods, $x_q^{(2)}$ the Athens stock exchange main general index, $y_q^{(3)}$ the exports of goods, $x_q^{(3)} = \sum_{k=1}^{3} \widetilde{x}_q^{(k)} + 0.2\widetilde{x}_q^{(4)}$ (for export of fuels $\widetilde{x}_q^{(1)}$, export of vessels $\widetilde{x}_q^{(2)}$, other exports $\widetilde{x}_q^{(3)}$ and travel receipts $\widetilde{x}_q^{(4)}$) and $y_q^{(4)}$ the imports of goods, $x_q^{(4)} = \sum_{k=5}^{7} \widetilde{x}_q^{(k)}$ (for $k = 5, 6, 7$ we denote the importation of fuels, importation of vessels and importation of other goods, respectively).

Then, we compute the one-step-ahead forecasts of $Y_q^{(k)}$ as $Y_{q+1\backslash q}^{(k)}$, for $k = 1, .., 4$ as well as the one-step-ahead forecasts of $Y_q^{(0)}$ as $Y_{q+1\backslash q}^{(0)} = \sum_{k=1}^{4} Y_{q+1\backslash q}^{(k)}$. Finally, we assume for the simulated process $Y_q^{(0)}$ that it can be estimated via a regression model that incorporates all the explanatory variables, i.e., $y_q^{(k)} = \beta_0^{(k)} + \sum_{i=1}^{4}\left(\beta_i^{(k)}(1-L)x_{i,q}^{(k)}\right) + \varepsilon_q^{(k)}$. So, we define the one-step-ahead forecasts of $Y_q^{(0)}$ from this regression as $Y_{q+1\backslash q}^{*(0)}$.

By design, the true data generating process of $Y_q^{(0)}$ is the aggregation of the components, or $Y_q^{(0)} = \left(\sum_{k=1}^{4} Y_q^{(k)}\right)$. So, in terms of statistical evaluation of forecasting accuracy, the $Y_{q+1\backslash q}^{(0)}$ forecasts must be more accurate compared to $Y_{q+1\backslash q}^{*(0)}$ forecasts according to the classical loss functions, despite the fact that we have imposed $k$ forecasting errors, $\varepsilon_{q+1\backslash q}^{(k)}$.

The simulations have been conducted for various values of parameters $\beta_0^{(k)}, \beta_1^{(k)}$ and of the magnitude of the error term,$\sigma_q^{2(k)}$, around the initially estimated values; $\beta_0^{(1)} = 0.001$, $\beta_0^{(2)} = -0.01$, $\beta_0^{(3)} = 0.002$, $\beta_0^{(4)} = -0.0007$, $0.1 \le \beta_1^{(1)} \le 1.8$, $0.095 \le \beta_1^{(2)} \le 1.595$, $0.02 \le \beta_1^{(3)} \le 1.42$ and$0.01 \le \beta_1^{(4)} \le 1.51$. Figure 6 presents the RMSE loss function for $Y_{q+1\backslash q}^{(0)}$ and $Y_{q+1\backslash q}^{*(0)}$ and the dispersion measure as well. The values of the RMSE loss function for the aggregated forecast $\sum_{k=1}^{4} Y_{q+1\backslash q}^{(k)}$ are stable across the various combinations of the parameter's values. On the other hand, the values of $RMSE = \sqrt{10.000^{-1} \sum_{q=1}^{10.000}\left(Y_{q+1}^{(0)} - Y_{q+1\backslash q}^{*(0)}\right)^2}$ are much higher (almost 6 times higher). Naturally, the dispersion measure is not related to the values of the RMSE loss function, because of the heterogeneity of the simulated framework in Eq. (41). However, as in the previous simulated framework, we reach a similar conclusion that the aggregation of the predictions provides more accurate one-step-ahead predictions, despite the inclusion of multiple sources of forecast errors.

## 10 Conclusions and further research

Literature has often highlighted that sophisticated models can rarely outperform the forecasting ability of a naive model; see D'Agostino et al. (2006) and Campbell (2007). Schumacher and Breitungth (2008) note that a sophisticated factor model is able to provide only moderate forecast performance in predicting German GDP, but as Schumacher (2010) notice, the preselection of international indicators may contain additional information in forecasting GDP. So, in contrast our paper contributes to the literature by providing both empirical and simulated evidence that more accurate nowcasting estimations of GDP require the use of a disaggregated multilayer mixed-frequency framework.

Indeed, the nowcasting ability of the AR(1) naive model is not better only if we define a sophisticated model framework. The proposed model framework requires the preselection of the explanatory variables. The explanatory variables must be related to the components of GDP based on a multilayer mixed-frequency framework, and we observe that even the daily available financial data are able to reduce the nowcasting error. So, we realize that the disaggregation into components reduces the forecasting error despite the inclusion of multiple sources of forecast errors.

Of course, there is, still, much to be done that could possibly improve the nowcasting accuracy. The induction of a supervised algorithm, like the Lasso model selection process, can probably identify the explanatory variables that are strongly associated with the nowcasted variable. One further avenue that could improve the nowcasting accuracy is to find a way of exploiting the cross-sectional information to get more accurate estimates or models.

Concluding, the estimation of the D-model[10] with the same structure among data and the same equations across layers but with data coming from another country will be like putting data into a black box. The construction of such framework requires the knowledge of data availability, their quality, and their interconnectedness. Thus, before the replication of the proposed model framework, the careful collection of the data and the construction of the appropriate connections among economic variables and across time is a necessity.

[10] This is the case for similar model framework having been proposed in the literature, as well.

## Appendix A

| Article | Country | Period | Technique | Explanatory variables | Results |
|---|---|---|---|---|---|
| Angelini et al (2011), Econometrics Journal | Euro Area | 1999Q1 2007Q2 | Pools of bridge equations and the 'bridging with factors' approach proposed by Giannone et al. (2008) for the backcast, nowcast and short-term forecast of euro area quarterly GDP growth | 85 macroeconomic time series | The factor model improves upon the pool of bridge equations |
| Angelini et al. (2008), ECB | Euro Area | 1993Q1 2006Q2 | A dynamic factor model based on Doz et al. (2011), which differs from other approaches (e.g., Stock & Watson, 2002a; Forni et al., 2000) | 85 macroeconomic time series | For GDP and a number of components, the factor model forecasts beat the forecasts from alternative model such as quarterly models and bridge equations |
| Antipa et al., (2012), Journal of Policy Modeling | Germany | 1993Q1 2007Q4 | Comparing the BMs and DFMs with a rolling forecast exercise in order to assess the forecasting performance | | Forecast errors of the BMs are smaller than those of the DFMs |
| Artis et al., (2005), Journal of Forecasting | UK | 1970Q1 1998Q3 | Dynamic factor model | 81 macroeconomic time series | 6 factors explain about 50% of the variability of 81 variables, the factors are related to groups of key variables, such as interest rates, price series, monetary aggregates, labor market variables and exchange rates |
| Baffigi et al., (2004), International Journal of Forecasting | Euro Area, Germany, France, Italy | 1980Q1 2002Q2 | Bridge model against three types of benchmark models: univariate ARIMA, multivariate VAR and structural models | Macroeconomic indicators for each country | BM performance is always better than benchmark models, provided that at least some indicators are available over the forecasting horizon |
| Bańbura and Rünstler (2011), International Journal of Forecasting | Euro Area | 1993Q1 1996Q2 | Dynamic factor model | 32 real activity series, 22 survey series, 22 financial series | Both forecast weights and forecast precision measures attribute an important role to survey data, whereas real activity data obtain rather low weights, apart perhaps from the backcasts. Financial data provide complementary information to both real activity and survey data for nowcasts and one-quarter ahead forecasts of GDP |
| Banerjee and Marcellino (2006), International Journal of Forecasting | USA | 1975Q1 2001Q4 | Dynamic factor model | 64 inflation indicators, 74 GDP growth indicators | All methods are systematically beaten by single indicator models both for inflation and GDP growth |

| Article | Country | Period | Technique | Explanatory variables | Results |
|---|---|---|---|---|---|
| Barhoumi et al. (2008), ECB | Selected European countries and the Euro Area | 1991m1 2006m6 | Bridge model and dynamic factor model | More than one hundred series for each country | For the euro-area countries models that exploit timely monthly releases fare better than quarterly models. Factor models, which exploit a large number of releases, generally do better than averages of bridge equations |
| Bessec (2012), Banque de France | France | 1990Q1 2010Q4 | Dynamic factor model | French GDP growth and 96 predictors. (surveys, indicators of real activity, monetary and financial variables) | Financial variables and survey variables are predominant at longer horizons, while the weight of real indicators increases at shorter ones. A pseudo real-time evaluation over the last decade shows again relative to factor models without preselection or with preselection made on the full dataset at least for large horizons |
| Boivin and Ng (2006), Journal of Econometrics | USA | 1971Q1 1997Q4 | A factor model, which focuses on the finite sample properties of the PC estimator in the presence of cross-section correlation in the idiosyncratic errors, which is a pervasive feature of the data | 147 series as in Stock and Watson (2002a) | In a real-time forecasting exercise, factors extracted from as few as 40 pre-screened series often yield satisfactory or even better results than using all 147 series |
| D'Agostino et al. (2006), ECB | USA | 1959m1 2003m12 | Random walk model, univariate forecasts, factor augmented forecast, in which the univariate models are augmented with common factors extracted from the whole panel of series. Pooling of bivariate forecasts: for each variable the forecast is defined as the average of 130forecasts obtained by augmenting the model with each of theremaining 130 variables in the data set | 131 monthly time series | The ability to predict several measures of inflation and real activity has declined remarkably, relative to naive forecasts, since the mid-1980s. The informational advantage of the Fed and professional forecasters is limited to the 1970s and the beginning of the 1980s |
| D'Agostino et al. (2012), OECD Journal of Business Cycle Measurement and Analysis | Irish | 1980Q1 1996Q4 | Dynamic factor model that produces nowcasts and backcasts of Irish quarterly GDP | Panel dataset of 35 indicators | The mean squared forecast errors for both the nowcasts and the backcasts based on DF model are considerably smaller than those of the benchmark model (average growth rate model) |
| Dahl et al. (2009), International Journal of Forecasting | Denmark | | Cyclical components factor model | 172 monthly and 74 quarterly series | Cyclical components factor model improves the forecast accuracy substantially relative to the regular diffusion index model for four Danish macroeconomic variables |

| Article | Country | Period | Technique | Explanatory variables | Results |
|---|---|---|---|---|---|
| Stock and Watson (2005a), NBER | USA | 1959m1 2003m12 | Static and dynamic factor models for VAR analysis | Monthly observations on 132 US macro time series | A large number of dynamic factors accounts for the movements in these data. Evidence against the VAR restrictions implied by the exact DFM. The data are well described by an approximate factor model but not an exact factor model. The structural FAVAR permits examination of overidentifying restrictions and diagnosis of modeling problems |
| Stock and Watson (2005b), Working Paper | USA | 1960m1 2003m12 | This paper compares the empirical accuracy of forecast combination, model selection, dynamic factor model forecasts, Bayesian model averaging, empirical Bayes methods | 131 monthly macro time series | The FAAR models and the principal component BMA models with small values of g put weight on a few of the principal components, resulting in more accurate forecasts |
| Favero et al. (2008), Journal of Applied Econometrics | USA, Euro Area; DE/IT/FR/ES | 1959m1 1998m12 (USA) 1982m1 1997m8 (Euro Area) | Static and dynamic factor models | 146 (USA) and 105 (DE/IT/FR/ES) time series | Factor models produce useful instruments for the estimation of forward-looking economic models. The DFM is more parsimonious than the static model, but the overall performance is similar |
| De Mol et al. (2008), Journal of Econometrics | USA | 1959m1 2003m12 | Bayesian shrinkage as an alternative to PCA | 131 macroeconomic variables (real and nominal variables, asset prices, surveys) | The forecasts provide a valid alternative to the PCA and are correlated with those obtained from the PCA. In addition, from an economic point of view, the results are not more interpretable than those of the PCA |
| Doz et al. (2011), Journal of Econometrics | Euro Area | 1993Q1 2006Q2 | The parameters of a DFM are estimated using OLS on PC and given the estimates the factors are estimated using a Kalman smoother | Simulation study for the DGP and 85 macroeconomic time series | This approach improves the estimation of the factors for small values of n |
| Giannone et al. (2008), Journal of Monetary Economics | USA | 1982Q1 2005Q1 | DFM using a two-step estimator for the factors: PCA followed by Kalman smoother | 200 macroeconomic time series | Precision of the nowcast increases monotonically as new data become available |
| Heij et al. (2008), International Journal of Forecasting | USA | 1959m3 1998m12 | Matched principal components regression (MPCR) | 146 macroeconomic predictor variables, dataset of Stock and Watson (2002a) | A modified PCM is proposed in order to improve the forecasting ability compared to the PCR. The MPCR maximizes the variance of the predictors during the estimation interval |

| Article | Country | Period | Technique | Explanatory variables | Results |
|---|---|---|---|---|---|
| Heij et al. (2011), International Journal of Forecasting | USA | 1959m1 2009m5 | An improved method for the construction of principal components in macroeconomic forecasting | 10 leading indicators and 4 coincident indicators | The proposed modification leads, on average, to more accurate forecasts than previously used principal component regression methods |
| Kuzin et al. (2011), International Journal of Forecasting | Euro Area | 1992Q1 2008Q1, 1992m1 2008m6 | Comparison between mixed-data sampling (Midas) and mixed-frequency VAR (MF-VAR) approaches | 20 monthly indicators from four main categories: industrial production, surveys, interest rates, exchange rates and money stocks, raw material prices and car registrations | Forecasting performance does not result in a clear winner. For short-term horizons AR-Midas performs better than Midas and MF-VAR, whereas for longer-term horizons MF-VAR outperforms the other two |
| Schumacher (2010), Economic Letters | Germany | 1980Q3 2004Q4 | Large factor model—factors areestimated by PC—targeted predictors | 531 variables: 123 quarterly indicators and data covering EA and G7 countries | International data improve forecasts only in the case that variables are preselected by LARS-EN (least-angle regression with elastic net) |
| Schumacher and Breitung (2008), International Journal of Forecasting | Germany | 1991Q2 2005Q1, 1991m4 2004m12 | Factors are estimated applying an EM logarithm combined with a PC estimator | 52 time series: 39 monthly series and 13 quarterly series | The mixed-frequency factor model performs slightly better in comparison with the balanced data factor models. The difference is more pronounced once the real-time factor model is compared to simple benchmark models |
| den Reijer (2005), De Nederlandsche Bank | The Netherlands | 1980Q1 2002Q4, GDP growth forecasts up to 8 quarters ahead | Large-scale factor model based on the static approach of Stock and Watson (2002a) and the dynamic approach of Forni et al. (2000) | 270 series underlying the Central Bank's macroeconomic structural model supplemented with leading indicator variables. Subset of 170 series | Full data sample: the factor models do not outperform the AR benchmark model. Data subsample: The forecasting performance of the factor models improves. The dynamic factor model systematically outperforms the AR benchmark model |
| Stakénas (2012), Lietuvos Bankas | Lithuania | 1996Q1 2011Q3, 2000Q2 2011Q1 for forecast evaluation | Principal components, generalized principal components and the state space model | 52 monthly indicators: survey, industry production, trade, price, financial variables, etc. | Factor models perform better than naïve benchmark models. The small-scale factor model (5 variables) outperforms the large-scale model comprising the whole dataset |
| Peña and Poncela (2004), Journal of Econometrics | European OECD countries; Belgium, France, Italy, the Netherlands, Spain | Annual real GNP 1949–1997. After 1981 forecasts were generated | A dynamic factor model with a common trend and a common AR(1) stationary factor | | The factor model provides substantial improvement in forecasts with respect to both univariate and shrinkage univariate forecasts |

| Article | Country | Period | Technique | Explanatory variables | Results |
|---|---|---|---|---|---|
| Iacoviello (2001), IMF | Italy | 1985Q22000Q2 forecasts from 1996Q2 | Indicator approach: bridge model (short-term forecasting) Econometric approach: Bayesian VAR (longer-term forecasting) | Bridge model: ind. Prod. index, coincident survey indicator, leading survey indicator BVAR model: real household cons., t-bill rate, coincident survey ind., exchange rate, cpi, German gdp | Based on forecasting performance, both models are useful tools |
| Stock and Watson (2002b), Journal of the American Statistical Association | USA | 1959m1 1998m12, 12-month ahead forecasts 1970m1 1997m12 | Principal components, factor model, univariate AR, VAR, leading indicator model, AR-augmented PCM | 149 monthly macroeconomic variables | The factor models offer substantial improvement stemming mainly from the first two or three factors. The leading indicator and the VAR models perform slightly better than the univariate AR |

## Appendix B

We are going through the 3rd month of the current quarter. Keeping in mind that the balance of payments is published with a lag of 2 months, or, $\boldsymbol{I}_m = \left\{x_{m-2}, pmi_{m-1}^{(exp)}\right\}$, the real-time nowcast of $x_m$ for the 3rd month of the quarter equals:

$$x_{m\backslash m-2} = e^{log\left(x_{m-1\backslash m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{43}$$

F the 1st month of the quarter, the $x_m$ is estimated as:

$$x_{m-1\backslash m-2} = e^{log\left(x_{m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-3}^{(exp)}\right)}. \tag{47}$$

When we are in the 1st month of the quarter, the real-time nowcast of $x_m$ for the 3rd month of the quarter is estimated by the first-order autoregressive model for $(1-L)log(x_m)$, as the $pmi^{(exp)}$ for the 3rd month has not been published. Thus:

$$x_{m\backslash m-4} = e^{(1+\gamma_1(m-4))\log\left(x_{m-1\backslash m-4}\right)-\gamma_1^{(m-4)}\log\left(x_{m-2\backslash m-4}\right)+\gamma_0^{(m-4)}\left(1-\gamma_1^{(m-4)}\right)}. \tag{48}$$

The real-time nowcast of $x_m$ for the 2nd month of the current quarter is:

$$x_{m-1\backslash m-2} = e^{log\left(x_{m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-3}^{(exp)}\right)}. \tag{44}$$

And for the 1st month of the quarter, the $x_m$ has been published already.

When we are in the 2nd month of the quarter, the real-time nowcast of $x_m$ for the 3rd month of the quarter equals to:

$$x_{m\backslash m-3} = e^{log\left(x_{m-1\backslash m-3}\right)+\gamma_0^{(m-3)}+\gamma_1^{(m-3)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{45}$$

The real-time nowcast of $x_m$. for the 2nd month of the quarter is computed as:

$$x_{m\backslash m-2} = e^{log\left(x_{m-1\backslash m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{46}$$

For the 2nd month of the quarter, the $x_m$ is estimated, based on $\boldsymbol{I}_m = \left\{x_{m-2}, pmi_{m-2}^{(exp)}\right\}$, as:

$$x_{m+1\backslash m-2} = e^{log\left(x_{m\backslash m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-1}^{(exp)}\right)}. \tag{49}$$

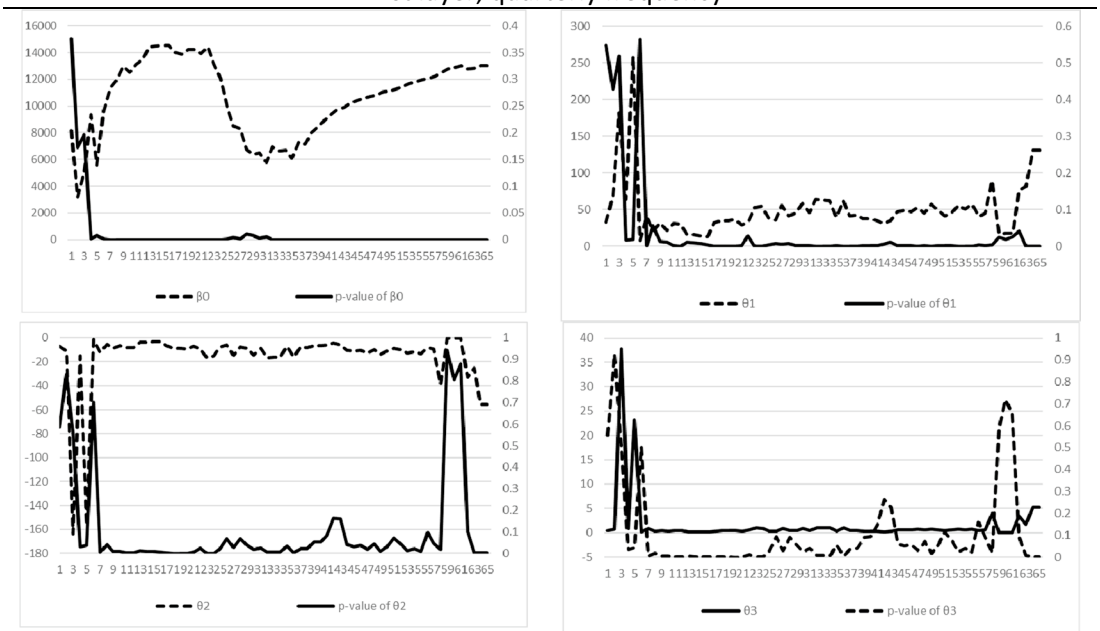
For the 1st month of the quarter, the $x_m$ is estimated as:

$$x_{m\backslash m-2} = e^{log\left(x_{m-1\backslash m-2}\right)+\gamma_0^{(m-2)}+\gamma_1^{(m-2)}(1-L)log\left(pmi_{m-2}^{(exp)}\right)}. \tag{50}$$

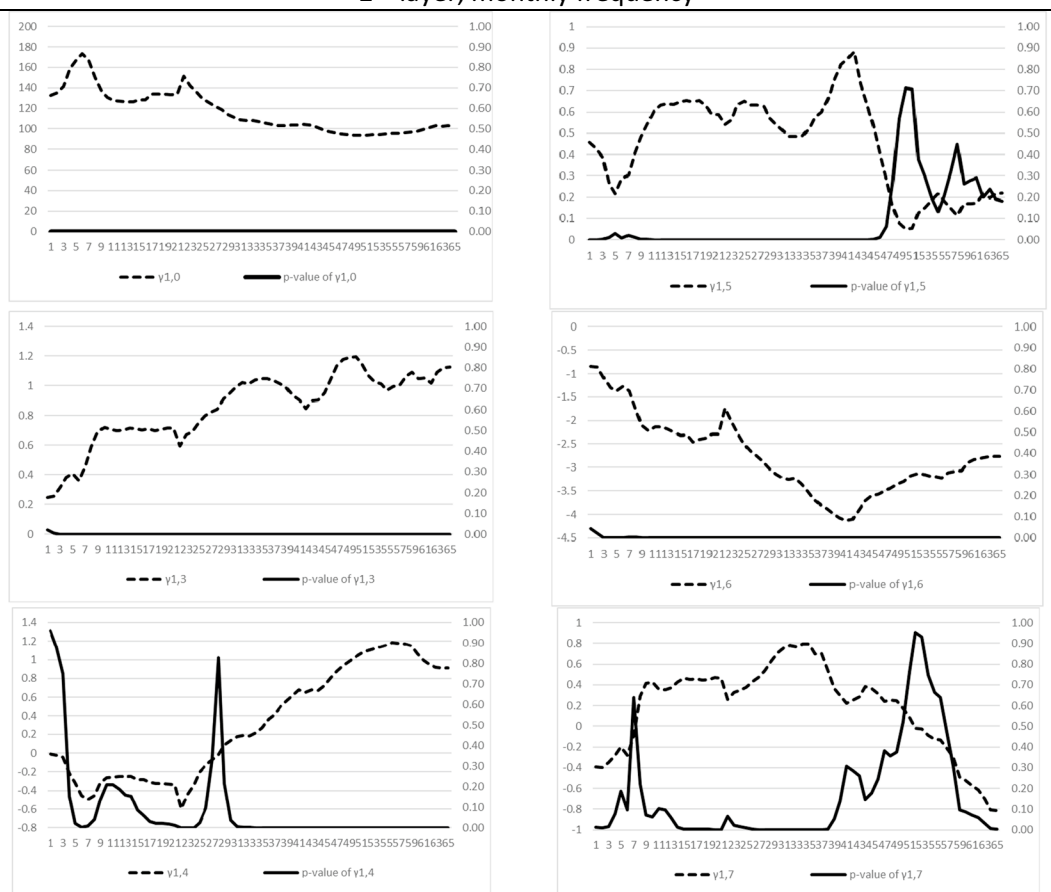
## Appendix C

The estimated coefficients (on the LHS) and their $p$ values (on the RHS), for the private consumption based on Eqs. 5 and 6. On the quarterly (monthly) frequency, the estimated parameters refer to current quarter (2nd month of current quarter), based on information available on the 3rd month of current quartet.

1st layer, quarterly frequency




2nd layer, monthly frequency



**Acknowledgements**
I would like to thank the editor Professor Rafael Lalive and the two anonymous reviewers for their helpful comments and suggestions. Their insights have improved the quality of the paper substantially. I would also like to thank Eleftheria Kafousaki, Thanos Petralias, Stelios Panagiotou, Dimitris Malliaropoulos, and Zacharias Bragoudakis for their useful comments. The views expressed in this paper are those of the author and not necessarily those of either the Bank of Greece or the Eurosystem.

**Author contributions**
S.D. is a Full Professor of Statistics and Econometrics at the Department of Economics and Regional Development at the Panteion University of Social and Political Sciences. He holds a Ph.D. in Statistics from Athens University of Economics and Business. His research interests revolve mainly around the financial and economic forecasting and energy economics. His research has received multiple research funding, that is, from FP7, Horizon 2020, and European Commission. He has also served as a consultant for US Energy Information Administration, Economic Chamber of Greece, Bank of Greece, and so on. Author read and approved the final manuscript.

**Funding**
The research has not been funded.

**Availability of data and materials**
The data are available upon request.

## Declarations

**Competing interests**
The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 18 March 2022 Accepted: 8 March 2023
Published online: 13 April 2023

**References**

Andreou, E., Ghysels, E., & Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics, 158*, 246–261.
Andreou, E., Ghysels, E., & Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business and Economic Statistics, 31*, 240–251.
Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., & Rünstler, G. (2011). Short-term forecasts of Euro Area GDP growth. *Econometrics Journal, 14*(1), 25–44.
Angelini, E., Bańbura, M., Rünstler, G. (2008). Estimating and forecasting the Euro Area monthly national accounts from a dynamic factor model. In *European Central Bank Working Paper, Series 953*.
Antipa, P., Barhoumi, K., Brunhes-Lesage, V., & Darne, O. (2012). Nowcasting German GDP: A comparison of bridge and factor models. *Journal of Policy Modeling, 34*(6), 864–878.
Artis, M. J., Banerjee, A., & Marcellino, M. (2005). Factor forecasts for the UK. *Journal of Forecasting, 24*, 279–298.
Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the Euro Area GDP. *International Journal of Forecasting, 20*(3), 447–460.
Bańbura, M., & Rünstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting, 27*(2), 333–346.
Banerjee, A., & Marcellino, M. (2006). Are there any reliable leading indicators for us inflation and GDP growth? *International Journal of Forecasting, 22*(1), 137–151.
Barhoumi, K., Benk, S., Cristadoro, R., Reijer, A. D., Jakaitiene, A., Jelonek, P., Rua, A., Rünstler, G., Ruth, K., & Nieuwenhuyze, C. V. (2008). Short-term forecasting of GDP using large monthly datasets: A pseudo real-time forecast evaluation exercise. *European Central Bank, Occasional Paper Series, 84*, 1–25.
Bessec, M. (2012). Short-term forecasts of French GDP: A dynamic factor model with targeted predictors. In *Banque de France, Working Paper 409*.
Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics, 132*(1), 169–194.
Campbell, S. (2007). Macroeconomic volatility, predictability, and uncertainty in the great moderation: Evidence from the survey of professional forecasters. *Journal of Business & Economic Statistics, 25*, 191–200.
Chernis, T., & Sekkel, R. (2017). A dynamic factor model for nowcasting Canadian GDP growth. *Empirical Economics, 53*(1), 217–234.
Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics, 138*, 291–311.
Clements, M. P., & Galvão, A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics, 24*(7), 1187–1206.
D'Agostino, A., McQuinn, K., & O'Brien, D. (2012). Nowcasting Irish GDP. *OECD Journal of Business Cycle Measurement and Analysis, 2*, 1–11.
D'Agostino, A., Giannone, D., Surico, P. (2006). (Un)predictability and macroeconomic stability. In *European Central Bank, Working Paper*, 605.
Dahl, C. M., Hansen, H., & Smidt, J. (2009). The cyclical component factor model. *International Journal of Forecasting, 25*(1), 119–127.
De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics, 146*(2), 318–328.
Den Reijer, A. H. J. (2005). Forecasting Dutch GDP using large scale factor models. In *De Nederlandsche Bank, Working Paper*, 28.
Diebold, F. X. (2020). Real-time real economic activity: Exiting the great recession and entering the pandemic recession. In *Working Paper 27482, National Bureau of Economic Research, July 2020*.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 20*(1), 134–144.
Diron, M. (2008). Short-term forecasts of euro area real GDP growth: An assessment of real-time performance based on vintage data. *Journal of Forecasting, 27*(5), 371–390.
Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics, 164*(1), 188–205.
Duarte, C., & Rua, A. (2007). Forecasting inflation through a bottom-up approach: How bottom is bottom? *Economic Modelling, 24*, 941–953.
Foroni, C., & Marcellino, M. (2014). A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates. *International Journal of Forecasting, 30*(3), 554–568.
Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics, 131*, 59–95.
Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics, 55*(4), 665–676.
Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics, 23*, 365–380.
Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica, 79*(2), 453–497.
Heij, C., van Dijk, D., & Groenen, P. J. (2008). Macroeconomic forecasting with matched principal components. *International Journal of Forecasting, 24*(1), 87–100.
Heij, C., van Dijk, D., & Groenen, P. J. (2011). Real-time macroeconomic forecasting with leading indicators: An empirical comparison. *International Journal of Forecasting, 27*(2), 466–481.
Iacoviello, M. (2001). Short-term forecasting: Projecting Italian GDP, one quarter to two years ahead. In *International Monetary Fund, IMF Working Papers 01/109*.
Jansen, W. J., Jin, X., & de Winter, J. M. (2016). Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts. *International Journal of Forecasting, 32*(2), 411–436.
Kim, H. H., & Swanson, N. R. (2018). Methods for backcasting, nowcasting and forecasting using factor-MIDAS: With an application to Korean GDP. *Journal of Forecasting, 37*(3), 281–302.
Kuzin, V., Marcellino, M., Schumacher, C. (2011). Midas vs. mixed- frequency VAR: Nowcasting GDP in the Euro Area. *International Journal of Forecasting, 27*(2), 529–542.

Marcellino, M., & Schumacher, C. (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics, 72*(4), 518–550.
Marcellino, M., Favero, C. A., & Neglia, F. (2005). Principal components at work: The empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics, 20*(5), 603–620.
Marcellino, M., Stock, J. H., & Watson, M. (2003). Macroeconomic forecasting in the euro area: Country specific versus euro wide information. *European Economic Review, 47*, 1–18.
Peña, D., & Poncela, P. (2004). Forecasting with nonstationary dynamic factor models. *Journal of Econometrics, 119*(2), 291–321.
Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association, 89*, 1303–1313.
Schumacher, C. (2007). Forecasting German GDP using alternative factor models based on large data sets. *Journal of Forecasting, 26*(4), 271–302.
Schumacher, C. (2010). Factor forecasting using international targeted predictors: The case of German GDP. *Economics Letters, 107*, 95–98.
Schumacher, C. (2016). A comparison of MIDAS and bridge equations. *International Journal of Forecasting, 32*(2), 257–270.
Schumacher, C., & Breitungth, J. (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting, 24*, 386–398.
Stakénas, J. (2012). Generating short-term forecasts of the Lithuanian GDP using factor models. In *Bank of Lithuania, Bank of Lithuania, Working Paper Series 13*.
Stock, J. H., & Watson, M. W. (2002a). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics, 20*, 147–162.
Stock, J. H., & Watson, M. W. (2002b). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97*(460), 1167–1179.
Stock, J. H. and M. W. Watson (2005a). Implications of dynamic factor models for VAR analysis, National Bureau of Economic Research, Working Paper No. 11467.
Stock, J. H., & Watson, M. W. (2005b). *An empirical comparison of methods for forecasting using many predictors.* Princeton University, Working paper.
Van Nieuwenhuyze, C. (2005). A generalised dynamic factor model for the Belgian economy: Identification of the business cycle and GDP growth forecasts. *Journal of Business Cycle Measurement and Analysis, 2*(2), 213–247.
White, H. (2000). A reality check for data snooping. *Econometrica, 68*, 1097–1126.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen journal and benefit from:**

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

**Submit your next manuscript at ► springeropen.com**