Bank of Japan Working Paper Series

Nowcasting Japanese GDPs

Kyosuke Chikamatsu [*]
kyousuke.chikamatsu@boj.or.jp

Naohisa Hirakata[*]
naohisa.hirakata@boj.or.jp

Yosuke Kido[*]
yousuke.kido@boj.or.jp

Kazuki Otaka[**]
kazuki.ootaka@boj.or.jp

No.18-E-18
November 2018

Bank of Japan
2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo 103-0021, Japan

[*] Research and Statistics Department
[**] Research and Statistics Department (currently at Personnel and Corporate Affairs Department)

Nowcasting Japanese GDPs

Kyosuke Chikamatsu*, Naohisa Hirakata†, Yosuke Kido‡, and Kazuki Otaka§

November 2018

Abstract

In this paper, we discuss the approaches to nowcasting Japanese GDPs, namely preliminary quarterly GDP estimates and revised annual GDP estimates. First, we look at nowcasting preliminary estimates of quarterly GDP using monthly indicators, ranging from hard data to soft data. In doing so, we compare a variety of mixed frequency approaches, a bridge equation approach, Mixed-Data Sampling (MIDAS) and factor-augmented version of these approaches, and also discuss the usefulness of forecast combination. Second, we work on nowcasting revised annual GDP, which is compiled with comprehensive statistics but only available after a considerable delay. In nowcasting the revised annual GDP, we employ several benchmarking methods, including Chow and Lin (1971), and examine the usefulness of monthly supply-side indicators to predict revised annual GDP. Our findings are summarized as follows. First, regarding nowcasting preliminary quarterly GDP, some of the mixed frequency models discussed in this paper record out-of-sample performance superior to an in-sample mean benchmark. Furthermore, there is a gain from combining model forecasts and professional forecasts. Second, regarding nowcasting revised annual GDP, some benchmarking models that exploit supply-side data serve as useful tools for predicting revised annual growth rates.

JEL classification: C53
Keywords: Nowcasting; Forecast Combination; Mixed-Data Sampling (MIDAS); Benchmarking

* Bank of Japan (E-mail: kyousuke.chikamatsu@boj.or.jp).
† Bank of Japan (E-mail: naohisa.hirakata@boj.or.jp).
‡ Bank of Japan (E-mail: yousuke.kido@boj.or.jp).
§ Bank of Japan (E-mail: kazuki.ootaka@boj.or.jp).

We are grateful for helpful comments from Naoko Hara, Hibiki Ichiue, Naoya Kato, Tomonori Murakoshi, Ichiro Muto, Shinsuke Ohyama and Toshitaka Sekine. The views expressed herein are those of the authors alone and do not necessarily reflect those of the Bank of Japan.

1. Introduction

Understanding the current state of economy is crucial for policy makers. However, due to the inevitable publication delays of some key economic data, such as GDP, policy makers are forced to set policies without knowing the current state, and sometimes, even without knowing the past state, of the economy.

Nowcasting, the prediction of the current state of the economy, consequently has a growing body of literature around it. Although GDP is compiled mostly at a quarterly frequency and released with a lag, many business cycle indicators are more timely and available at higher frequencies; e.g., monthly industrial production data, high-frequency financial data, or big-data obtained from internet/electronic transactions. Economists want to exploit such data in the most efficient way to monitor current state of economy in a timely manner.

Another important yet often neglected problem is data revision. While the vast majority of nowcast literature focuses on predicting the preliminary estimate of GDP, in reality, there are subsequent data revisions, which can sometimes be sizeable. Therefore, nowcasting the revised version of GDP, or in other words, nowcasting the subsequent revision of GDP, is equally important. Indeed, in the case of Japan, which we focus on in this paper, the preliminary estimates of GDP are subject to relatively large revisions at the annual update, which is released with a delay of up to two years.

In this paper, we study two types of nowcasting in the context of Japanese GDP. First, we discuss nowcasting the preliminary estimates of quarterly GDP. To effectively utilize timely indicators which are available monthly, ranging from hard data (e.g., industrial production index) to soft data (business surveys), we employ mixed frequency approaches, Mixed-Data Sampling (MIDAS) and bridge equation approach. In doing so, we also examine factor models that utilize a novel sparse principal component approach, and also examine combination of models and professional forecasts. Second, we discuss the nowcasting of the revised annual GDP, which is compiled based on several annual statistics which are comprehensive, but only available after a considerable delay. In nowcasting the revised annual GDP, we employ benchmarking methods popularly used in national accounts compilation (e.g., Chow and Lin, 1971; and Denton, 1971), exploiting timely indicators.

Our findings are outlined as follows. First, in nowcasting the preliminary quarterly GDP estimates, we find that out-of-sample forecasts produced by the employed models, namely MIDAS and bridge equation models, outperform those of an in-sample mean benchmark. In addition, we find that there is a gain from employing a sparse principal component approach and from combining individual model forecasts and professional forecasts. Second, regarding nowcasting the revised annual GDP, benchmarking methods that exploit supply-side monthly indicators serve as useful tools for predicting the revised annual GDP.

As for previous studies of the Japanese economy, Hara and Yamane (2013), Urasawa (2014), and Bragoli (2017) develop short-term forecasting models to conduct real-time GDP forecasts for Japan and assess performance. These papers show that the forecasts generated by the proposed model are comparable with or outperform the simple univariate model/professional forecasts. Our paper contributes to the literature through the following points. First, while the approaches are dynamic or static factor models, we employ a variety of mixed frequency models, including MIDAS models and Factor MIDAS models, and employ a novel sparse principal component analysis (SPCA) approach in extracting factors. Second, we consider forecast combination schemes and examine the usefulness of combining model forecasts and professional forecasts. Another important feature of this paper is that, while previous literature focuses on the estimation of preliminary quarterly GDP estimates, we also discuss the method of predicting revised annual GDP, which is more accurate but only available after a considerable delay.

Some central banks utilize nowcasting as a method of capturing current economic conditions in a timely fashion. For example, the Federal Reserve Bank of New York (NY Fed) releases the New York Fed Staff Nowcast regularly. In the NY Fed, the aim of the nowcast is to provide a model-based counterpart to the forecasts, which have traditionally been based on expert judgment. The Federal Reserve Bank of Atlanta (Atlanta Fed) releases GDPNow. In the Atlanta Fed, GDPNow is best viewed as a running estimate of real GDP growth based on available data for the current measure quarter. These nowcasts are updated quite frequently: the New York Fed Staff Nowcast is updated every Friday (except on Federal holidays) and the GDPNow forecast is updated six or seven times a month.[1] As for the methodology, in both cases, a dynamic factor model is employed to effectively distill the information contained in large data sets through a small number of common factors.[2] Bank of England (BOE) reports the nowcast of GDP in the Inflation Report. While the nowcast that represents the MPC estimate of GDP growth in the current quarter is ultimately judgmental, they are heavily informed by statistical models, such as MIDAS models.[3] Furthermore, in the inflation report, BOE also utilizes backcasts in the final estimate of GDP. In addition, Norges Bank regularly releases short-term forecasts for GDP growth rates and inflation rates produced by the system for averaging models (SAM). SAM combines the forecast of three types of models: Vector Autoregression (VAR); a leading indicator model; and a factor model (Aastveit et al., 2011).

It should be stressed that the models discussed in this paper are no magic wand. Like other central banks' models, or other models in the nowcasting field, our models are accompanied by some estimation errors (in our case, the root mean square error of around 0.4). Moreover, the optimal models can be changed in the future, as more data become available. Therefore, models proposed in this paper should be regularly reviewed and updated if necessary.

The paper is organized as follows. In Section 2 we describe the revision processes of Japanese GDP. Section 3 discusses approaches to nowcasting preliminary quarterly GDP estimates. Section 4 discusses approaches to nowcasting revised annual GDP. Section 5 concludes.

2. Japanese GDPs

National statistical offices publish quarterly estimates of GDP exploiting timely indicators available at a high frequency prior to the release of high quality estimates of

[1] It should be noted that in both banks, the nowcasts are not an official forecast but a staff forecast. For example, the Atlanta Fed emphasizes that the nowcasts are based solely on the mathematical results of the model and are not adjusted by judgment or subjective factors.

[2] Higgins (2014) provides a detailed description of the data sources and methods implemented in the GDPNow model.

[3] Anesti et al. (2017) describe the details of the approach to GDP nowcasting.

annual GDP. There is inevitably a delay to the release of the relevant annual benchmark, thus, compilation of initial quarterly estimates requires some forecasts. The published quarterly estimates are then revised when the annual benchmark becomes available. In other words, quarterly estimates of GDP can be viewed as a nowcast/forecast of annual GDP.

Like other countries, in Japan, the initial quarterly estimate of GDP is compiled based on timely indicators, then the initial quarterly estimate is revised along with the increasing availability of more relevant data. Specifically, GDP series are revised in the following manner:

1. Quarterly Estimates of GDP: First Preliminary Estimates (**QE1**)

   QE1 is released with a delay of around 45 days; e.g., the first quarter QE1 is released in mid-May.

2. Quarterly Estimates of GDP: Second Preliminary Estimates (**QE2**)

   QE2 is released with a delay of around 75 days; e.g., the first quarter QE2 is released in mid-June.

3. Annual Report on National Accounts (**ARNA GDP**)

   ARNA GDP is an annual GDP series released with a delay of up to two years; i.e., this year's ARNA GDP will be released in December of next year.[4]

In each revision, the quality of the estimates improves as more information become available. The compilation of the ARNA GDP utilizes comprehensive supply-side annual data (e.g. the Census of Manufacture; and the Economic Census), which leads to improved accuracy.[5] On the other hand, the compilation of Quarterly Estimates (QE1 and QE2) partly relies on demand-side data (e.g., the Family Income and Expenditure Survey; and the Financial Statements Statistics of Corporations by Industry). Although GDP measured with demand-side data and measured with supply-side data are conceptually identical, in reality they do not coincide. As a consequence, Quarterly

---

[4] To be precise, the first estimate of ARNA GDP is initially released and is subsequently revised in the second and third annual estimates. Also, it should be noted that ARNA GDP can be revised due to methodological changes, including the update of the national accounts compilation manuals.

[5] In this paper, supply-side data and demand-side data are defined according to Cabinet Office (2018). Supply-side data are used for the commodity-flow method of estimating GDP, and are typically collected from firms that supply goods or services. On the other hand, demand-side data are typically collected from households and firms that purchase (demand) goods and services.

Estimates are subject to some revision when ARNA GDP is released.

Thus, in this paper, we discuss two central issues. In Section 3, as in most of the nowcast literature, we discuss the method of nowcasting QEs. In doing so, we take a variety of mixed frequency approaches, a bridge equation approach, Mixed-data sampling (MIDAS) approach and factor-augmented version of these approaches, exploiting monthly indicators (e.g., the index of industrial production) partly available prior to the release of QE1.

In Section 4, we conduct nowcast exercises of the annual revised GDP series, ARNA GDP. As Quarterly Estimates are subject to subsequent revision, it can be potentially misleading to conduct business cycle analysis solely with Quarterly Estimates. To mitigate such data uncertainty due to revisions, we discuss the methods used to predict the ARNA GDP prior to its release. In doing so, we consider popularly used benchmarking methods and compare their performances with Quarterly Estimates.

3. Nowcasting Quarterly Estimates (QEs)

This section focuses on forecasting the preliminary quarterly GDP growth of Japan, which is only available with a delay of two months. To estimate quarterly GDP growth, economists want to efficiently utilize data available at higher frequencies such as industrial production or survey data available at monthly. In this section, we compare several mixed frequency approaches: the bridge equation approach, Mixed-Data Sampling (MIDAS), and factor-augmented version of these approaches (see, for example, Bańbura et al., 2013; Schumacher, 2016 for survey).[6,7] Then we discuss forecast combination.

[6] One of other popular approaches is a dynamic factor model approach, which specifies a joint model for the variable of interest and for predictors. See, for example, Bańbura and Modugno (2014) for dynamic factor model.

[7] In this paper, the estimations of MIDAS are done using the R package *midasr* developed by Ghysels et al. (2016).

3.1 Nowcast models

Basic Bridge Equation Model

One of the econometric approaches to forecast in the presence of mixed-frequency data is a bridge equation approach. Bridge equations are linear regressions that link high frequency variables, such as industrial production or survey data, to low frequency ones, e.g., the quarterly real GDP growth, providing some estimates of current and short-term developments in advance of the release (see e.g. Baffigi et al., 2004). This method allows the computation of early estimates of low-frequency variables by using high frequency indicators. A bridge equation model is defined as:

$$y_t = \alpha + \sum_{i=1}^{N} \beta_i x_{i,t}^{Q} + \epsilon_t \ , \tag{1}$$

where y_t is GDP growth in quarter t, $x_{i,t}^{Q}$ is a monthly economic indicator i converted to quarterly value at quarter t. For growth rates, conversion can be done with the following equation (Mariano and Murasawa, 2003):

$$x_{i,t}^{Q} = \frac{1}{3}(x_{i,3t}^{M} + 2x_{i,3t-1}^{M} + 3x_{i,3t-2}^{M} + 2x_{i,3t-3}^{M} + x_{i,3t-4}^{M}) \ , \tag{2}$$

where $x_{i,3t}^{M}$ is a monthly growth rate of indicator i in $3t$th month.

Estimation for equation (1) is performed by the Ordinary Least Square (OLS) method. In the forecasting exercise, forecasters typically face ragged edge data, where monthly indicators are only partially available. To handle ragged edge data, auxiliary equations are employed to forecast missing values.

Bridge equation models can be further extended to include lagged dependent variables. However, Stock and Watson (2002) show that the introduction of lagged dependent variables creates efficiency losses. Thus, in this paper we do not include lagged dependent variables.

Bridge Equation Model with Factors

An extension of the basic bridge equation model is the introduction of factor terms. Factor approaches are employed in many empirical exercises, with the aim of capturing movements reflected in a large dataset. The information included in a large dataset can be summarized by a few factors that represent key economic driving forces. Factor models read as

$$y_t = \alpha + \sum_{i=1}^{N} \beta_i \, x_{i,t}^{Q} + \sum_{i=1}^{M} \gamma_i \, F_{i,t}^{Q} + \epsilon_t, \tag{3}$$

where y_t is GDP growth in quarter t, $x_{i,t}^{Q}$ is a selected monthly economic indicator i converted to quarterly value at quarter t, $F_{i,t}^{Q}$ is a ith factor extracted from monthly indicators converted to quarterly value at quarter t. Monthly-quarterly conversion is done using equation (2).

In this paper, we consider two approaches for estimating factors. The first approach is the standard principal component analysis (PCA) approach employed by a large number of forecasting exercises (e.g., Stock and Watson, 2002). The second approach is sparse principal component analysis (SPCA) proposed by Zou et al. (2006).

Under standard PCA, the factor loading coefficients are all typically nonzero, making interpretation of the estimated components difficult. Zou et al. (2006) address this issue by proposing a modified method in which the LASSO (or elastic net) is used to construct principal components with sparse loadings (see Appendix A for methodological detail). A number of recent studies apply SPCA for forecasting exercises and often obtain improvements in forecasting accuracy compared with standard PCA approaches (Kristensen, 2017; Kim and Swanson, 2018). In estimating SPCA, we obtain factors based on LASSO-type loss function, and select a tuning parameter based on Bayesian Information Criterion (BIC) following Kristensen (2017).[8]

[8] In this paper, the calculation of SPCA is done using the R package *sparsepca.*

MIDAS Model

Mixed-Data Sampling, or MIDAS, regressions are a recently developed framework (Andreou et al., 2010) for handling regressions where the dependent variable is of a lower frequency to the explanatory variables. MIDAS models have been employed and have shown promise in forecasting contexts (e.g., Foroni and Marcellino, 2014; Kuzin et al., 2013; and Schumacher, 2016) as well as in structural analysis (Ferrara and Guérin, 2018).

MIDAS models can be applied to a context such as ours, where monthly variables (e.g., industrial production, survey indices, and possibly other variables) are used to forecast quarterly GDP growth. Unlike bridge equation models, which rely on auxiliary regressions to forecast explanatory variables, MIDAS models are a direct forecasting tool that directly estimates the current quarter with a lag structure which corresponds to ragged edge data. This yields different models for different forecasting horizons. A MIDAS model reads as:

$$y_t = \alpha + \sum_{i=1}^{k} \sum_{j=0}^{l_i} \beta_{i,j} x_{i,3t-j}^{M} + \epsilon_t. \tag{4}$$

In (4), y_t denotes a GDP growth in quarter t. $x_{i,3t-j}^{M}$ denotes a monthly economic indicator i in $3t-j$th month. k stands for the number of indicators, and l_i denotes the number of lags for the indicator i in terms of month.

In the standard MIDAS literature, functional lag polynomials are employed to avoid parameter proliferation in the case of long high-frequency lags (Andreou et al., 2010). This approach can have the advantage of greater parsimony — if a large number of parameters needed be estimated — otherwise it comes at the cost of flexibility.

On the other hand, unrestricted MIDAS (U-MIDAS), a variant of the MIDAS approach considered by Foroni et al. (2015), does not impose restrictions on parameters with a certain functional form. Foroni et al. (2015) argues that U-MIDAS outperforms standard MIDAS where differences in sampling frequencies are small, as is the case with our quarterly-monthly data. Therefore, in this paper, we employ U-MIDAS.

To deal with the possible presence of misspecification and the availability of numerous indicators, forecast combination is often applied in MIDAS models.

Factor MIDAS Model

It is possible to augment the MIDAS regressions with the factors extracted from a high frequency dataset in order to exploit large high frequency datasets for predicting low-frequency variables. While the basic MIDAS framework consists of a regression of a low-frequency variable on a set of high-frequency indicators, the Factor-MIDAS approach exploits estimated factors rather than single or small groups of economic indicators as repressors. Factor MIDAS has been employed in recent forecasting papers and often improves forecast performances (Foroni and Marcellino, 2014; Schumacher, 2016; and Kim and Swanson, 2017). Factor MIDAS can be defined as

$$y_t = \alpha + \sum_{i=1}^{k_1}\sum_{j=0}^{l_{1,i}} \beta_{i,j} x_{i,3t-j}^{M} + \sum_{i=1}^{k_2}\sum_{j=0}^{l_{2,i}} \gamma_{i,j} F_{i,3t-j}^{M} + \epsilon_t, \quad (5)$$

where $F_{i,3t-j}^{M}$ is a ith factor extracted from monthly indicators in $3t-j$ th month. k_1 denotes the number of indicators, and k_2 denotes the number of factors, respectively. $l_{1,i}$ denotes the number of lags for indicator i, and $l_{2,i}$ denotes the number of lags for factor i in terms of month.

Because of small differences in sampling frequencies, we choose U-MIDAS following Foroni et al. (2015) instead of standard MIDAS models.

3.2 Forecast Combination

In Section 3.1, we discuss individual forecasting models. In this subsection, we discuss the ways to combining multiple forecasts. There is a large body of literature that suggests that forecast combinations can provide more accurate forecasts by combining multiple models rather than relying on a specific model (see Hendry and Clements, 2004; and Timmermann, 2006).[9] One justification for using forecast combinations methods is the presence of model uncertainty that forecasters face due to different sets

[9] See, for example, Ohyama (2001) for the application of forecast combination to Japanese data.

of predictors and different modeling approaches. In addition, forecast combinations can deal with model instability and structural breaks (Hendry and Clements, 2004).

Given N individual forecasting models, a linear version of forecast combination can be defined as follows

$$Y_t = \sum_{j=1}^{N} w_{j,t} Y_{j,t}, \tag{6}$$

where $w_{j,t}$ is a weight given to the jth forecast model in period t.

Forecast combinations have frequently been found in empirical studies to produce better forecasts than methods based on individual forecasting models. However, there is no consensus concerning how to form the forecast weights. Thus, in this paper we consider variety of weighting schemes.

A Simple Average

First, we consider a simple average method that ignores historical performance. Despite more refined combination schemes existing, much of the literature finds that simple averages perform well.[10] Under this simple scheme, a weight is given as

$$w_{i,t} = \frac{1}{N}. \tag{7}$$

Triangular Kernel Approach

Simple rank-based weighting schemes are another popular approach. The most common scheme in this class is to use the median forecast (Hendry and Clements, 2004). Alternatively, forecasters can consider a triangular weighting scheme that sets the combination weights inversely proportional to the models' performance rank (Aiolfi and Timmermann, 2006). Since ranks are likely to be less sensitive to outliers, this weighting scheme can be expected to be more robust. With this scheme, weights can be computed as

[10] This phenomenon is often called the "forecast combination puzzle" (Stock and Watson, 2004).

$$w_{i,t} = \frac{1/r_{i,t}}{\sum_{j=1}^{N} 1/r_{j,t}}, \tag{8}$$

where $r_{j,t}$ denotes the model *j*'s rank at period t, which is determined according to mean square forecast error (MSFE).

Inverse Mean Square Error

Another popular approach is to determine a weight proportional to the inverse of mean square error (MSE). Theoretically, this weighting scheme corresponds to the optimal weighting scheme discussed by Bates and Granger (1969) when individual forecasts are uncorrelated. The weights are computed as:

$$w_{i,t} = \frac{1/MSFE_{i,t}}{\sum_{j=1}^{N} 1/MSFE_{j,t}}, \tag{9}$$

where $MSFE_{i,t}$ denotes the model i's mean square forecast error (MSFE) at t.

Discounted Mean Square Forecast Errors

Stock and Watson (2004) propose the discounted mean square forecast error (dMSFE) forecast combination method. Each individual predictor is given a weight according to its historical performance and the weight is inversely proportional to the predictor's dMSFE. The discount factor attaches greater weight to the recent predictive ability of the individual predictor. The weights are given as

$$w_{i,t} = \frac{1/m_{i,t}}{\sum_{j=1}^{N} 1/m_{j,t}}, \tag{10}$$

$$m_{i,t} = \sum_{s=T_0}^{t} \delta^{t-s} \left(Y_s - \hat{Y}_{i,s}\right)^2,$$

and $\hat{Y}_{i,s}$ denotes model i's forecast for time s. In this paper, the dMSFE forecasts are computed for the value of $\delta =$ 0.95, following Stock and Watson (2004).

Bayesian Model Averaging

Buckland et al. (1997) propose popular Bayesian Model Averaging (BMA) approaches for combining models. Under this approach, weights are defined as

$$w_{i,t} = P(M_i|D_t) = \frac{P(D_t|M_i) * P(M_i)}{\sum_{j=1}^{N} P(D_t|M_j) * P(M_j)}, \tag{11}$$

where $P(D_t|M_i)$ denotes model i's $(M_i's)$ marginal likelihood for the dataset $D_t = (Y_1, \dots, Y_t)$ and $P(M_i)$ denotes prior weight for model i.

Our BMA weights are set as:

$$w_{i,t} = P(M_i|D_t) \approx \frac{\exp\left(-\frac{BIC_{i,t}}{2}\right)}{\sum_{j=1}^{N} \exp\left(-\frac{BIC_{j,t}}{2}\right)}, \tag{12}$$

where $BIC_{i,t}$ denotes Bayesian Information Criterion (BIC) for model i. This is an approximate BMA for the case of equal model priors.

3.3 Empirical Application

3.3.1 Data

Typically, short-term indicators, such as business surveys or industrial production indexes, are released at a monthly frequency and are partially available before the release of GDP growth data. Thus, in this paper, we consider a set of monthly indicators for forecasting GDP growth (see Appendix B for data description).

For hard data, we employ the Index of Industrial Production (IIP), the Index of Tertiary Industry Activity (ITA), and the Current Survey of Commerce (CSC). The IIP captures monthly production activity of the manufacturing sector and is widely utilized in a forecasting context (see, e.g., Bragoli, 2017 for Japan's case). The ITA captures monthly activity of the service sector and is often utilized in forecasting Japan's GDP growth (Hara and Yamane, 2013; and Bragoli, 2017). In addition, we also consider the

Current Survey of Commerce (CSC), which the captures sales value of firms.[11] The CSC is closely related to the ITA, but is complementary as it sometimes exhibits different movement and is available slightly earlier than the ITA.

For soft data, following previous research, we utilize a set of business survey indicators obtained from Reuters' Tankan Survey (Urasawa, 2014; and Bragoli, 2017).[12] In doing so, we extract factors from a set of survey indicators using a sparse principal component analysis, rather than relying on a single indicator. Survey data can be useful as they are promptly available. However, it is unclear whether such survey data improve forecasting performance. Thus, we also examine the case where only hard data are included in explanatory variables.

It should be noted that we intentionally use final vintage data rather than real-time vintage. In Japan, GDP has undergone a large revision due to methodological changes, including the update of National Accounts Compilation Manual.[13] This suggests that, if we employ real-time data, models may not be suitably estimated for forecasting GDP because GDP is compiled with the updated methodology. For that reason, we choose final vintage data, which are compiled in line with the updated methodologies.

3.3.2 Forecast evaluation schemes

In this subsection, we evaluate out-of-sample forecasting performances of the bridge equation models and MIDAS models described in Section 3.1. In the empirical application, we evaluate model forecasts at two months (60 days) and one month (30 days) prior to GDP release dates, based on increasingly available information from the indicators. That is, in our forecast experiment, we consider the ragged-edge of the monthly indicators according to the data publication lag reported in Appendix B. For MIDAS models, GDP forecasts are directly computed with ragged edge data. For bridge models, ragged edge data are handled in following ways. First, we extrapolate missing data due to publication lag, exploiting already-published data of the corresponding

[11] For the CSC, we use wholesale value and deflate it with producer price index.
[12] We also utilize the Economic Watcher Survey to extrapolate the ITA using bridge equation models. See Appendix C for the auxiliary equations for the extrapolation.
[13] For example, in 2016, Japanese GDP was revised due to the introduction of 2008SNA. Hara and Ichiue (2011) describe details of other major methodological changes.

month based on auxiliary regressions (see Appendix C for methodology). In the case of two-month ahead forecasts, where no data in the third month of the quarter are available, we simply assume that the third month of GDP level is equal to the average of first month and second month of GDP. With this simple assumption, we calculate quarterly GDP growth rate according to the monthly-quarterly growth rate conversion of Mariano and Murasawa (2003) (see Appendix D for detail).[14]

In our recursive forecasting experiment, we split the sample (from 2000Q2 to 2018Q1) into an estimation subsample and an evaluation subsample. First, we estimate parameters with the subsample between 2000Q2 and 2012Q4, and then extend the estimation subsample recursively.

For benchmark comparison, we compute in-sample mean of GDP growth, which is also recursively re-estimated for every estimation sample. In the forecast literature, this benchmark has turned out to be a strong competitor to more sophisticated approaches (Giannone et al., 2008; Kuzin et al., 2013). In evaluating forecast performances, we conduct the Diebold-Mariano test (Diebold and Mariano, 1995) and compare predictive accuracy between the model forecasts and this simple benchmark.[15]

3.3.3 Forecast evaluation (individual models)

Table 1 reports root mean square error (RMSE) of bridge equation models. For most of the models considered, RMSE is lower than the simple benchmark model. Furthermore, in some cases, predictive accuracy is statistically better than the simple benchmark according to the Diebold-Mariano test. The table also compares bridge equation models with factors estimated with PCA and SPCA.[16] Although SPCA improves the predictive accuracy of some models, the difference is not large. Among others, the model with the IIP, the ITA and the first factor of surveys extracted by SPCA records the best performance.

[14] We also examine the approaches employed by Altissimo et al. (2010), where third-month GDP level (or third-month growth rate) is equal to that of second-month. We find that the approach considered in this paper (i.e., third-month GDP level is equal to the average of first-month GDP level and second-month GDP level) works better.

[15] Throughout the paper, we apply quadratic loss function for the Diebold-Mariano test.

[16] In some cases, the results obtained with PCA and SPCA are identical since the tuning parameter of SPCA is estimated as 0 (see Appendix A).

For MIDAS models, we report forecast performances of each single indicator models, in addition to multi-indicator models that correspond to the aforementioned bridge equation models (Table 2).[17] In some models, RMSE is lower than the simple benchmark model, with the predictive accuracy statistically better than the benchmark. Overall, however, results are mixed, suggesting that some of the models suffer misspecification or parameter instability. Thus, following the MIDAS literature (Clements and Galvão, 2008; Andreou et al., 2013; Kuzin et al., 2013; Anesti et al., 2017), in the next section, we apply forecast combination to MIDAS models.

3.3.4 Forecast evaluation (forecast combination)

Combined MIDAS

To deal with the possible presence of misspecification and the availability of many indicators, forecast combination is often applied to MIDAS models (Clements and Galvão, 2008; Andreou et al., 2013; Kuzin et al., 2013; Anesti et al., 2017). Thus, we pool MIDAS models evaluated in Section 3.3.3 and combine them using the weighting schemes discussed in Section 3.2. For comparison, we report two sets of combinations: the combination of MIDAS models and Factor MIDAS models estimated with PCA listed in Table 2; and the combination of MIDAS models and Factor MIDAS models estimated with SPCA listed in Table 2. As with Section 3.3.3, we evaluate models' out-of-sample performances, with the evaluation sample from 2013Q1 to 2018Q1. In doing so, we update weights for combination recursively according to the past out-of-sample performances.[18]

Table 3 reports out-of-sample forecast performances of combined MIDAS models. As reported, in many cases, combined MIDAS models work well compared with the individual MIDAS models discussed in Section 3.3.3. It is also found that the combination that utilizes SPCA records superior performances compared with the combination that utilize regular PCA. Among other combinations, the simple average of

[17] MIDAS models considered in this paper include at most three predictor variables to prevent proliferation of number of parameters.

[18] To compute weights in this way, we need out-of-sample performances prior to 2013Q1. Hence, we compute out-of-sample performances from 2008Q1 to 2012Q4 and utilize them for computing weights.

MIDAS models and Factor MIDAS models estimated with SPCA shows the best performances for both forecasts at 1-month and 2-month prior to GDP release dates.

Forecast combination with professional forecasters

An interesting question is whether a combination of model forecasts and judgmental forecasts from the professional forecasters produces better results than the best single ones. This approach can be appealing, given that the sets of information available from models and judgmental forecasts differ. In particular, judgmental survey forecasts may exploit subjective information as well as anticipated policy effects. Aiolfi et al. (2011) find that the combination of model-based forecasts and judgmental survey forecasts improve forecast accuracies.

Thus, in this subsection, we also combine model forecasts and surveys conducted by Bloomberg and Japan Center for Economic Research (JCER). Bloomberg conducts a survey and collects forecasts from professional forecasters in order to produce predictions for GDP and other market-relevant variables before their release dates. Likewise, JCER conducts JCER ESP Forecast, a survey on short-term forecasts of Japanese GDP, and collects forecasts from professional economists. JCER ESP Forecast is published monthly, and is utilized in many empirical exercises on Japan (see, for example, Miyamoto et al, 2018; and Bragoli, 2017).

As for model forecasts, we utilize the combined MIDAS model that records the best performance (i.e., the combination that employs SPCA and a simple weighting approach) and the bridge equation model that also records the best performance (i.e., the bridge equation model using the IIP, ITA, and the first factor of SPCA). We examine all possible combinations of forecasts and apply weighting schemes as listed in Section 3.2. Since this combination includes model-free forecasts, we do not calculate Bayesian Model Averaging. We evaluate out-of-sample performances, with the evaluation sample from 2013Q1 to 2018Q1. As done previously, we update weights for combination recursively according to the past out-of-sample performances.

Table 4 reports out-of-sample performances of combined forecasts and individual forecasts.[19] As reported in the table, in many cases, combined forecasts outperform

[19] RMSE for the surveys is also calculated based on final vintage GDP growth rates.

individual forecasts. Among the variety of combinations, the combination of the Bridge equation model and combined MIDAS weighted with simple average records the best performances at 2-month prior to GDP release dates, with predictive accuracy statistically better than the simple benchmark model. At 1-month prior to GDP release dates, however, the combination of combined MIDAS and JCER survey weighted with simple average records the best performance, suggesting that combining judgmental forecasts with model forecasts can improve predictive accuracy. On average, the combination of the Bridge equation model, combined MIDAS and JCER survey weighted with simple average records the best performances (Table 4, Figures 1 and 2).

4. Nowcasting the ARNA GDP

The previous section discusses the methods used to nowcast the Quarterly Estimate (QE) of Japanese GDP. However, as discussed in Section 2, QE is subject to subsequent revisions at the release of ARNA. In that sense, QE is viewed as one way to nowcast ARNA, and its predictive accuracy could possibly be improved. Thus, this section discusses alternative ways to nowcast the ARNA GDP.

In this section, we discuss nowcasting the ARNA GDP, employing benchmarking methods popularly used in national accounts compilation. Benchmarking methods derive quarterly estimates of an annual aggregate, exploiting preliminary series available at a higher frequency. The methods can be split into two parts: distribution and extrapolation. The former procedure is used to generate quarterly series which are consistent with annual values (i.e., the quarter values sum up to the annual value), while mimicking the movement of preliminary series. Extrapolation refers to the calculation of quarterly series based on preliminary indicators before the annual benchmark, the ARNA GDP, become available. Thus, benchmarking methods can be used to calculate an early projection of the ARNA GDP before it become available.

In this section, we discuss the method used to nowcast the ARNA GDP. In doing so, we assess the predictive accuracy of popularly used benchmarking methods. The following subsection discusses the methods and empirical results.

4.1 Benchmarking Methods

4.1.1 Regression Based Method

Let $\mathbf{x}$ be an unobserved quarterly GDP series to be estimated, and let $\mathbf{Z}$ be the matrix of quarterly indicator series with $4M \times N$, where M is the sample size in year, and N is the number of indicator variables used for estimating the unobserved quarterly GDP $\mathbf{x}$. We specify the relationship between $\mathbf{x}$ and $\mathbf{Z}$, as follows:

$$\mathbf{x} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{u} \tag{13}$$

$$\text{s.t. } E(\mathbf{u}) = 0,\ E(\mathbf{u}\mathbf{u}') = \mathbf{V}.$$

Next, we introduce a transformation matrix $\mathbf{C}$, which annualizes quarterly series and is defined as:

$$\mathbf{C} \equiv \mathbf{I_M} \otimes \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}. \tag{14}$$

We introduce $\mathbf{y}$, which denotes the $M \times 1$ vector of observed annual GDP series for M years. The relationship between the observed annual GDP series $\mathbf{y}$ and the unobserved quarterly GDP series $\mathbf{x}$ can be expressed as:

$$\mathbf{y} = \mathbf{C}\mathbf{x}. \tag{15}$$

Using (13) and (15), we obtain the following equation:

$$\mathbf{y} = \mathbf{C}\mathbf{Z}\boldsymbol{\beta} + \mathbf{C}\mathbf{u}. \tag{16}$$

Given that covariance matrix $\mathbf{V}$ in (13), the GLS estimator $\hat{\boldsymbol{\beta}}_{\mathbf{GLS}}$ can be obtained as:

$$\hat{\boldsymbol{\beta}}_{\mathbf{GLS}} = [\mathbf{Z}'\mathbf{C}'(\mathbf{CVC}')^{-1}\mathbf{CZ}]^{-1}\mathbf{Z}'\mathbf{C}'(\mathbf{CVC}')^{-1}\mathbf{y}. \tag{17}$$

The GLS estimator $\hat{\boldsymbol{\beta}}_{\mathbf{GLS}}$ of (17) is proposed by Chow and Lin (1971). The estimated quarterly GDP series $\hat{\mathbf{x}}$ is given by:

$$\hat{\mathbf{x}} = \mathbf{Z}\widehat{\boldsymbol{\beta}} + \mathbf{VC}'(\mathbf{CVC}')^{-1}\big[\mathbf{y} - \mathbf{CZ}\widehat{\boldsymbol{\beta}}\big]. \tag{18}$$

(18) states that the quarterly fitted value is composed of the value predicted by preliminary series $\mathbf{Z}\widehat{\boldsymbol{\beta}}$ and forecast error for annual value adjusted by variance-covariance matrix $\mathbf{VC}'(\mathbf{CVC}')^{-1}\big[\mathbf{y} - \mathbf{CZ}\widehat{\boldsymbol{\beta}}\big]$. Specifically, when $\boldsymbol{V} = \mathbf{I}$, $\boldsymbol{\beta}_{\mathbf{GLS}}$ is reduced to be an OLS estimator $\widehat{\boldsymbol{\beta}}_{\mathbf{OLS}}$:

$$\widehat{\boldsymbol{\beta}}_{\mathbf{OLS}} = [\mathbf{Z}'\mathbf{C}'(\mathbf{CC}')^{-1}\mathbf{CZ}]^{-1}\mathbf{Z}'\mathbf{C}'(\mathbf{CC}')^{-1}\mathbf{y}. \tag{19}$$

Variance-Covariance Matrix Based on Chow-Lin

Chow and Lin (1971) assume that the error term u_t is governed by stationary AR(1) process,

$$\begin{gathered} u_t = \rho u_{t-1} + \varepsilon_t, \\ \text{where } |\rho| < 1, \varepsilon_t \sim iid(0, \sigma_\varepsilon^2). \end{gathered} \tag{20}$$

Given (20), the variance-covariance matrix $\mathbf{V}_{\boldsymbol{\rho}}$ can be written as:

$$\mathbf{V}_{\boldsymbol{\rho}} = \frac{\sigma_\varepsilon^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{4M-1} \\ \rho & 1 & \rho & \cdots & \rho^{4M-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{4M-1} & \rho^{4M-2} & \cdots & \cdots & 1 \end{pmatrix}. \tag{21}$$

There are several ways to estimate the autoregressive parameter ρ in the Chow-Lin method. Here, we examine two methods, the maximum likelihood approach and the minimization of the weighted residual sum of squares. First, as for the maximum likelihood approach, Bournay and Laroque (1979) suggest the maximization of the following likelihood:

$$\hat{\rho} = \underset{\rho}{\operatorname{argmax}} \frac{\exp\left[-\frac{1}{2}\mathbf{u}'\mathbf{C}'\left(\mathbf{C}\mathbf{V}_{\boldsymbol{\rho}}\mathbf{C}'\right)^{-\mathbf{1}}\mathbf{C}\mathbf{u}\right]}{(2\pi)^{M/2}\cdot\det\left(\mathbf{C}\mathbf{V}_{\boldsymbol{\rho}}\mathbf{C}'\right)^{1/2}}. \tag{22}$$

The second approach is the minimization of the weighted residual sum of squares, as suggested by Barbone et al. (1981):

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}}\ \mathbf{u}'\mathbf{C}'\left(\mathbf{C}\mathbf{V}_{\boldsymbol{\rho}\mathbf{2}}\mathbf{C}'\right)^{-\mathbf{1}}\mathbf{C}\mathbf{u}, \tag{23}$$

where

$$\mathbf{V}_{\boldsymbol{\rho}\mathbf{2}} = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{4M-1} \\ \rho & 1 & \rho & \cdots & \rho^{4M-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{4M-1} & \rho^{4M-2} & \cdots & \cdots & 1 \end{pmatrix}. \tag{24}$$

Variance-Covariance Matrix Based on Fernández

Fernández (1981) suggests the models that assume that the quarterly residuals follow a unit root process:

$$u_t = u_{t-1} + v_t\,. \tag{25}$$

In the case of Fernández (1981), the variance-covariance matrix is expressed as follows:

$$\mathbf{V_0} = \sigma_\varepsilon^2[\mathbf{D}'\mathbf{D}]^{-1}, \tag{26}$$

where $\mathbf{D}$ is the matrix that makes variables first difference:

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}. \tag{27}$$

Variance-Covariance Matrix Based on Litterman

In addition to the unit root process, Litterman (1983) further assumes that v_t follows AR (1) process:

$$v_t = \alpha v_{t-1} + \varepsilon_t. \tag{28}$$

The variance-covariance matrix is expressed as follows:

$$\mathbf{V}_{\boldsymbol{\alpha}} = \sigma_\varepsilon^2 [\mathbf{D}'\mathbf{H}'\mathbf{H}\mathbf{D}]^{-1}, \tag{29}$$

where

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\alpha & 1 & 0 & \cdots & 0 \\ 0 & -\alpha & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\alpha & 1 \end{bmatrix}. \tag{30}$$

α is obtained by solving the following optimization problem:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmax}} \frac{\exp\left[-\frac{1}{2}\mathbf{u}'\mathbf{C}'(\mathbf{C}\mathbf{V}_{\boldsymbol{\alpha}}\mathbf{C}')^{-\mathbf{1}}\mathbf{C}\mathbf{u}\right]}{(2\pi)^{M/2} \cdot \det(\mathbf{C}\mathbf{V}_{\boldsymbol{\alpha}}\mathbf{C}')^{1/2}}. \tag{31}$$

4.1.2 Proportional Denton Method

In the Denton methods, the benchmarked estimates $\mathbf{x}$ are obtained by allocating the discrepancy between the sum of four preliminary quarterly estimates $\mathbf{z}$ and corresponding annual estimates $\mathbf{y}$ to the four quarters in each year. Specifically, for the Proportional First Differences (PFD) version of the Denton method, where the proportional period-to-period changes of the benchmarked series $\mathbf{x}$ are solved by directly linking those of the preliminary variables $\mathbf{z}$, benchmarked estimates can be computed by minimizing the following quadratic loss function:

$$\min_{\mathbf{x}} (\mathbf{x} - \mathbf{z})'\mathbf{Q}(\mathbf{x} - \mathbf{z}), \tag{32}$$

$$\text{s.t. } \mathbf{y} = \mathbf{Cx}, \tag{33}$$

where $\mathbf{Q} = \text{diag}(\mathbf{z})^{-1}\boldsymbol{\Delta}'\boldsymbol{\Delta}\ \text{diag}(\mathbf{z})^{-1}$ and (34)

$$\boldsymbol{\Delta} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{35}$$

In the above, $\boldsymbol{\Delta}$ is the first difference matrix. For the optimization problem described by (32) and (33), Di Fonzo (1994) shows the matrix form solution as follows:

$$\hat{\mathbf{x}} = [\text{diag}(\mathbf{z}) \quad \mathbf{0}] \begin{bmatrix} \boldsymbol{\Delta}'\boldsymbol{\Delta} & \text{diag}(\mathbf{z})\mathbf{C}' \\ \mathbf{C}\text{diag}(\mathbf{z}) & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix}. \tag{36}$$

Given a preliminary variable $\mathbf{z}$, we can extrapolate $\mathbf{x}$ using the estimated quarterly values $\hat{\mathbf{x}}$. Unlike the regression based methods discussed in Section 4.1.1, the proportional Denton method does not model the relationship between preliminary variable and annual benchmark. Instead, it directly links preliminary variable and annual benchmark, smoothing the changes in their proportions two between neighboring quarters.

4.2 Empirical Application

Prior to presenting the estimation results, we describe the data used for the estimation and the models to be evaluated in our exercises.[20]

4.2.1 Data

The ARNA GDP is compiled using annual statistics, such as the Census of manufacturing and the Economic Census for Business Activity, which comprehensively captures Japanese economic activity from supply-side. Our strategy is to employ preliminary series available at a higher frequency that similarly measure Japanese

[20] The estimations use the R package *tempdisagg* developed by Sax and Steiner (2013).

economic activity from the supply-side. Specifically, we consider the following indicators:

- Index of Industrial Production (IIP);
- Index of Tertiary Industry Activity (ITA);
- Index of Construction Industry Activity (ICA); and
- Index of All Industry Activity (IAA).

In the above, all four indicators are available at a monthly frequency (see Appendix B for data description). These indicators capture Japanese economic activity from the supply-side. The first three indicators are sector-specific indicators and should be used jointly to comprehensively capture Japanese economic activity. On the other hand, the IAA can serve as a single preliminary indicator as it captures overall Japanese economic activity.

4.2.2 Models

We estimate unobservable quarterly GDP $\mathbf{x}$ using annual ARNA GDP and the preliminary series listed in Section 4.2.1. We evaluate the performance of the following six methods elaborated on above:

i. **Ordinary Least Square (OLS)**

The parameters $\boldsymbol{\beta}$ are obtained by (19).

ii. **Chow-Lin method with maximum likelihood**

The parameters $\boldsymbol{\beta}$ are estimated by (17). The variance covariance matrix (21) is used and ρ is estimated by (22).

iii. **Chow-Lin method with minimization of residual**

The parameters $\boldsymbol{\beta}$ are estimated by (17). The variance covariance matrix (24) is used and ρ is estimated by (23).

iv. **Litterman method**

The parameters $\boldsymbol{\beta}$ are estimated by (17). The variance covariance matrix (29) is used and α is estimated by (31).

v. Fernández method

The parameters $\boldsymbol{\beta}$ are estimated by (17). The variance covariance matrix (26) is used.

vi. Proportional Denton method

The quarterly GDP values are calculated by (36).

The first five methods are regression based approaches, while the Proportional Denton method is not regression based. For i, ii and iii, we estimate using both levels and log first difference of the indicators. For iv, v and vi we estimate using the level of the indicators.

4.2.3 Estimation results

In this subsection, we evaluate out-of-sample forecasting performances of the benchmarking models listed in Section 4.2.2. In this exercise, we evaluate forecasting performances at the one- and two- year horizon r prior to ARNA GDP release dates. In doing so, we evaluate predictive accuracy of models in terms of predicting ARNA GDP growth rates, based on root mean squared error (RMSE).

In our recursive forecasting experiment, we split the sample (from 1988 to 2016) into an estimation subsample and an evaluation subsample.[21][22] First, we estimate parameters with the subsample between 1988 and 2000, and then extend the estimation subsample recursively.

For comparison, we also compute RMSE of QE1 and QE2.[23] As discussed, QEs are

[21] As for the models that use log first differences, the sample starts from 1989. Also, in the case that the ICA is used, the sample starts from 1993 in the case of level, or from 1994 in the case of log first difference, due to the availability of the ICA.
[22] ARNA GDP starts from 1994. To increase sample size, we calculate GDP from 1988 to 1993 using the growth rates of provisional estimates of GDP released by Cabinet Office.
[23] For QE1 and QE2, annual growth rates are computed from their quarterly growth rates. Prior to 2016/Q3 data, Japanese GDP was compiled based on the System of National Account, 1993 (SNA 1993). To adjust the revisions stemming from compilation methodologies, from 2000 to 2014, we

viewed as a nowcast of ARNA. We conduct Diebold-Mariano tests and compare the predictive accuracy of the benchmarking model forecasts and QEs.

Table 5 shows forecasting performance measured by RMSE calculated in terms of growth rates. As shown in the seventh row of the tables, the models using the IAA and time trend record smaller RMSE, indicating that IAA's out-of-sample performances is superior to that of other indicators. Among the models using the IAA and time trend, Fernández and Litterman record the best performance in terms of both one-year ahead and two-year ahead forecasts; RMSE of 0.464 in one-year ahead forecast and RMSE of 0.438 in two-year ahead forecast.

The last two rows of Table 5 show RMSE for QE1 and QE2. These RMSEs are 0.833 and 0.840 for one-year ahead out-of-sample forecast, and 0.838 and 0.823 for two-year ahead out of sample forecast, respectively. The Fernández and Litterman models using the IAA and time trend record smaller RMSE than QEs.

Table 5 also lists the parameter estimates of ρ and α in (20) and (28). In the Litterman model using the IAA and time trend, the estimate value of α is zero, indicating that the estimated model of Litterman is reduced to the Fernández model in our results.[24] In addition, as stated above, RMSEs of the two models are exactly same.

Figure 3 shows out-of-sample forecasting performance of the Fernández model using the IAA and trend. According to Figure 3, we find that the predicted annual values of each period generated by the Fernández model are in line with the actual ARNA GDP. Figures 4 and 5 plot the in-sample quarterly predicted values in terms of level and growth rate. The values in level and growth rate of quarterly predicted values are broadly in line with quarterly ARNA GDP.

5. Conclusion

In this paper, we discuss the approaches used to nowcast Japanese GDPs, namely

---

use SNA 1993 version of ARNA GDP growth rates to evaluate performances of QEs.

[24] In estimating ρ and α, we employ the restrictions that the $\rho(\alpha)$ is nonnegative and less than 1. When we estimate under the restriction that $|\rho|(|\alpha|)$ is less than 1, optimization procedures don't work well.

preliminary quarterly estimates and revised annual estimates.

First, we attempt to nowcast the preliminary estimates of quarterly GDP using monthly indicators, such as the index of industrial production. In doing so, we employ a variety of mixed frequency approaches, bridge equation approach, MIDAS approach, and factor-augmented version of them, to utilize those data effectively. We find that those models outperform an in-sample mean benchmark. Furthermore, we find that there is a gain from employing sparse principal component analysis in extracting factors and from combining individual model forecasts and survey forecasts.

Second, we work on nowcasting the revised annual GDP (ARNA GDP), which is compiled based on comprehensive annual statistics, but only available after a considerable lag. In nowcasting the revised GDP, we apply several benchmarking methods, including Chow and Lin (1971) and Fernández (1981). We find that some benchmarking models that utilize timely monthly supply-side indicators serve as useful tools for predicting ARNA GDP growth rates.

References

Aastveit, K. A., Gerdrup, K. R., and Jore, A. S. (2011). "Short-term Forecasting of GDP and Inflation in Real-time: Norges Bank’s System for Averaging Models." Norges Bank Staff Memo 9/2011.

Aiolfi, M., and Timmermann, A. (2006). "Persistence in Forecasting Performance and Conditional Combination Strategies." *Journal of Econometrics,* 135(1-2), 31-53.

Aiolfi, M., Capistrán, C., and Timmermann, A. (2011). "Forecast Combination." Ch. 12. *The Oxford Handbook of Economic Forecasting*, Edited by M.P. Clements and D.F. Hendry.

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., and Veronese, G. (2010). "New Eurocoin: Tracking Economic Growth in Real Time." *The Review of Economics and Statistics*, 92(4), 1024-1034.

Andreou, E., Ghysels, E., and Kourtellos, A. (2010). "Regression Models with Mixed Sampling Frequencies." *Journal of Econometrics,* 158(2), 246-261.

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

Anesti, N., Hayes, S., Moreira, A., and Tasker, J. (2017). "Peering into the Present: the Bank's Approach to GDP Nowcasting." *Quarterly Bulletin* 2017 Q2, Bank of England.

Baffigi, A., Golinellib, R., and Parigia, G. (2004). "Bridge Models to Forecast the Euro Area GDP." *International Journal of Forecasting,* 20, 447- 460.

Bańbura, M., and Modugno, M. (2014). "Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics*, 29(1), 133-160.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). "Now-casting and the Real-Time Data Flow." *Handbook of Economic Forecasting*, 2(Part A), 195-237.

Bates, J. and Granger, C. (1969). "The Combination of Forecasts." *Journal of the Operational Research Society,* 20(4), 451-468.

Barbone, L., Bodo, G., and Visco, I. (1981). "Costi e Profitti nell'Industria in Senso

Stretto: un'Analisi su Serie Trimestrali, 1970-1980." *Bolletino della Banca d'Italia*, 467-510.

Bragoli, D. (2017). "Now-casting the Japanese Economy." *International Journal of Forecasting*, 33, 390-402.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). "Model Selection: an Integral Part of Inference." *Biometrics*, 603-618.

Bournay, J., and Laroque. G. (1979). "Réflexions sur la Méthode d'Élaboration des Comptes Trimestriels." *Annales de l'INSÉÉ*, 36:3-30.

Cabinet Office. (2018). "Methodology of Compiling Preliminary Quarterly GDP CY2011 Benchmark Edition."

Chow, G. C., and Lin, A. L. (1971). "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." *Review of Economics and Statistics*, 372-375.

Clements, M. P., and Galvão, A. B. (2008). "Macroeconomic Forecasting with Mixed-frequency Data: Forecasting Output Growth in the United States." *Journal of Business and Economic Statistics*, 26(4), 546-554.

Denton, F. T. (1971). "Adjustment of Monthly or Quarterly Series to Annual Totals: an Approach based on Quadratic Minimization." *Journal of the American Statistical Association*, 66(333), 99-102.

Diebold, F., and Mariano, R. (1995). "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, 13(3), 253-263.

Di Fonzo, T. (1994). "Temporal Disaggregation of a System of Time Series when the Aggregate Is Known: Optimal vs. Adjustment Methods." In Workshop on Quarterly National Accounts, pp 63-77, Paris, Dec. 1994. Eurostat.

Fernández, R. B. (1981). "A Methodological Note on the Estimation of Time Series." *Review of Economics and Statistic*s, 63(3):471-476.

Ferrara, L., and Guérin, P. (2018). "What Are the Macroeconomic Effects of High-frequency Uncertainty Shocks?" *Journal of Applied Econometrics*, 33(5), 662-679.

Foroni, C., and Marcellino, M. (2014). "A Comparison of Mixed Frequency Approaches for Nowcasting Euro Area Macroeconomic Aggregates." *International Journal of*

Forecasting, 30(3), 554-568.

Foroni, C., Marcellino, M. G., and Schumacher, C. (2015). "Unrestricted Mixed Data Sampling (MIDAS): MIDAS Regressions with Unrestricted Lag Polynomials." *Journal of the Royal Statistical Society Series A*, vol. 178(1), 57-82.

Ghysels, E., Kvedaras, V., and Zemlys, V. (2016). "Mixed Frequency Data Sampling Regression Models: The R Package midasr." *Journal of Statistical Software*, 72(i04).

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

Hara, N., and Ichiue, H. (2011). "Real-Time Analysis on Japan's Labor Productivity." *Journal of the Japanese and International Economies*, 25, 107-130.

Hara, N., and Yamane, S. (2013). "New Monthly Estimation Approach for Nowcasting GDP Growth: The Case of Japan." Bank of Japan Working Paper Series, No.13-E-14.

Hendry, D., and Clements, M. (2004). "Pooling of Forecasts." *Econometrics Journal*, 7(1), 1-31.

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

Kim, H. H., and Swanson, N. R. (2017). "Methods for Pastcasting, Nowcasting and Forecasting using Factor-MIDAS: With an Application to Korean GDP." *Journal of Forecasting*, 37(3), 281-301.

Kim, H. H., and Swanson, N. R. (2018). "Mining Big Data using Parsimonious Factor, Machine Learning, Variable Selection and Shrinkage Methods." *International Journal of Forecasting*, 34(2), 339-354.

Kristensen, J. T. (2017). "Diffusion Indexes with Sparse Loadings." *Journal of Business and Economic Statistics*, 35(3), 434-451.

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

Mariano, R. S., and Murasawa, Y. (2003). "A New Coincident Index of Business Cycles based on Monthly and Quarterly Series." *Journal of Applied Econometrics*, 18(4), 427-443.

Miyamoto, W., Nguyen, T. L., and Sergeyev, D. (2018). "Government Spending Multipliers under the Zero Lower Bound: Evidence from Japan." *American Economic Journal: Macroeconomics*, 10(3), 247-77.

Litterman, R. B. (1983). "A Random Walk, Markov Model for the Distribution of Time Series." *Journal of Business and Economic Statistics*, 1(2), 169-173.

Ohyama, S. (2001). "Forecasting Japanese GDP using Forecast Combination." Bank of Japan Working Paper Series, 01-3 (in Japanese).

Sax, C., and Steiner, P. (2013). "Temporal Disaggregation of Time Series." *The R Journal*, 5(2), 80-87.

Schumacher, C. (2016). "A Comparison of MIDAS and Bridge Equations." *International Journal of Forecasting*, 32(2), 257-270.

Stock, J. H., and Watson, M. W. (2002). "Macroeconomic Forecasting using Diffusion Indexes." *Journal of Business and Economic Statistics*, 20(2), 147-162.

Stock, J. H., and Watson, M. W. (2004). "Combination Forecasts of Output Growth in a Seven-country Data Set." *Journal of Forecasting*, 23(6), 405-430.

Timmermann, A. (2006). "Forecast Combinations." *Handbook of Economic Forecasting*, Volume 1, Edited by Graham Elliott, Clive W.J. Granger and Allan Timmermann.

Urasawa, S. (2014). "Real-time GDP Forecasting for Japan: A Dynamics Factor Model Approach." *Journal of the Japanese and International Economies*, 34, 116-134.

Zou, H., Hastie, T., and Tibshirani, R. (2006). "Sparse Principal Component Analysis." *Journal of Computational and Graphical Statistics*, 15(2), 265-286.

Appendix A: Sparse Principal Component Analysis (SPCA)

Stock and Watson (2002) propose a technique to forecast objective variables with factors obtained from a large number of macroeconomic variables by means of principal component analysis. They show that the subsequent forecasts perform well when compared with other models and leading indicators. This appendix provides an explanation of sparse principal component analysis (SPCA), a variant of the principal component approach proposed by Zou et al. (2006).[25]

A standard factor model is represented as follows:

$$\mathbf{X_t} = \mathbf{\Lambda F_t} + \mathbf{e_t} \tag{A1}$$

where $\mathbf{X_t} = \{x_{it}\}$ for $t = 1, \ldots, T$ is a $n \times 1$ vector of observed variables. r is defined as the number of factor. Hence, $\mathbf{\Lambda} = \{\lambda_{ij}\}$ is a $n \times r$ matrix of the factor loadings, $\mathbf{F_t}$ is a $r \times 1$ vector of common latent factors. To simplify the descriptions, we introduce the matrix notations, $\mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_T})'$.

In classical principal component analysis (PCA), as each factor is a linear combination of all variables, some of the loadings may be typically nonzero. Even though the estimated factors allow us to be very parsimonious in the forecasting equations, the factors are by no means parsimonious. Therefore, the traditional PCA can be modified such that the estimated loadings will be sparse, which we will denote a sparse principal component analysis (SPCA).

Zou et al. (2006) employ LASSO and elastic net penalization by developing a regression optimization framework. For the LASSO penalization case, consider first the problem of estimating a single factor. Augmenting the least squares criterion with the penalty terms will give us the following objective function:

$$V^{\mathrm{LASSO}}(\mathbf{F}, \mathbf{\Lambda}; \mathbf{X}, \psi_T) = \frac{1}{nT}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \boldsymbol{\lambda}_\mathbf{i}\mathbf{F_t})^2 + \psi_T \sum_{i=1}^{n}|\boldsymbol{\lambda}_\mathbf{i}|\right]. \tag{A2}$$

[25] This appendix follows the argument of Kristensen (2017).

$\boldsymbol{\lambda}_{\mathbf{i}}$ is the ith row of $\boldsymbol{\Lambda}$, say, a $1 \times r$ vector**.** Note that the functions are written in terms of $\mathbf{F}$ and $\boldsymbol{\lambda}$ in order to make it explicit that we are only estimating single factors. Furthermore, the objective functions now also depend on the tuning parameter ψ_T.[26]

One of the appealing features of SPCA is that the estimated factor will be a linear combination of the observed variables $\mathbf{X}$ just as in the PCA case. For example, in the LASSO case, subsequent factors are given as:

$$\hat{\mathbf{F}}_{\mathbf{1}}^{\mathbf{LASSO}} = \frac{\mathbf{X}\hat{\boldsymbol{\lambda}}_{\mathbf{1}}^{\mathbf{LASSO}}}{n}. \tag{A3}$$

However, the crucial difference is that the loadings will now be sparse, in the sense that some of the entries of $\hat{\boldsymbol{\lambda}}_{\mathbf{1}}^{\mathbf{LASSO}}$ will be zero. Hence, the factor may depend only on some selected variables.

The factors can be estimated in a sequential approach as detailed in the following manner. The SPCA of the first factor and associated loadings are obtained by:

$$\left(\widehat{\mathbf{F}_{\mathbf{1}}}, \hat{\boldsymbol{\lambda}}_{\mathbf{1}}\right) = \underset{\mathbf{F},\boldsymbol{\lambda}}{\operatorname{argmin}}\, V^{\mathrm{LASSO}}(\mathbf{F}, \boldsymbol{\lambda}; \mathbf{X}, \psi_T) \quad s.t. \quad \boldsymbol{\lambda}'\boldsymbol{\lambda}/n = 1.$$

Let the residuals from the estimation of the kth factor be defined as e_k, then for $k > 1$ the subsequent estimates are given as

$$\left(\hat{\mathbf{F}}_{\mathbf{k}}, \hat{\boldsymbol{\lambda}}_{\mathbf{k}}\right) = \underset{\mathbf{F},\boldsymbol{\lambda}}{\operatorname{argmin}}\, V^{\mathrm{LASSO}}(\mathbf{F}, \boldsymbol{\lambda}; \mathbf{e}_{\mathbf{k-1}}, \psi_T) \quad s.t. \quad \boldsymbol{\lambda}'\boldsymbol{\lambda}/n = 1.$$

Hence the SPCA factor estimates of r factors and the loading matrix are given as $\hat{\mathbf{F}} = \left(\hat{\mathbf{F}}_{\mathbf{1}}, \dots, \hat{\mathbf{F}}_{\mathbf{r}}\right)$ and $\hat{\boldsymbol{\Lambda}} = (\hat{\boldsymbol{\lambda}}_{\mathbf{1}}, \dots, \hat{\boldsymbol{\lambda}}_{\mathbf{r}})$.

[26] Following Kristensen (2017), the tuning parameter is set to optimize the Bayesian Information Criterion (BIC) of forecasting models.

Appendix B: Data Description

(1) Series used in the nowcasting models

Name	Frequency	Source	Transformation	Reporting lag (months)	(days)
Hard indicators:					
Index of industrial production	Monthly	METI	sa, m/m change	1	27− 31
Index of tertiary industry activity	Monthly	METI	sa, m/m change	2	40− 51
Current survey of commerce (sales value, wholesale)	Monthly	METI	sa, m/m change	1	27− 30
[Deflator for sales value] producer price index	Monthly	BOJ	-	1	10− 16
Official GDP estimate	Quarterly	CAO	sa, q/q change	-	-
Soft indicators:					
Reuters Tankan DI (manufacturers)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (materials)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (textile and paper)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (chemicals)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (oil refinery and ceramics)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (steel and metals)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (manufactured products)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (food)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (metal and machinery)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (electric machinery)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (transport equipment)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (precision machinery)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (non-manufacturers)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (real estate)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (wholesalers)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (retailers)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (transport and utility)	Monthly	TR	level	0	-24− -7
Reuters Tankan DI (other services)	Monthly	TR	level	0	-24− -7

(2) Series used only for extrapolation

Name	Frequency	Source	Transformation	Reporting lag (months)	(days)
Survey of Production Forecast	Monthly	METI	sa, m/m change	0	-3− 0
Economy watchers survey (DI for current conditions)	Monthly	CAO	level	1	8− 12

(3) Series used in the benchmarking models

Name	Frequency	Source	Transformation	Reporting lag (months)	(days)
Index of all industry activity	Quarterly	METI	sa, level	2	50− 56
Index of industrial production	Quarterly	METI	sa, level	1	27− 31
Index of tertiary industry activity	Quarterly	METI	sa, level	2	40− 51
Index of construction industry activity	Quarterly	METI	sa, level	2	50− 56
Official GDP estimate (Annual Report on National Accounts)	Annual	CAO	level	-	-

Notes:
1. METI: Ministry of Economy, Trade and Industry; CAO: Cabinet Office; TR: Thomson Reuters; BOJ: Bank of Japan.
2. The survey of production forecast figure is calculated based on METI's adjusted value.
3. Reporting lag indicates the lag between month-end (quarter-end) dates and data release dates.

Appendix C: Extrapolation in bridge equation approaches

The monthly indicators used for nowcasting preliminary quarterly estimates have different publication lags. As discussed in Section 3, some indicators have to be extrapolated using bridge model approaches. First, the IIP is extrapolated simply by using the result of the Survey of Production Forecast (SPF), conducted by Ministry of Economy, Trade and Industry (see Appendix B).[27] Second, we extrapolate the Current Survey of Commerce (CSC) using the following auxiliary equation

$$\mathrm{dlog}\left(CSC_{t,m}\right) = \beta_0 + \beta_1 \mathrm{dlog}\left(SPF_{t,m}\right) + \varepsilon_{t,m}, \tag{C1}$$

where $CSC_{t,m}$ denotes the CSC at mth month of the quarter t, $SPF_{t,m}$ denotes the SPF at mth month of the quarter t.[28]

Thirdly, we extrapolate the Index of Tertiary Activities (ITA). In doing so, we utilize the Economic Watcher Survey (EWS), which is available prior to the release of the ITA (see Appendix B). The auxiliary equation reads as:

$$\begin{gathered}\mathrm{dlog}\left(ITA_{t,m}\right) = \beta_0 + \beta_1 \mathrm{dlog}\left(SPF_{t,m}\right) + \beta_2 \mathrm{dlog}\left(\widehat{CSC}_{t,m}\right) \\ +\beta_3 EWS_{t,m} + \varepsilon_{t,m},\end{gathered} \tag{C2}$$

where $ITA_{t,m}$ denotes the ITA at mth month of the quarter t, $\widehat{CSC}_{t,m}$ denotes fitted value of the CSC in (C1) for the mth month of the quarter t, and $EWS_{t,m}$ denotes the EWS at mth month of the quarter t.[29]

[27] The SPF is the survey conducted for the purpose of projecting near-term IIP. The survey covers about 760 firms and around 80 percent of total production. To avoid bias of the survey, we use the adjusted value of the SPF, which is also published by Ministry of Economy, Trade and Industry.

[28] Since the sample size of the SPF is relatively small, equations (C1) and (C2) are estimated with actual values of IIP.

[29] The CSC is released prior to the release of the ITA (see Appendix B). When the CSC is available, the fitted value of the CSC is replaced by actual value of the CSC.

Appendix D: Approximation of the quarter-on-quarter growth rate using monthly growth rates

This appendix discusses the method used to obtain quarter-on-quarter growth rates from monthly growth rates in bridge equation models. Mariano and Murasawa (2003) propose approximating quarter-on-quarter growth rates of GDP using three-month growth rates of each month in the quarter. Let q_t represent the quarter-on-quarter growth rates at quarter t. Then, q_t can be approximated as:

$$q_t = \frac{1}{3}\left(m_{t,1}^{(3)} + m_{t,2}^{(3)} + m_{t,3}^{(3)}\right), \tag{D1}$$

where $m_{t,i}^{(3)}$ is three-month growth rate at ith month of quarter t, which can be expressed as the sum of $m_{t,j}$, monthly growth rates at jth month of quarter t

$$\begin{aligned} m_{t,1}^{(3)} &= m_{t-1,2} + m_{t-1,3} + m_{t,1} \\ m_{t,2}^{(3)} &= m_{t-1,3} + m_{t,1} + m_{t,2} \\ m_{t,3}^{(3)} &= m_{t,1} + m_{t,2} + m_{t,3}. \end{aligned} \tag{D2}$$

At two months prior to GDP release dates, where information on $m_{t,3}$ is not available, we simply assume that the monthly GDP level for third month is the average of the values at the first and second months:

$$m_{t,3} = \frac{\frac{M_{t,1} + M_{t,2}}{2} - M_{t,2}}{M_{t,2}} = \frac{1}{2}\left(\frac{M_{t,1}}{M_{t,2}} - 1\right) = \frac{1}{2}\left(\frac{1}{1 + m_{t,2}} - 1\right) \approx -\frac{1}{2}m_{t,2}, \tag{D3}$$

where $M_{t,k}$ denotes the monthly level of GDP for the kth month of quarter t.

Thus, at two months prior to GDP release dates, the quarter-on-quarter growth rates can be expressed as

$$q_t = \frac{1}{3}\left(m_{t,1}^{(3)} + m_{t,2}^{(3)} + m_{t,3}^{(3)}\right)$$

$$= \frac{1}{3}\left\{\left(m_{t-1,2} + m_{t-1,3} + m_{t,1}\right) + \left(m_{t-1,3} + m_{t,1} + m_{t,2}\right)\right.$$

$$\left. + \left(m_{t,1} + m_{t,2} - \frac{1}{2}m_{t,2}\right)\right\} \tag{D4}$$

$$= \frac{1}{3}\left(m_{t-1,2} + 2m_{t-1,3} + 3m_{t,1} + \frac{3}{2}m_{t,2}\right).$$

Similarly, at one month prior to GDP release dates, where all values for the quarter are available:

$$q_t = \frac{1}{3}\left(m_{t,1}^{(3)} + m_{t,2}^{(3)} + m_{t,3}^{(3)}\right)$$

$$= \frac{1}{3}\left\{\left(m_{t-1,2} + m_{t-1,3} + m_{t,1}\right) + \left(m_{t-1,3} + m_{t,1} + m_{t,2}\right)\right. \tag{D5}$$

$$\left. + \left(m_{t,1} + m_{t,2} + m_{t,3}\right)\right\}$$

$$= \frac{1}{3}\left(m_{t-1,2} + 2m_{t-1,3} + 3m_{t,1} + 2m_{t,2} + m_{t,3}\right).$$

Table 1: Bridge Model RMSE

Model	Predictors	RMSE 2-month prior to GDP release dates	RMSE 1-month prior to GDP release date
Bridge	IIP , ITA	0.470	0.449*
	IIP, CSC	0.536	0.520
	IIP	0.570	0.564
	ITA	0.610	0.494*
	CSC	0.596	0.485
Bridge with PCA	IIP, ITA, PC1, PC2	0.473	0.455*
	IIP, CSC, PC1, PC2	0.544	0.533
	IIP, ITA, PC1	0.467	0.443*
	IIP, CSC, PC1	0.539	0.521
	IIP, PC1, PC2	0.588	0.587
	ITA, PC1, PC2	0.703	0.585
	CSC, PC1, PC2	0.613	0.493
	IIP, PC1	0.575	0.571
	ITA, PC1	0.669	0.542*
	CSC, PC1	0.613	0.492
	PC1, PC2	0.807	0.819
	PC1	0.710	0.712
Bridge with SPCA	IIP, ITA, SPC1, SPC2	0.469	0.453*
	IIP, CSC, SPC1, SPC2	0.547	0.533
	IIP, ITA, SPC1	0.464	0.439*
	IIP, CSC, SPC1	0.538	0.520
	IIP, SPC1, SPC2	0.592	0.590
	ITA, SPC1, SPC2	0.660	0.534*
	CSC, SPC1, SPC2	0.613	0.493
	IIP, SPC1	0.574	0.569
	ITA, SPC1	0.652	0.526*
	CSC, SPC1	0.612	0.489
	SPC1, SPC2	0.801	0.812
	SPC1	0.713	0.714
In-sample mean (Benchmark)		0.684	0.684

Note: The shading indicates the lowest RMSE. PC1 and PC2 stand for first and second factor of the surveys extracted with the principal component approach. SPC1 and SPC2 stand for first and second factor of the surveys extracted with sparse principal component analysis. Asterisks indicate statistical significance of the Diebold-Mariano test compared with the benchmark model (in-sample mean).
* Denotes significance at the 10% level.

Table 2: MIDAS Model RMSE

Model	Predictors	RMSE 2-month prior to GDP release dates	 1-month prior to GDP release dates
MIDAS	IIP, ITA, CSC	0.554	0.501
	IIP, ITA	0.571	0.574**
	IIP, CSC	0.512	0.581
	ITA, CSC	0.707	0.814
	IIP	0.513*	0.571*
	ITA	0.486	0.411*
	CSC	0.636	0.548
Factor MIDAS with PCA	IIP, ITA, PC1	0.627	0.596
	IIP, CSC, PC1	0.517	0.570
	ITA, CSC, PC1	0.625	0.723
	IIP, PC1	0.529	0.578
	ITA, PC1	0.500*	0.405*
	CSC, PC1	0.498	0.465
	PC1	0.591	0.637
Factor MIDAS with SPCA	IIP, ITA, SPC1	0.636	0.597
	IIP, CSC, SPC1	0.498	0.557
	ITA, CSC, SPC1	0.544	0.703
	IIP, SPC1	0.524	0.565
	ITA, SPC1	0.500*	0.405*
	CSC, SPC1	0.465	0.480
	SPC1	0.591	0.637
In-sample mean (Benchmark)		0.684	0.684

Note: The shading indicates the lowest RMSE. PC1 stands for the first factor of the surveys extracted with principal component analysis. SPC1 stands for the first factor of the surveys extracted with sparse principal component analysis. Asterisks indicate statistical significance of the Diebold-Mariano test compared with the benchmark model (in-sample mean).

* Denotes significance at the 10% level.

** Denotes significance at the 5% level.

Table 3: Combined MIDAS RMSE

		RMSE	
		2-month prior to GDP release dates	1-month prior to GDP release dates
with PCA	Simple Average	0.426*	0.412*
	Triangular Kernel Approach	0.462*	0.495*
	Inverse Mean Squared Error	0.433*	0.430*
	Discounted Mean Squared Forecast Errors (δ=0.95)	0.433*	0.428*
	Bayesian Model Averaging	0.512*	0.569
with SPCA	Simple Average	0.414*	0.406*
	Triangular Kernel Approach	0.455*	0.485*
	Inverse Mean Squared Error	0.421*	0.425*
	Discounted Mean Squared Forecast Errors (δ=0.95)	0.421*	0.423*
	Bayesian Model Averaging	0.512*	0.569

Note: The shading indicates the lowest RMSE. Asterisks indicate statistical significance of the Diebold-Mariano test compared with the benchmark model (in-sample mean). "With PCA" indicates the combination of the MIDAS models and the Factor MIDAS models estimated with principal component analysis listed in Table2. "With SPCA" indicates the combination of the MIDAS models and the MIDAS models estimated with sparse principal component analysis listed in Table 2.

* Denotes significance at the 10% level.

Table 4: Forecast Combination RMSE

	Simple Average		Triangular Kernel Approach		Inverse Mean Squared Error		Discounted Mean Squared Errors	
	2-month prior to GDP release dates	1-month prior to GDP release dates	2-month prior to GDP release dates	1-month prior to GDP release dates	2-month prior to GDP release dates	1-month prior to GDP release dates	2-month prior to GDP release dates	1-month prior to GDP release dates
Bridge, C-MIDAS	0.413*	0.410*	0.424*	0.417*	0.426*	0.417*	0.424*	0.416*
Bridge, JCER	0.432*	0.425*	0.433*	0.423*	0.432*	0.423*	0.432*	0.423*
Bridge, Bloomberg	0.456*	0.434*	0.451*	0.430*	0.451*	0.430*	0.451*	0.430*
C-MIDAS, JCER	0.429*	0.393*	0.445*	0.409*	0.435*	0.398*	0.435*	0.397*
C-MIDAS, Bloomberg	0.444*	0.406*	0.464*	0.397*	0.447*	0.404*	0.446*	0.404*
Bridge, C-MIDAS, JCER	0.414*	0.400*	0.420*	0.410*	0.418*	0.406*	0.417*	0.406*
Bridge, C-MIDAS, Bloomberg	0.428*	0.408*	0.434*	0.412*	0.430*	0.412*	0.430*	0.411*
Bridge, JCER, Bloomberg	0.453*	0.437*	0.439*	0.427*	0.443*	0.429*	0.444*	0.430*
C-MIDAS, JCER, Bloomberg	0.452*	0.415*	0.467	0.420*	0.456*	0.417*	0.455*	0.416*
Bridge, C-MIDAS, JCER, Bloomberg	0.432*	0.411*	0.430*	0.413*	0.430*	0.412*	0.430*	0.412*
Bridge	0.464	0.439*						
C-MIDAS	0.414*	0.406*						
JCER	0.489	0.467						
Bloomberg	0.515	0.480						

Note: The left column indicates forecasts included in forecast combination. The shading indicates the lowest RMSE. Asterisks indicate statistical significance of the Diebold-Mariano test compared with the benchmark model (in-sample mean). "C-MIDAS" indicates the combination of the MIDAS models and the Factor MIDAS models with SPCA listed in Table 2 weighted with simple average.

* Denotes significance at the 10% level.

Table 5: Benchmarking Model RMSE

	Predictors	i. Ordinary Least Square (OLS)		ii. Chow-Lin method with maximum likelihood			iii. Chow-Lin method with minimization of residual		
		1st year	2nd year	1st year	2nd year	ρ	1st year	2nd year	ρ
dlog(GDP)	dlog4(IAA)	0.483**	0.475**	0.483**	0.475**	0	0.485***	0.440***	0.599
dlog(GDP)	dlog4(IIP), dlog4(ITA)	0.534***	0.529***	0.535***	0.529***	0	0.559**	0.517**	0.603
dlog(GDP)	dlog4(IIP), dlog4(ITA), dlog4(ICA)	0.691	0.804	0.691	0.804	0	0.746	0.771	0.601
GDP	IAA	4.750	1.910	0.958	1.045	0.998	2.325	2.142	0.934
GDP	IIP, ITA	2.322	0.716	1.007	0.886	0.944	1.391	1.130	0.890
GDP	IIP, ITA, ICA	1.987	0.992	1.068	1.046	0.992	1.353	1.145	0.879
GDP	IAA, trend	0.657*	0.488**	0.537**	0.464***	0.747	0.497***	0.449***	0.719
GDP	IIP, ITA, trend	0.932	0.588**	0.687	0.653	0.873	0.698	0.647	0.777
GDP	IIP, ITA, ICA, trend	0.752	0.712	0.702	0.699	0.557	0.658*	0.679	0.665

	Predictors	iv. Litterman method			v. Fernández method		vi. Proportional Denton method	
		1st year	2nd year	α	1st year	2nd year	1st year	2nd year
dlog(GDP)	dlog4(IAA)							
dlog(GDP)	dlog4(IIP), dlog4(ITA)							
dlog(GDP)	dlog4(IIP), dlog4(ITA), dlog4(ICA)							
GDP	IAA	0.755	0.831	0.512	0.767	0.834	0.902	0.981
GDP	IIP, ITA	0.729	0.765	0.373	0.728	0.764		
GDP	IIP, ITA, ICA	0.773	0.835	0.534	0.779	0.846		
GDP	IAA, trend	0.464***	0.438***	0	0.464***	0.438***		
GDP	IIP, ITA, trend	0.583**	0.571**	0	0.583**	0.571**		
GDP	IIP, ITA, ICA, trend	0.692*	0.713	0	0.692*	0.713		
	QE1	0.833	0.838					
	QE2	0.840	0.823					

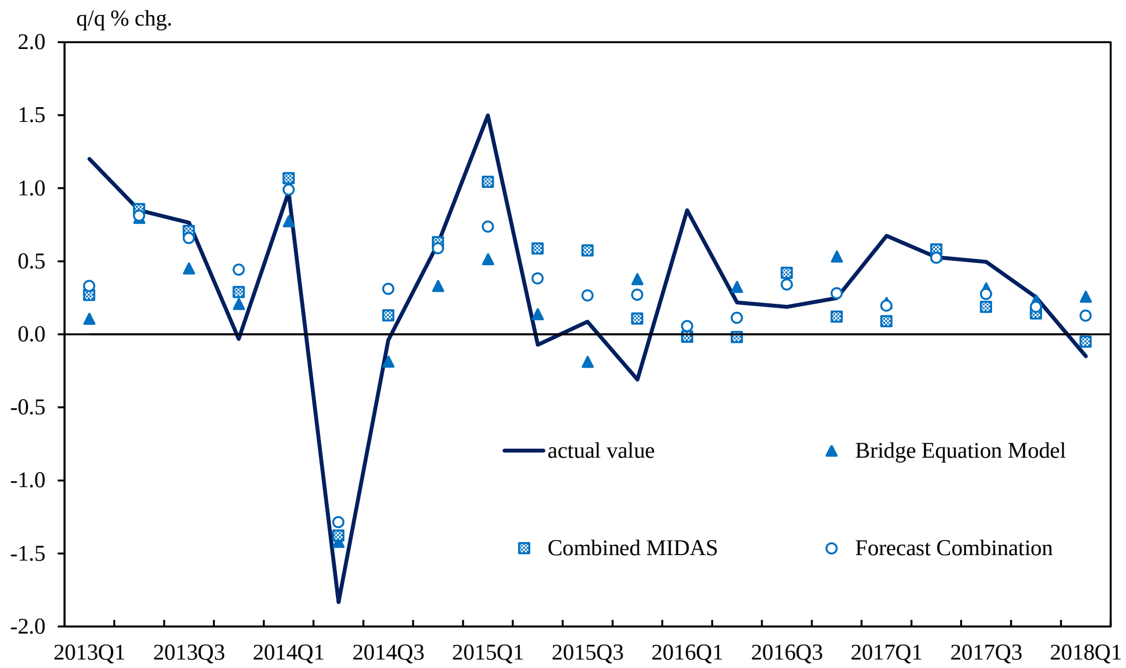
Note: The shading indicates the lowest RMSE, and dlog4 stands for four-quarter log-difference growth rates. RMSE is computed in terms of growth rates. Asterisks indicate the statistical significance of the Diebold-Mariano test compared with the QE2.

* Denotes significance at the 10% level.

** Denotes significance at the 5% level.

*** Denotes significance at the 1% level.

Figure 1: Forecast combination performances (2-month prior to GDP release dates)

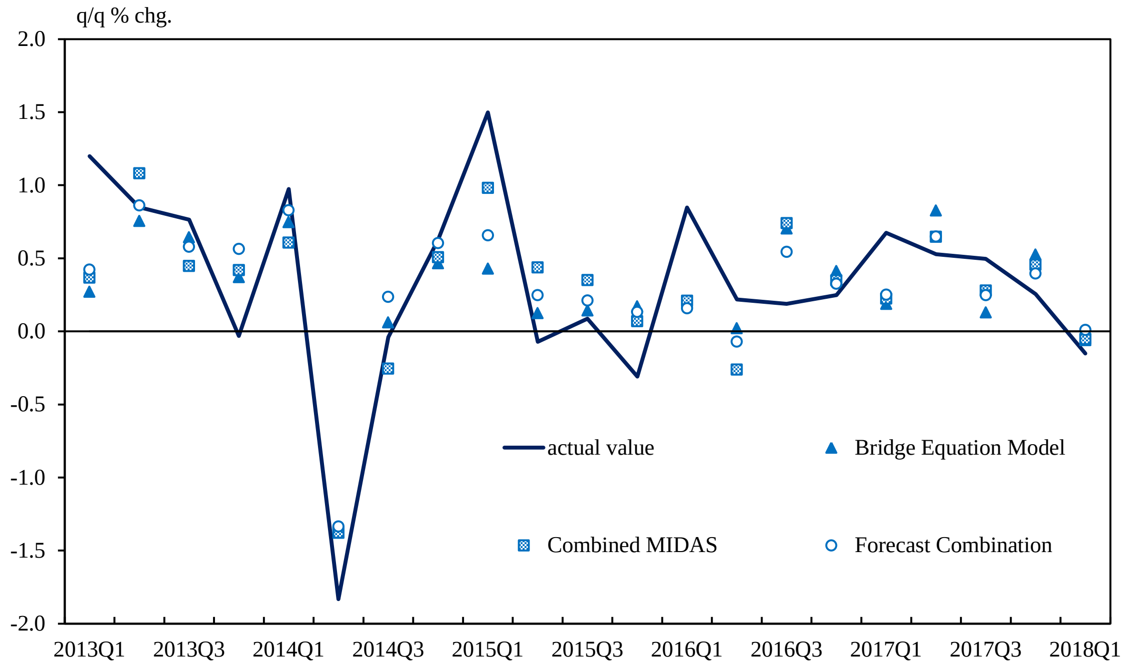


Note: "Combined MIDAS" indicates combination of MIDAS models and Factor MIDAS models with SPCA listed in Table 2 weighted with simple average.

"Bridge Equation Model" indicates Bridge Equation Model with IIP, ITA and first factor extracted with sparse principal component analysis (SPC1).

"Forecast Combination" indicates the combination of Combined MIDAS, Bridge Equation Model and JCER weighted with simple average.

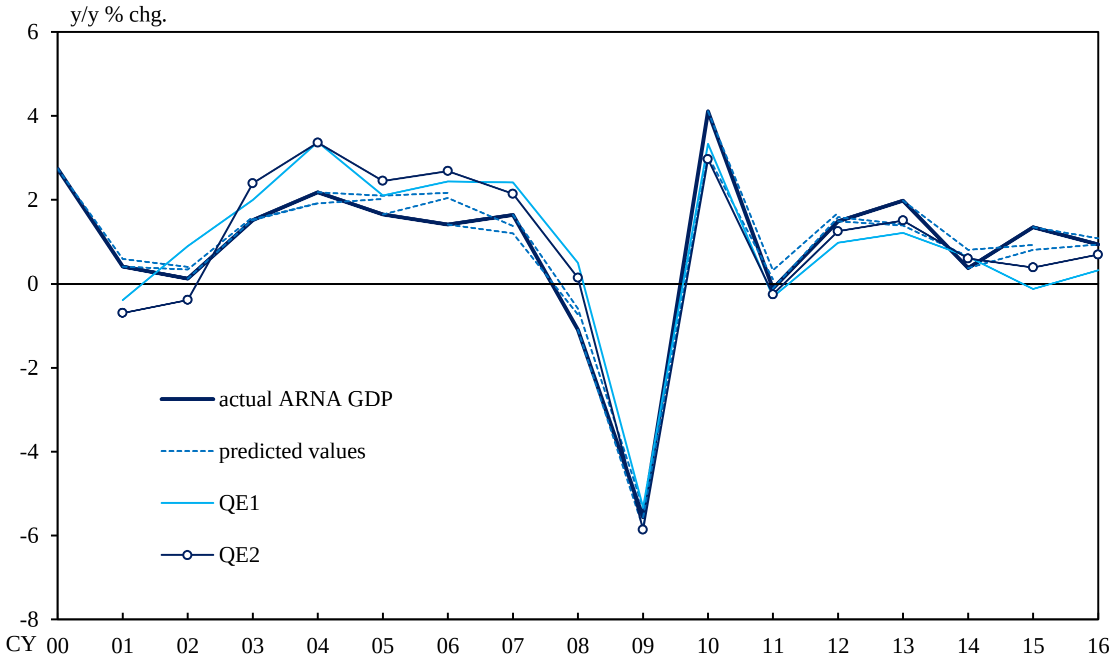
Figure 2: Forecast combination performances (1-month prior to GDP release dates)



Note: “Combined MIDAS” indicates combination of MIDAS models and Factor MIDAS models with SPCA listed in Table 2 weighted with simple average. “Bridge Equation Model” indicates the Bridge Equation Model using IIP, ITA and first factor extracted with sparse principal component analysis (SPC1).

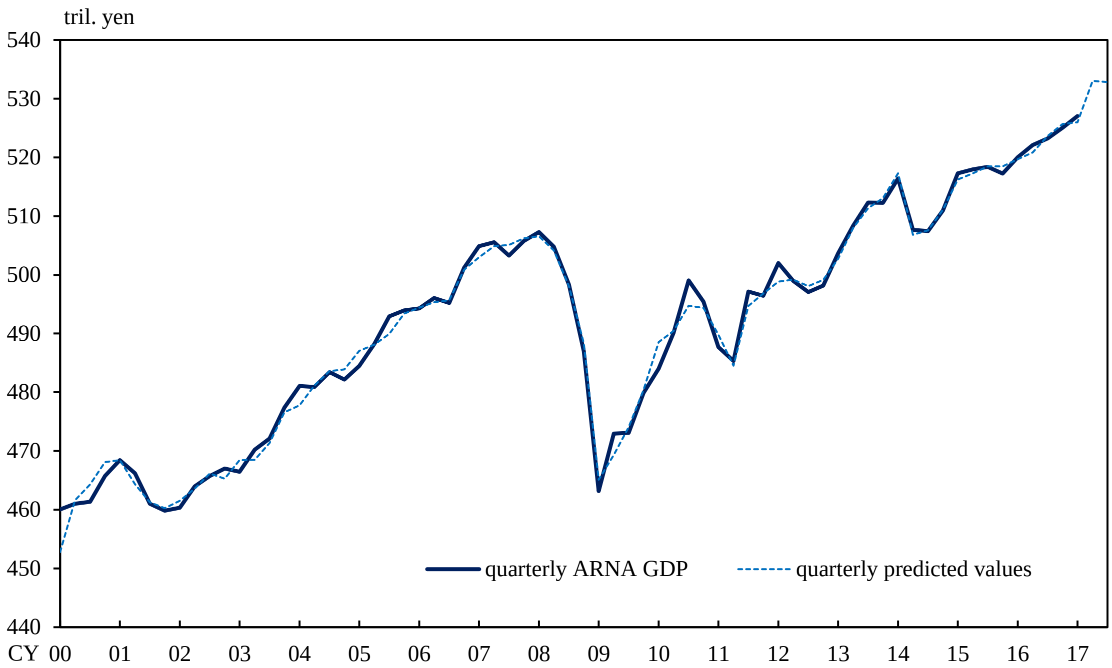
“Forecast Combination” indicates the combination of Combined MIDAS, Bridge Equation Model and JCER weighted with simple average.

Figure 3: Benchmarking model out-of-sample forecasts (year-on-year growth rates)



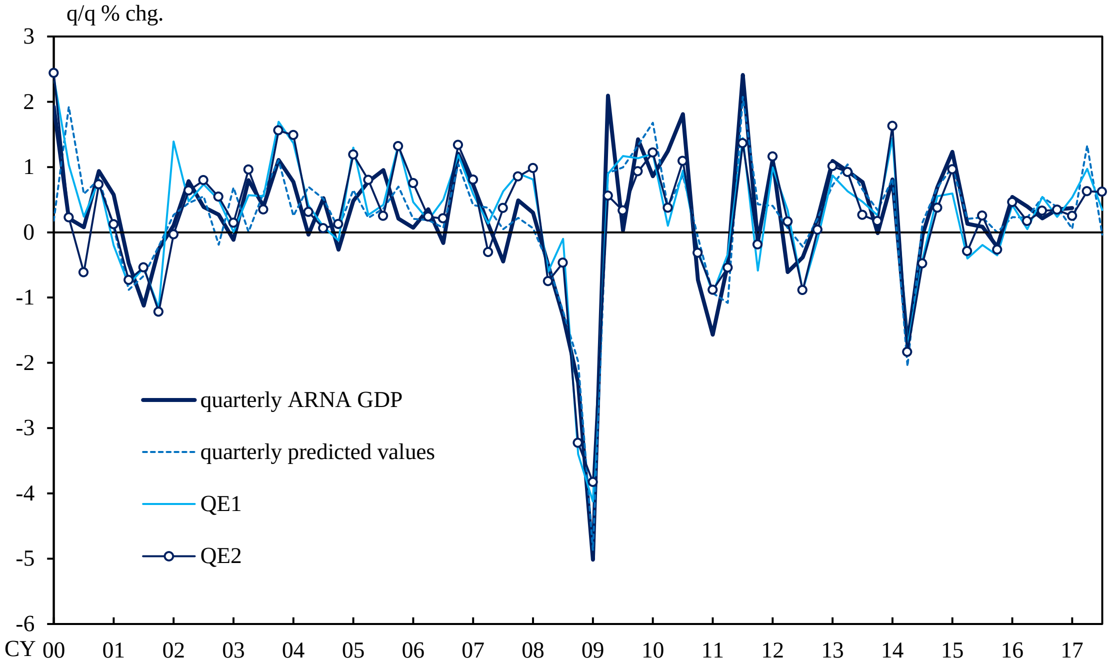
Note: "Predicted values" indicates the out-of-sample forecasts produced by the Fernández method with the preliminary variable of the IAA and time trend.

Figure 4: Benchmarking model quarterly fitted values (level)



Note: “Quarterly predicted values” indicates the fitted values produced by the Fernández method with the preliminary variable of the IAA and time trend. The fitted values from 2000 to 2016 are in-sample values, and the fitted values from 2017 are out-of-sample values.

Figure 5: Benchmarking model quarterly fitted values (quarter-on-quarter growth rates)



Note: "Quarterly predicted values" indicates the fitted values produced by the Fernández method with the preliminary variable of the IAA and time trend. The fitted values from 2000 to 2016 are in-sample values, and the fitted values from 2017 are out-of-sample values.