Nowcasting Japan’s GDP

Fumio Hayashi\textsuperscript{1} · Yuta Tachi\textsuperscript{2}

Received: 16 November 2021 / Accepted: 20 August 2022 / Published online: 20 September 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
This paper backtests a nowcast of Japan’s real GDP growth. It has three contributions: (i) use of genuine real-time data, (ii) implementation of a new method for the revision analysis that relates the revision of the nowcast to not only new observations but also data revisions, and (iii) a benchmarking of the nowcast to a market consensus forecast at monthly forecasting horizons. Our nowcast’s forecast accuracy is comparable to that of the consensus at most, but not all, monthly horizons. Our revision analysis of the March 2011 earthquake finds the nowcast reacting to a steep post-quake decline in car production. In contrast, the consensus hardly budged, most likely because the decline was correctly viewed as temporary. The onset of COVID-19 triggers the consensus to take a precipitous descent. The nowcast, despite timely red flags from “soft” (i.e., survey-based) indicators, does not respond immediately in full, because it took a month or more for “hard” (i.e., non-survey-based) indicators to register sharply reduced economic activities.

Keywords Nowcasting · Real-time data · Dynamic factor models · Revision analysis

JEL Classification E17 · E37 · C53

1 Introduction and summary

Many key economic variables, most notably GDP, are published with considerable delay. Inferring their current value from other indicator variables that are more fre-
quently or more promptly observed is called nowcasting. This paper backtests a nowcast of Japan’s quarterly real GDP growth.

The literature on nowcasting for the USA, the UK, and the Euro-area countries has been surveyed by Bânbura et al. (2013) and, more recently, by Bok et al. (2018). The standard statistical model employed in the literature is the dynamic factor model. The model is particularly suitable for nowcasting because, when cast in a state-space form, it provides a natural way to handle data with mixed frequencies. Our GDP nowcast utilizes the model of Mariano and Murasawa (2003), a dynamic factor model designed for handling monthly and quarterly frequencies.

The paper has three contributions. As forcefully argued by Croushore (2011), the real-time nature of the backtest requires us to collect all past vintages of each indicator variable, which is a major undertaking because coverage of Japan by available off-the-shelf vintage sources is quite limited. The paper’s first contribution, relative to the literature on nowcasting Japan’s GDP (to be reviewed in Sect. 2), is that it is the first real-time backtest for Japan with a standard set of monthly and quarterly indicators, 19 in number, typically considered in the literature for other countries.

The paper’s second contribution is that it implements an extension, recently developed by Hayashi and Tachi (2021), of the revision analysis of how a nowcast for a fixed target date evolves in time in response to various data releases and changes in the model parameters. A data release typically entails not only new observations of the indicator variable but also data revisions. In the existing revision analysis, only the effect of new observations can be broken down into contributions from individual variables. The new method allows us to do the breakdown for data revisions and parameter changes as well. This revision analysis is applied to three episodes of disasters, the Lehman crisis, the 2011 earthquake, and the COVID-19 voluntary lockdown of 2020.

The paper’s third contribution is a benchmarking of the nowcast to a market consensus forecast, which has not been done in previous studies for Japan. The consensus forecast paired with nowcasts in most US studies is quarterly, while our consensus forecast...
forecast is monthly. Since our consensus growth forecast is for the current and subsequent quarters, it is possible to compare our nowcast of a given quarterly growth with the consensus at monthly forecasting horizons. The horizon ranges from −1 month when forecasts are as of 1 month after the quarter but before the “flash” (the first official estimate) released about 6 weeks after the quarter, to 11 months before the quarter’s end.

A summary of the paper’s results is as follows. Our nowcast’s RMSE (root mean squared error) is comparable to that of the consensus at most, but not all, monthly horizons. Our revision analysis of the Lehman crisis finds the GDP nowcast to be revised down substantially when a string of negative data releases brings about parameter changes. For the March 2011 earthquake, our nowcast of the 2011:Q1 to Q2 growth reacts to a steep decline in car production following the quake. In contrast, the consensus hardly budged, most likely because the decline was correctly viewed as temporary. The onset of COVID-19 triggers the consensus to take a precipitous descent. The nowcast, despite timely red flags from “soft” (i.e., survey-based) indicators, does not respond immediately in full, because it took a month or more for “hard” (i.e., non-survey-based) indicators to register sharply reduced economic activities. In all these episodes, the data-revisions effect cannot be ignored when compared to the new-observations effect for indicator variables with data revisions.

The organization of the paper is as follows. The next section is a brief survey of the literature on nowcasting Japan’s GDP. Section 3 quickly reviews the language of real-time data by an example of just two indicator variables. Section 4 reviews the Mariano and Murasawa (2003) model. Section 5 describes the real-time database that we constructed from primary sources. It is followed by an explanation of the structure of the monthly multi-horizon consensus forecast in Sect. 6. Sections 7 and 8 report our results on the multi-horizon RMSE comparison and on the revision analysis of the three disasters. Section 9 consists of a brief conclusion and an agenda for future research.

2 Brief review on nowcasting Japan’s GDP

There exist only a handful of papers on nowcasting Japan’s real GDP. There are two studies, Hara and Yamane (2013) and Chikamatsu et al. (2021), that use the bridge equation/MIDAS (mixed-data sampling) approach. The former does not conduct a backtest because forecast error calculation is in-sample, based on parameter values estimated on the full sample. The latter is not a real-time backtest either because, although the forecast error calculation is out-of-sample, the data are only “pseudo”

---

5 A widely used consensus forecast for US GDP is the SPF (Survey of Professional Forecasts), a quarterly survey compiled and made publicly available by the Federal Reserve Bank of Philadelphia. Nowcast studies that use the SPF include Giannone et al. (2008), Liebermann (2014), and Bok et al. (2018). There are monthly surveys, notably one by Blue Chip and Consensus Economics. Nowcast studies utilizing monthly surveys are relatively scarce, perhaps because those surveys are proprietary. A rare exception is Higgins (2014) which uses the Blue Chip survey.

6 Essentially, this approach estimates monthly components of quarterly real GDP growth from monthly indicator variables via a relation that can be inferred from a regression of real GDP growth on those indicator variables aggregated to quarterly frequency.
real-time in that past vintages are assumed to be equal to the most recent vintage. This assumption is appropriate for “soft” (survey-based) indicators, but not for “hard” (non-survey-based) indicators such as GDP and IIP (the index of industrial production) whose initial values get revised in subsequent data releases.

More relevant to us are Urasawa (2014) and Bragoli (2017), both of which, like our paper, use mixed-frequency dynamic factor models. Their real-time backtests are constrained due to the limited coverage of Japan by available off-the-shelf vintage sources. The backtesting period in Urasawa (2014), 2012:Q2-2013:Q2, is very short, due most likely to the difficulty of obtaining pre-2012 vintages. The set of indicator variables consists of only six variables including real GDP.

The backtesting period in Bragoli (2017) is 2005–2015. The set of indicators for her “Model A” has 18 variables, but her model A backtest is only pseudo real-time. The variable set for her genuine real-time backtest has eleven variables, but only five of them are hard indicators. Past vintages of those variables, which she obtains from the OECD real-time database, misses a number of data releases because at most one data release per month is recorded in that database. Her benchmark forecast is not a consensus forecast by market economists; it is a model-based quarterly forecast by a private research institute.

3 The language of real-time data

A **vintage** of a variable is the time-series provided in a data release. The **revision matrix** of the variable is a matrix whose columns are vintages of the variable. To illustrate, Table 1 has the revision matrix for Japan’s GDP growth. Take, for example, the time series in the column labeled “2009/5/20.” It is the vintage released on that day. The reference point indicates the end-of-quarter month. For example, the March 2009 entry of $-16.49$ is the annualized quarterly GDP growth in percent from 2008:Q4 to 2009:Q1 (whose end-of-quarter month is March 2009). By construction, each row collects growth estimates for the same period from different data releases. Its first or left-most non-empty element is the “flash” (first official) estimate, indicated in italics. For example, the flash estimate of the 2008:Q4 to 2009:Q1 growth is $-16.49\%$ from the 2009/5/20 vintage.

Three features of the table are noteworthy. First, Japan’s GDP estimates are revised substantially, particularly in the first two years after the initial release.\(^7\) Second, it takes about six weeks (as opposed to four for the USA) for the flash to be released. Third, the COVID-19 voluntary lockdown in the first half of 2020 has produced a jaw-dropping flash estimate, released on 2020/8/17, of $-32.56\%$ per year for June 2020.\(^8\)

---

\(^7\) See, for example, Figure 3 of an OECD study by Zwijnenburg (2015). The revision size of quarterly real GDP growth, measured by the mean absolute revision to the first published estimate, for Japan is about twice that for the USA.

\(^8\) The flash estimate released by the Cabinet Office, which is $-27.79\%$ per year, looks less extreme because the annualization by the Cabinet Office is by the formula $100 \times \left(\frac{Y_t - 3}{Y_{t-1}}\right)^4 - 1$ where $Y_t$ is real GDP for the quarter whose end-of-quarter month is $t$. 
### Table 1: Revision matrix for Japan’s GDP growth

| Reference point     | Vintages (data releases) |        |        |        |        |        |
|---------------------|--------------------------|--------|--------|--------|--------|--------|
|                     |                          | 2009/5/20 | 2011/5/19 | 2020/6/8 | 2020/8/3 | 2020/8/17 |
| March 1990          |                         | 0.61    | -1.91  | -2.85  | -2.85  | -2.85  |
| June 1990           |                         | 9.95    | 12.42  | 10.45  | 10.45  | 10.45  |
| March 2009          | -16.49                   |        | -20.20 | -19.61 | -19.61 | -19.62 |
| March 2020          |                          |        |        | -2.27  | -2.24  |        |
| June 2020           |                          |        |        |        | -32.56 |        |

Annualized quarterly growth rate of seasonally adjusted real GDP in percent. That is, if $t$ is the month indicated by the reference point, the GDP growth rate for $t$ equals $400 \times [\log(Y_t) - \log(Y_{t-3})]$ where $Y_t$ is the level of real GDP for the quarter whose end-of-quarter month is $t$. The “flash” (i.e., the first preliminary) estimates, which are released about 6 weeks after the conclusion of the quarter, are in italics.
Table 2  Two samples from real-time data

| Sample as of | July 6, 2020 | August 7, 2020 |
|--------------|--------------|---------------|
|              | IIP          | GDP           | IIP          | GDP           |
|              | Vintage      | 2020/6/30     | 2020/6/8     | 2020/7/31     | 2020/8/3      |
| January 1990 | −0.87        | −0.87         | −0.87        | −0.87         |
| February 1990| 0.87         | 0.87          | 0.87         | 0.87          |
| March 1990   | 1.90         | −2.85         | 1.90         | −2.85         |
| February 2020| −0.30        |               | −0.30        |               |
| March 2020   | −3.79        | −2.27         | −3.79        | −2.24         |
| April 2020   | −10.33       |               | −10.33       |               |
| May 2020     | −8.83        |               | −9.33        |               |
| June 2020    | 2.63         |               |              |               |

Growth rates in percent. The GDP growth rate is an annualized quarterly growth rate, defined as 400 times the quarterly difference in logs. The IIP growth rate is the monthly growth rate, defined as 100 times the monthly difference in logs.

A real-time database is a collection of revision matrices, one for each variable in the set of indicator variables. The sample as of the vantage point v is a collection of latest vintages available as of date v for the set of variables. The sample may be of mixed frequencies. To illustrate, Table 2 shows two samples for just two indicator variables: IIP (the index of industrial production), which is monthly, and GDP, which is quarterly, for Japan.

Regarding the table’s first sample, which is as of the vantage point of July 6, 2020, and which is shortly after the end of 2020:Q2, the latest vintage of IIP is imported from the revision matrix for IIP (the IIP revision matrix not shown here). It is from a data release on 2020/6/30. For GDP, as seen from Table 1, the latest release as of July 6, 2020 is on 2020/6/8. Because GDP is quarterly, values are missing for months that are not the last month of a quarter. The second sample is as of August 7, 2020, which is about 1 month after the end of 2020:Q2. For GDP, the latest vintage is 2020/8/3. For IIP, it is 2020/7/31. The second sample differs from the first sample in two respects. First, there is a new observation: 2.63% for IIP growth for June 2020 (from May to June 2020). Second, a data revision occurs: for IIP, the May 2020 value was revised from −8.83% in the 2020/6/30 vintage to −9.33% in the 2020/7/31 vintage; for GDP, the March 2020 value was revised from −2.27% to −2.24%.

To construct a real-time database from which the sample can be derived, one needs to collect all past vintages. An alternative is pseudo real-time data, which obtains if values in all vintages are assumed to be equal to those in the latest vintage. For example, in Table 1, for each row, all values are set equal to the value in the right-most column. Therefore, provided that all past data release dates are known, a pseudo real-time database can be constructed from a single vintage. However, as forcefully

---

9 Normally, there should be no data release after the June release until the mid-August. For 2020, there was an extra data release on August 3.
argued by Croushore (2011), use of pseudo real-time data is inappropriate for real-time backtesting because the forecast is given the benefit of hindsight. This is particularly so if the forecast is to be benchmarked to market forecasts.

4 The Mariano–Murasawa model

This section reviews the mixed-frequency dynamic factor model of Mariano and Murasawa (2003). We start out with a refresher on state-space models because a dynamic factor model is a particular state-space model with a small number of state variables.

4.1 The canonical state-space model

The notation, with minor modifications, borrows from Durbin and Koopman (2012). There is a set of $p$ observable variables, called the observation vector $\mathbf{y}_t$, which are related to $m$ unobservable state variables assembled in the state vector $\mathbf{\alpha}_t$, by the observation equation:

$$ (\text{observation equation}) \quad \mathbf{y}_t = \mathbf{d} + \mathbf{Z} \mathbf{\alpha}_t. $$

The state vector $\mathbf{\alpha}_t$ obeys the state equation:

$$ (\text{state equation}) \quad \mathbf{\alpha}_t = \mathbf{T} \mathbf{\alpha}_{t-1} + \mathbf{R} \mathbf{\eta}_t, \quad \mathbf{\eta}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}). $$

$(\mathbf{d}, \mathbf{Z}, \mathbf{T}, \mathbf{R}, \mathbf{Q})$ are called the system matrices.

Write the sample as of the vantage point $v$ as $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$ where $\mathbf{y}_n$ is the latest observation as of $v$. In Table 2 example, for $v = \text{July 6, 2020}$, we have: $t = 1$ for January 1990 and $t = n$ for May 2020. The standard apparatus of Kalman filtering and smoothing can be used to compute $\mathbf{\alpha}_{t|n}$, defined as $\mathbf{\alpha}_{t|n} = \mathbb{E}(\mathbf{\alpha}_t | \mathbf{y}_1, \ldots, \mathbf{y}_n)$, given the parameter value. If, as illustrated in Table 2, the sample is littered with missing values, there is a well-known procedure for accommodating them. “Appendix 2” describes the procedure and proves its validity.

With $\{\mathbf{\alpha}_{t|n}\}_{t=1}^n$ in hand, we can compute the nowcast as follows. Suppose that the value of variable $j$ for the target date $t$, $\mathbf{y}_{jt}$, is missing. Its nowcast is the expected value of $\mathbf{y}_{jt}$ conditional on the sample. Computing it involves either interpolation or extrapolation.

(a) Suppose $t \leq n$, so the target date $t$ falls within the sample period. The observation equation (1) implies that the nowcast of $\mathbf{y}_{jt}$ is the $j$-th element of $\mathbf{d} + \mathbf{Z}\mathbf{\alpha}_{t|n}$.

(b) Suppose, instead, that $t > n$. Compute $\mathbf{\alpha}_{t|n}$ by operating on a recursion starting with $\mathbf{\alpha}_{n|n}$. The recursion is obtained by taking the expectation, conditional on the

The observation equation here admits no measurement error, but the measurement error can be included in the state vector.
sample, of the state equation (2):

\[ \alpha_{s|n} = T \alpha_{s-1|n}, \quad s = n + 1, n + 2, \ldots, t. \]  

(3)

The nowcast of \( y_{jt} \) is again given by the \( j \)-th element of \( d + Z \alpha_{t|n} \).

The GDP nowcast is the nowcast of \( y_{jt} \) for \( j \) corresponding to GDP.

4.2 The Mariano–Murasawa (2003) monthly quarterly model

The Mariano–Murasawa model obtains if the system matrices of the state-space model are properly specialized. The unit time interval is a month. Assume, for now, that all indicator variables including GDP are observable monthly. The observation of the \( j \)-th variable for month \( t \) is \( y_{jt}^* \). It is related to two unobservable common factors \((f_{t}^{(1)}, f_{t}^{(2)})\) and an observation error \( u_{jt} \) (also called the specific factor):\(^{11}\)

\[ y_{jt}^* = \mu_j^* + \beta_j^{(1)} f_{t}^{(1)} + \beta_j^{(2)} f_{t}^{(2)} + u_{jt}, \quad j = 1, 2, \ldots, p. \]  

(4)

Those factors follow first-order autoregressions:

\[ f_{t}^{(i)} = \phi_f^{(i)} f_{t-1}^{(i)} + \xi_t^{(i)}, \quad i = 1, 2, \]  

(5)

\[ u_{jt} = \phi_j u_{j,t-1} + \xi_{jt}, \quad \text{Var}(\xi_{jt}) = \sigma_j^2, \quad j = 1, 2, \ldots, p. \]  

(6)

The shocks to those autoregressions are jointly normally distributed, serially uncorrelated and uncorrelated with each other. Without loss of generality, the variance of the shock to the unobservable common factor, \( \xi_t^{(i)} \), is normalized to unity. The model parameters are:

\[ (\mu_1^*, \mu_2^*, \ldots, \mu_p^*, \beta_1^{(1)}, \beta_2^{(1)}, \ldots, \beta_p^{(1)}, \beta_1^{(2)}, \beta_2^{(2)}, \ldots, \beta_p^{(2)}, \phi_f^{(1)}, \phi_f^{(2)}, \phi_1, \phi_2, \ldots, \phi_p, \sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2). \]

They are \( 5p + 2 \) in number. In our specification below, \( p = 19 \), so there are 97 parameters.

Now assume that what is observable for variable \( j \), \( y_{jt} \), is temporally aggregated, as in

\[ y_{jt} \equiv \omega_{j0} y_{jt}^* + \omega_{j1} y_{j,t-1}^* + \cdots + \omega_{j,\ell_j} y_{j,t-\ell_j}^* + \sum_{s=0}^{\ell_j} \omega_{js} y_{j,t-s}^*, \quad j = 1, 2, \ldots, p. \]  

(7)

\(^{11}\) In Mariano and Murasawa (2003) and also in Bragoli (2017), there is only one unobservable common factor. Here we allow two. We will address this issue of the choice of the number of common factors later in the paper, in Sect. 7.1.
The weights \((\omega_{j0}, \ldots, \omega_{j,\ell_j})\) can depend on the variable \(j\). Examples include:

\[
\begin{align*}
\text{(no aggregation)} & \quad \ell_j = 0, \quad \omega_{j0} = 1. \quad (8) \\
\text{(level)} & \quad \ell_j = 2, \quad \omega_{j0} = \frac{1}{3}, \quad \omega_{j1} = \frac{1}{3}, \quad \omega_{j2} = \frac{1}{3}. \quad (9) \\
\text{(growth rate)} & \quad \ell_j = 4, \quad \omega_{j0} = \frac{1}{3}, \quad \omega_{j1} = \frac{2}{3}, \quad \omega_{j2} = 1, \quad \omega_{j3} = \frac{2}{3}, \quad \omega_{j4} = \frac{1}{3}. \quad (10)
\end{align*}
\]

The “level” weights (9) obtain if the variable \(j\) is a quarterly variable that takes the average over the three constituent months. The “growth rate” weights (10) require some explanation. Let \(y^*_j t\) be the monthly growth rate of variable \(j\) from month \(t - 1\) to \(t\), and let \(y_{jt}\) be the quarterly growth rate which is the log difference in the 3-month average of the variable from the previous 3-month period (months \(t - 5, t - 4,\) and \(t - 3\)) to the current 3-month period (months \(t - 2, t - 1,\) and \(t\)). As Mariano and Murasawa (2003) noted, we have:

\[
y_{jt} \approx \frac{1}{3} y^*_j t + \frac{2}{3} y^*_j, t - 1 + y^*_j, t - 2 + \frac{2}{3} y^*_j, t - 3 + \frac{1}{3} y^*_j, t - 4,
\]

with the weights as given in (10). This will be the aggregation weights for quarterly real GDP growth.

To derive the observation equation after temporal aggregation, substitute the \(j\)-th observation equation in (4) into (7) to obtain

\[
y_{jt} = \mu_j + \beta^{(1)}_j \sum_{s=0}^{\ell_j} \omega_{js} f_{j,t-s}^{(1)} + \beta^{(2)}_j \sum_{s=0}^{\ell_j} \omega_{js} f_{j,t-s}^{(2)} + \sum_{s=0}^{\ell_j} \omega_{js} u_{j,t-s}, \quad j = 1, 2, \ldots, p. \quad (12)
\]

where

\[
\mu_j \equiv (\omega_{j0} + \cdots + \omega_{j,\ell_j}) \mu^*_j.
\]

This is the observation equation in the Mariano–Murasawa model. The system matrices in a state-space representation of the model are in “Appendix 1.”

5 The real-time database

This section is a summary description of our real-time database. 12

5.1 The indicator variables

The set of indicator variables is composed of 19 monthly and quarterly variables listed in the “name” column of Table 3. It is a standard set of variables used in many

---

12 A full description is in a document downloadable from https://drive.google.com/file/d/1_nWMMGz8HrYwEk69OoBv-HXvS2QAurte/view?usp=sharing.
nowcasting studies and followed by market participants. All non-survey-based (i.e., “hard”) indicators are seasonally adjusted at the source, except for two variables \((j = 7, 8)\) which are seasonally adjusted by us. We do not include financial indicators such as stock prices and interest rates, because most nowcasting studies report their limited contribution to nowcasting performance.

The set includes five “soft” indicators derived from surveys \((j = 14, \ldots, 18)\). They are diffusion indices about business conditions. We follow the suggestion of Vermeulen (2012) to take the logistic transformation of these diffusion indexes. Respondents for the Economy Watchers’ Survey \((j = 14)\) and the Consumer Confidence Survey \((j = 16)\) are individuals. The Tankan Survey \((j = 17)\) covers a wide cross section of firms. For the Local Business Outlook survey \((j = 15)\) and the Large Firms Survey \((j = 18)\), the respondents are small-to-medium sized firms in the former survey and large firms in the latter. It is reasonably clear from the questionnaire that the Tankan survey is about the level of business conditions while the Large Firms Survey is about the change.

Many of the indicators appear to have trends. Following most other nowcasting studies, we take the first difference (or the first log difference as the case may be) for those indicators. The “transformation” column of Table 3 shows how we transformed the original monthly and quarterly time series before feeding them to the Mariano–Murasawa dynamic factor model. In particular, the real GDP growth rate is a log difference.

The last column reports the publication lag in days between the end of the reference period and the time of the first data release. For all monthly indicators, the publication lag is shorter than that for GDP of 46 days.

13 The makeup of our indicator variable set is similar to that for “Model A” of Bragoli (2017), all of which score high on Bloomberg’s relevance index. There are two differences. We replace the PMI (Purchasing Managers Index) in Bragoli’s set, which is not freely available, by the Tankan, a business condition index compiled by the Bank of Japan. The Tankan is widely followed by forecasters and prominently featured in the media. We replace the SME survey, whose publication has been suspended for several years, by the “LOBO” (the Local Business Outlook survey conducted by Japan’s Chamber of Commerce). Neither our set nor Bragoli’s include the CPI (consumer price index) because, as Bragoli (2017) reports, its inclusion does not improve the nowcast, which is not surprising given that the set of indicator variables already includes both real and nominal variables.

14 See, e.g., Section 3 of Banbura et al. (2013). Bragoli (2017) finds the same for Japan.

15 A diffusion index is an aggregate of categorical responses. Let \(x_i = (1, 2, \ldots, K)\) be the categorical response by respondent \(i = (1, 2, \ldots, I)\). A diffusion index is defined as the share-weighted average of the response, i.e.,

\[
\sum_{k=1}^{K} \frac{\# \{i | x_i = k\}}{I} k,
\]

where \(#A\) is the cardinality (the number of elements) of set \(A\).

16 If \(x\) is the diffusion index in percent, its logistic transformation is \(100 \times \log\left[\frac{U-x}{x-L}\right]\). We take \((U, L) = (100, 0)\) for Economy Watchers’ Survey \((j = 14)\) and Consumer Confidence Survey \((j = 16)\), and \((U, L) = (100, -100)\) for Local Business Outlook \((j = 15)\), Tankan \((j = 17)\), and for Large Firms Survey \((j = 18)\). We also take the logistic transformation of Capacity Utilization index \((j = 2)\), with \(U = 150\) and \(D = 0\). The upper bound is set to 150 rather than 100 because the index frequently exceeds 100.
| Variable no. (j) | Name                               | Original unit | Transformation | Approx. days between end of reference period and release date |
|-----------------|------------------------------------|---------------|----------------|-------------------------------------------------------------|
| 1               | Monthly earnings                   | Index         | Δ log          | 34                                                          |
| 2               | Capacity utilization               | Index         | Δ              | 45                                                          |
| 3               | Construction orders                | Yen           | Δ log          | 30                                                          |
| 4               | Housing starts                     | Housing units | Δ log          | 30                                                          |
| 5               | Machinery orders                   | Yen           | Δ log          | 42                                                          |
| 6               | Unemployment                       | #persons      | Δ              | 29                                                          |
| 7               | Passenger car production           | #vehicles     | Δ log          | 29                                                          |
| 8               | Vehicle sales                      | #vehicles     | Δ log          | 2                                                           |
| 9               | Exports                            | Yen           | Δ log          | 21                                                          |
| 10              | Imports                            | Yen           | Δ log          | 21                                                          |
| 11              | IIP (Index of Industrial Production) | Index     | Δ log          | 29                                                          |
| 12              | Tertiary Industry Index            | Index         | Δ log          | 45                                                          |
| 13              | Retail Sales Index                 | Index         | Δ log          | 29                                                          |
| 14              | Economy Watchers’ Survey          | Index         | Δ              | 9                                                           |
| 15              | Local Business Outlook Survey      | Index         | Δ              | −1                                                          |
| 16              | Consumer Confidence Survey         | Index         | Δ              | 8                                                           |
| 17              | Tankan Survey                      | Index         | Δ              | −3                                                          |
| 18              | Large Firms Survey                 | Index         | None           | −17                                                         |
| 19              | Real GDP                           | Constant yen  | Δ log          | 46                                                          |

The frequency is monthly for the first 16 indicators ($j = 1, 2, \ldots, 16$) and quarterly for $j = 17, 18, 19$. The last column reports the publication lag in days between the end of the reference period and the time of the first data release. It can be negative if the data release occurs before the end of the reference period. For example, the result of the Tankan survey for the fourth quarter is released in mid-December.
5.2 Temporal aggregation weights

The set of indicators includes three quarterly variables \((j = 17, 18, 19)\). We need to specify the vector of aggregation weights \((\omega_{j0}, \ldots, \omega_{j,\ell})\), defined in (7), that converts underlying monthly values into quarterly values.

- We have already indicated in Sect. 4.2 that the weights for real GDP \((j = 19)\), which aggregates monthly growth rates to quarterly growth rates, should be the “growth rate” weights (10).
- The weights for the Large Firms Survey \((j = 18)\) should be the same as those for real GDP, because the “change” in the questionnaire refers to the change from the previous quarter to the current quarter.
- Regarding the Tankan survey \((j = 17)\), we interpret “the current business condition” in the questionnaire to be about the current month. So if we take the first difference in Tankan’s (quarterly) diffusion index, it is a 3-month change from the previous survey. Since a 3-month change (stated at a monthly rate) is the average of the most recent three successive monthly changes, the “level” weights are appropriate.

6 The ESP forecast

As the benchmark forecast against which our nowcast is made to compete, we consider the ESP forecast, a monthly market consensus forecast compiled by the JCER (Japan Center for Economic Research). It is a survey that solicits approximately forty prominent professional forecasters for their forecasts of quarterly GDP growth and other macro variables over several successive quarters. The respondents are required to turn in their forecasts early in the month. The JCER then publishes the mean forecast in their monthly report.

The structure of the monthly ESP survey can be seen from Table 4. Forecasts of current and subsequent GDP growth at given points in time are in the columns of the table. For example, look at the column corresponding to the May 2020 report whose survey deadline is May 11, 2020. Its March 2020 entry of \(-4.74\%\) is the ESP consensus forecast, as of May 11, of the annualized real GDP growth from 2019:Q4 to 2020:Q1 (whose end-of-quarter month is March 2020). It is the last forecast for March 2020 before the flash (first official estimate) was released on 2020/5/18. The June 2020 entry of \(-23.99\%\) is the consensus forecast, formed on May 11, of the 2020:Q1 to Q2 growth. The other entries of the column are longer-horizon forecasts formed on the same date.

The staggered structure of the ESP forecast means that each row of the table records the evolution of the forecast for a fixed target end-of-quarter month over several successive monthly horizons. Consider, for example, the 2020:Q1 to Q2 growth as the fixed target. The evolution of its nowcast is in the row for June 2020 (the ending

17 For a description of the Tankan survey, visit http://www.boj.or.jp/en/statistics/outline/exp/tk/extk03.htm/.
18 For more details on the ESP forecast, visit https://www.jcer.or.jp/en/esp-forecast-top.
### Table 4 Real GDP growth forecasts in the ESP survey

| End-of-quarter month | Deadline for professional forecasters |
|----------------------|---------------------------------------|
|                      | Feb. 5, 2020  | Mar. 10, 2020 | Apr. 8, 2020 | May 11, 2020 | June 11, 2020 | July 6, 2020 | Aug. 7, 2020 |
| December 2019        | −4.13        |              |             |              |              |              |              |
| March 2020           | 0.33         | −2.93        | −4.14        | −4.74        |              |              |              |
| June 2020            | 1.53         | 2.28         | −11.74       | −23.99       | −26.16       | −26.83       | −30.91       |
| September 2020       | 1.49         | 2.86         | 4.98         | 7.92         | 8.66         | 8.99         | 12.45        |
| December 2020        | 0.64         | 0.97         | 3.83         | 5.27         | 4.71         | 4.59         | 4.60         |

Annualized quarterly growth rates in percent. The source is the ESP survey published by JCER (Japan Center for Economic Research). The real GDP growth rate reported in the ESP survey is the rate of change in percent. The GDP growth rate here has been converted from the rate of change to the log difference. The conversion formula is $100 \times \log(1 + z/100)$ where $z$ is the ESP consensus forecast in annualized percent as originally published by JCER.
month of 2020:Q2). The right-most entry of $-30.91\%$ is the consensus forecast, as of August 7, of the 2020:Q1 to Q2 growth (whose flash turned out to be, as shown in Table 1, the jaw-dropping $-32.56\%$ released on 2020/8/17). We say that its horizon in months is $-1$ because the vantage point, which is the survey deadline of August 7, is approximately 1 month after the conclusion of the quarter (of 2020:Q2). The entry of $-26.83\%$ to the left of it has a horizon of 0 because its vantage point of July 6 is shortly after the quarter’s end. Similarly, the forecast of 1.53\% in the February 2020 entry of the June 2020 row has a horizon of 5 months. This row shows that the forecasters became aware of the devastating effect of COVID-19 sometime between March 10 (whose horizon is 4 months) and April 8, 2020 (whose horizon is 3 months).

7 Results from multi-horizon RMSE comparison

This section reports a comparison of our nowcast with the ESP consensus forecast in terms of the root mean squared error (RMSE). For the target GDP growth that goes into the RMSE calculation, we will consider two measures: the flash (the first official GDP estimate) and the “final” which we define as the official estimate released two years after flash. For example, as can be read off from Table 1, the flash for the 2008:Q4 to 2009:Q1 growth is $-16.49\%$ released on 2009/5/20 and the final is $-20.20\%$ released on 2011/5/19.

7.1 Preliminaries

We have described in Sect. 4 how to compute our nowcast from the sample given the model parameters. We set the sample size to 20 years (240 months) in size. We follow the practice of the New York Fed Staff nowcast to revise the parameter vector at the beginning of every quarter. Parameter estimation is by maximum likelihood. See “Appendix 2” for the evaluation of the Gaussian likelihood function with missing values, and “Appendix 3” for details on the numerical optimization algorithm.

As assumed in the observation equations (4) or (12), the number of common factors is two. The information criteria of Bai and Ng (2002) for determining the number of factors is not directly applicable here because they do not accommodate mixed frequencies. Andreou et al. (2019), which allows for mixed frequencies, is not applicable either, because the aggregation weights are not the same across quarterly variables in our model. Given these constraints, we proceed as follows. We apply the $IC_{p1}$, $IC_{p2}$ and $IC_{p3}$ information criteria of Bai and Ng (2002) on a balanced panel of the monthly indicators ($j = 1, 2, \ldots, 16$) listed in Table 3, with the maximum number of common
factors set to 5. The number of common factors selected by $IC_{p1}$ is the maximum of 5, while the number selected by $IC_{p2}$ and $IC_{p3}$ is 1. We assume the common factors to be two in number, to allow for an additional factor that affects the three quarterly indicators ($j = 17, 18, 19$) but not the monthly indicators ($j = 1, 2, \ldots, 16$).

Availability of vintages determines the backtesting period for the RMSE calculation. We take the earliest GDP target end-of-quarter month to be March 2008. That is, the earliest GDP growth to forecast is from 2007:Q4 to 2008:Q1. It is an earliest time when vintage time series of most indicator variables are available. The backtesting period’s latest target end-of-quarter month is March 2020 (i.e., the latest GDP growth to forecast is from 2019:Q4 to 2020:Q1), because the final (the GDP estimate released two years after the flash) for 2020:Q2 is not yet available at this time of writing (of June 2022).

We define the RMSE for forecasting horizon $k$ as

$$RMSE_k = \sqrt{\frac{1}{\# \mathcal{R}} \sum_{t \in \mathcal{R}} (y_t - \hat{y}_t|v_k(t))^2},$$

$$\mathcal{R} \equiv \{ \text{March 2008, June 2008, \ldots, March 2020} \},$$

where $\mathcal{R}$, a collection of target end-of-quarter months, is the backtesting period, “$\# \mathcal{R}$” is the cardinality (the number of elements) of $\mathcal{R}$ (which is 49), $y_t$ is the target quarterly real GDP growth to forecast for the quarter ending in month $t$, $\hat{y}_t|v$ is the forecast (be it the nowcast or the ESP consensus forecast) of $y_t$ as of vantage point $v$, and $v_k(t)$ is the vantage point whose horizon relative to $t$ is about $k$ months (so $\hat{y}_t|v_k(t)$ is the $k$-month ahead forecast of $y_t$). This $v_k(t)$ is taken from the ESP survey. For example, as seen from Table 4’s row for $t = June 2020$ (corresponding to the 2020:Q1 to Q2 growth), we have $v_k(t) = August 7$ if $k = -1$, July 6 if $k = 0$, and February 5 if $k = 5$.

19 The balanced panel we consider is the sample as of the time of this writing (June 3, 2022). This choice of the vantage point is dictated by the fact that, for the Consumer Confidence Survey ($j = 16$), the time series starts only since May 2004 in all of its vintages. We take the latest possible vantage point to maximize the sample size.

20 The earliest vantage point, $v_k(t)$ for $t = March 2008$ and $k = 11$, is $v = May 7, 2007$. As already mentioned in the text, we revise the parameter estimate by ML (maximum likelihood) at the beginning of every quarter. So the first ML parameter estimate is as of April 1, 2007. The model parameters are re-estimated every quarter thereafter. In our real-time database, the date of the earliest available vintage differs across indicators, so we also need to re-estimate parameters when a new indicator enters the set after April 1, 2007. Such dates are: April 17, 2007 (which is the release date of the earliest available vintage of Consumer Confidence Survey), June 20, 2007 (Large Firms Survey), December 20, 2007 (Exports and Imports), January 11, 2008 (Economy Watcher), January 16 (Retail Sales), January 29 (Unemployment), January 31 (Construction Orders and Housing Starts) of 2008. For a few early vantage points (i.e., for $v_k(t)$’s for early $t$’s and relatively long $k$’s), the sample size is slightly less than 20 years, because, for all indicator variables, the earliest reference point of its vintages (time series) is January 1990 or later. For example, for the earliest vantage point of $v = May 7, 2007$, the sample as of this $v$ runs from January 1990 to April 2007. Its sample size is 208 months rather than 240 months.
### 7.2 Accuracy comparison of our nowcast with the consensus at short horizons

Table 5 reports the RMSE given by (14) at short horizons, $k = 0$ (shortly after the conclusion of the quarter) and $k = -1$ (1 month after), for our nowcast and the ESP consensus forecast. The first two columns of numbers show that, when the GDP growth measure $y_t$ is the final, our nowcast slightly outperforms (i.e., the RMSE is smaller than) the consensus at $k = 0$ but not at $k = -1$. However, as the Diebold-Mariano (Diebold and Mariano (1995)) $t$-value indicates, the difference is not statistically significant.

As the middle columns of the table show, the consensus forecast beats the nowcast when the GDP growth measure is the flash. At the shortest horizon ($k = -1$), the difference is statistically significant. A possible explanation of the consensus’s superiority is that it is optimized to the flash.21 Shortly after the end of the quarter, it may be possible for the professional forecasters to guess the government’s preliminary estimate of the underlying demand components of GDP that would go into the official calculation of the flash. This is particularly true for $k = -1$. Consider, for example, a Q1 to Q2 growth. As illustrated in Table 2, the vantage point with $k = -1$ is in early August, slightly more than a month after the last month (June) of the ending quarter but before the release of the flash (thanks to the six-week lag). By then, for a large number of key monthly indicators in addition to those in our set of indicators, their June values are already out. Those additional indicators may be useful at $k = -1$.22

We note in passing that GDP nowcasting for Japan exhibits less forecast accuracy than reported elsewhere for the USA and the Euro Area. Our nowcast’s RMSE for the final GDP growth measure at the shortest horizon ($k = -1$) is 2.93% per year, or 0.73% per quarter. The target real GDP measure in the real-time backtesting for the USA in Bańbura et al. (2013) is the most recent revision, similar to our “final” measure here. Their RMSE, if the vantage point is in the middle of the first month following the end of the quarter, is about 0.5% per quarter. The RMSE for Euro Area GDP in Table V of Bańbura and Modugno (2014) for a similar vantage point is, somewhat puzzlingly, as low as 0.2%.23 In contrast, the RMSE for Japan’s real GDP growth in Bragoli (2017) is slightly higher than ours, about 0.8% per quarter (which can be read off from her Figure 2).

The backtesting period (the $R$ in (14)) does not include the COVID-19 end-of-quarter months from June 2020 on because for those months the “final” GDP growth estimate is not yet available at this time of writing (of June 2022). However, the “flash” is available for those months. The last two columns of Table 5 contain the RMSE comparison if the backtesting period is composed of end-of-quarter target

---

21 The coveted forecasters-of-the-year award, doled out by the entity (the JCER) that disseminates the consensus forecast, goes to forecasters who are successful in forecasting the flash.

22 The superiority of market consensus forecasts at short horizons is not limited to forecasting Japan’s GDP. Drawing on Sims (2002), Antolin-Díaz et al. (2021) states: “Human forecasters are a high benchmark for econometric models in particular for short-run forecasts, as they have access to a large information set including news, political developments, and information about structural change. Consensus measures, …, are an even higher benchmark as they aggregate the opinions of a multitude of forecasters.”

23 However, their backtest is only “pseudo” real-time, rather than fully real-time, which amounts to assuming that most recent vintages were known at the time of forecasting in the past.
Table 5  RMSE comparison, nowcast versus consensus

| Backtesting period (\(R\)) | Pre-COVID \(R\) \(=\) [March 2008,…, March 2020] | COVID \(R\) \(=\) [June 2020,…, March 2022] |
|---------------------------|---------------------------------|---------------------------------|
| GDP growth measure        | Final                           | Flash                           |
| Horizon \((k)\)           | \(k = 0\)                       | \(k = 0\)                       |
|                          | \(k = -1\)                      | \(k = -1\)                      |
| RMSE for Nowcast          | 3.36                            | 4.05                            |
|                          | 2.93                            | 3.96                            |
| RMSE for Consensus        | 3.40                            | 4.46                            |
|                          | 2.79                            | 2.14                            |
| \(t\)-value for the difference | −0.11                          | −                           |
|                           | 0.40                            | 2.14                            |
|                           | 1.65                            | 2.78                            |

The RMSE calculated by (14) where the target real GDP growth \((y_t)\) is an annualized log difference in percent. That is, \(y_t = 400 \times [\log(Y_t) - \log(Y_{t-3})]\) where \(Y_t\) is the level of real GDP for the quarter whose ending month is \(t\). The GDP growth measure is either the “flash” (the first preliminary estimate released by the Cabinet Office about 6 weeks after the conclusion of the ending quarter) or the “final” (the revised GDP estimate released 2 years after the flash). “Nowcast” refers to our dynamic factor model-based nowcast of real GDP growth. “Consensus” refers to the ESP consensus forecast. For both the ESP forecast and the GDP release by the Cabinet Office, the growth rate is the annualized rate of change by the formula \(100 \times \left( \frac{Y_t}{Y_{t-3}} \right)^4 - 1\). To be consistent with how \(y_t\) is defined, we convert the annualized rate of change to the annualized log difference in percent. The conversion formula is \(100 \times \log(1 + \frac{\epsilon}{100})\) where \(\epsilon\) is the annualized rate of change in percent. The estimate of the long-run variance, which is needed to calculate the Diebold–Mariano (Diebold and Mariano 1995) \(t\) statistic, includes up to third-order autocovariances. The Mariano–Diebold \(t\) statistic is not reported for the COVID period because of the small sample size.
months from June 2020 through March 2022. Not surprisingly, the forecast accuracy is relatively poor under COVID. As in the pre-COVID backtesting period, the consensus forecast is significantly more accurate at the shortest horizon \(k = -1\), but somewhat surprisingly, our nowcast outperforms the consensus at \(k = 0\).

**7.3 RMSE comparison at longer horizons**

Figure 1 shows the RMSE profile (the plot of \(RMSE_k\) against horizons \(k = -1, 0, 1, \ldots, 11\)) for our nowcast and the consensus forecast. The measure of the target real GDP growth \(y_t\) is the final here. Two notable features about the figure are:

- For the consensus forecast, the forecast accuracy as measured by RMSE steadily improves as horizon \(k\) gets shorter. That is, the RMSE profile is monotonically declining. This is what one would expect.
- The forecast accuracy of our nowcast is about as good as that of the consensus forecast at most monthly horizons, except at \(k = 2\) and \(k = 3\) where the profile has a hump.

To anticipate, the revision analysis of the next section will show that the hump is largely due to the nowcast’s large forecast error for the 2011:Q1 to Q2 target growth made during the days ensuing the March 2011 earthquake.

![Fig. 1 Nowcast versus consensus, GDP measure is the final. Notes: The RMSE calculated by (14) where the target real GDP growth \((y_t)\) is an annualized log difference in percent. That is, \(y_t = 400 \times \left[\log(Y_t) - \log(Y_{t-3})\right]\) where \(Y_t\) is the level of real GDP for the quarter whose end-of-quarter month is \(t\). The measure of GDP growth is the “final” (the revised GDP estimate released 2 years after the “flash” (the first official estimate) which in turn is released about 6 weeks after the end of the quarter). “Nowcast” refers to our dynamic factor model-based nowcast of real GDP growth. “Consensus” refers to the ESP consensus forecast. For both the ESP forecast and the GDP release by the Cabinet Office, the growth rate is the annualized rate of change by the formula \(100 \times \left[\left(Y_t / Y_{t-3}\right)^4 - 1\right]\). To be consistent with how \(y_t\) is defined, we convert the annualized rate of change to the annualized log difference in percent. The conversion formula is \(100 \times \log(1 + z/100)\) where \(z\) is the annualized rate of change in percent](image)
The next figure (Fig. 2) differs from Fig. 1 in that the measure of the target GDP growth is the flash rather than the final. Reflecting the fact that revisions subsequent to the flash are large for Japan, the RMSE is substantially lower for both forecasts and for all horizons. The profile for the consensus lies significantly below that for the nowcast at the short end of the horizon, the fact we already noted for Table 5.

8 A revision analysis of three disasters

In the revision analysis, the interest is how the nowcast \( \hat{y}_t | v \) of the target variable \( y_t \) for a fixed given target end-of-quarter month \( t \) evolves as the vantage point \( v \) progresses in time. This section carries out the revision analysis for three choices of the target month \( t \) associated with three disasters: the Lehman crisis, the March 2011 earthquake and the ensuing Fukushima Daiichi nuclear disaster, and the COVID-19 voluntary lockdown of 2020.

8.1 The new-observations and data-revisions effects in the revision analysis

The nowcast \( \hat{y}_t | v \) is a function of the sample as of \( v \) and the parameter vector. Consider an interval \( (v', v'') \) defined by two successive vantage points \( v' \) and \( v'' \). Since, as illustrated in Table 2, new observations and data revisions constitute the difference in the sample between the two endpoints \( v' \) and \( v'' \) of the interval, the nowcast change over the interval, \( \hat{y}_t | v'' - \hat{y}_t | v' \), is composed of three effects: the new-observations, data-revisions, and parameter-revisions effects. In the existing revision analysis, only the new-observations effect is broken down into contributions from individual indicator
variables. The method of Hayashi and Tachi (2021) allows one to break down the other two effects as well.\textsuperscript{24}

In the revision analysis below, we follow the practice of using a bar graph to show a breakdown. That is, for each interval \((v', v'')\), the height of the bar equals \(\widehat{y}_{t|v''} - \widehat{y}_{t|v'}\). The bar is made up of segments representing the breakdown. To prevent the bar graph from becoming exceedingly cluttered, we aggregate the set of indicator variables into two groups: “hard” variables \((j = 1, 2, \ldots, 13, 19\), see Table 3 for the identity of those variables) and “soft” (i.e., survey-based) variables \((j = 14, \ldots, 18\). The contribution from the hard variables as a whole will be divided between the new-observations and data-revisions effects. For the soft variables, the data-revisions effect should be zero or close to zero because surveys involve no data revisions (except for occasional seasonal re-adjustment due to increased lengths of each vintage). For this reason, we won’t make the new-observations vs data-revisions distinction for the soft variables. We could, but will not, break down the parameter-revisions effect to maintain the readability of the bar graph. Therefore, the bar will have a 4-way makeup, consisting of four segments to be labeled as: “Hard, new observations,” “Hard, data revisions,” “Soft,” and “Parameter revisions.”\textsuperscript{25}

8.2 The Lehman crisis

The first episode for our revision analysis is the Lehman crisis. We take the fixed target end-of-quarter month \(t\) to be March 2009 (so the target growth \(y_t\) is from 2008:Q4 to 2009:Q1) when the real GDP growth registered the (then) postwar low.

Figure 3 displays the result. The hollow circle and the “x” at the right of the figure are the two measures of target GDP growth, the flash (released on 2009/5/20) and the final (released two years later, on 2011/5/19). Their difference is large: the flash of 49% per year which was revised down to -20.20% in two years. The filled-in circle tracks the evolution of the nowcast \(\widehat{y}_{t|v}\) as the vantage point \(v\) proceeds in time. We consider 16 vantage points varying from \(v = \text{December 2, 2008 to v = May 12, 2009, forming 15 intervals. Each interval is about 10 days wide. Included in those vantage points are monthly ESP vantage points. The filled-in triangle tracks the evolution of the consensus over those ESP vantage points.

A notable feature of Fig. 3 is the sharp downward revision of the nowcast \(\widehat{y}_{t|v}\) at \(v = \text{April 6, 2009}\). The makeup of the bar for the interval from March 23 to April 6 shows that the major contributor is the parameter revision on April 1, 2009. Many monthly revision analyses do the same only for revisions, and parameter-revisions effects from variable \(j\) that the major contributor is the parameter revision on April 1, 2009. Many monthly

\textsuperscript{24} To elaborate, Hayashi and Tachi (2021) show that \(\widehat{y}_{t|v''} - \widehat{y}_{t|v'}\) can be written as \(J \sum_{j=1}^{J} n_j + J \sum_{j=1}^{J} r_j + \sum_{j=0}^{J} p_j\) where \(J\) is the number of indicator variables, \((n_j, r_j, p_j)\) are the new-observations, data-revisions, and parameter-revisions effects from variable \(j\). This decomposition obtains because the nowcast \(\widehat{y}_{t|v}\) is the least-squares projection of \(y_t\) on the sample as of \(v\). The summation for \(p_j\) is from \(j = 0, \) not from \(j = 1, \) to account for the intercept in the least-squares projection. Hayashi and Tachi (2021) also show how to compute each element of \((n_j, r_j, p_j)\) from real-time data. The conventional revision analysis does the same only for \((n_j)\).

\textsuperscript{25} Using the notation in footnote 24, the bar will be composed of: \(\sum_{j \in \mathcal{H}} n_j, \sum_{j \in \mathcal{H}} r_j, \sum_{j \in \mathcal{S}} (n_j + r_j), \) and \(\sum_{j=0}^{J} p_j\), where \(\mathcal{H} = \{1, 2, \ldots, 13\}\) and \(\mathcal{S} = \{14, \ldots, 18\}\). The bar in the New York Fed Staff Nowcast Report is composed of: \(n_j\) for a partition of \(\{1, 2, \ldots, J\}, \sum_{j=1}^{J} r_j, \) and \(\sum_{j=0}^{J} p_j\).
Nowcasting Japan’s GDP

Fig. 3 Revision analysis, target is 2008:Q4 to 2009:Q1 growth. Notes: The target real GDP growth is from 2008:Q4 to 2009:Q1. The growth rate is annualized and is defined as 400 times the quarterly log difference. The two measures of the target GDP growth—the “flash” (i.e., the first preliminary estimate) and the “final” (the revised estimate 2 years after the flash)—are shown at the right of the figure. “Nowcast” refers to our dynamic factor model-based nowcast of the real GDP growth. The height of a bar in the bar graph equals the change in the nowcast from the previous vantage point to the current. Each bar is divided into four segments. The “Hard, new observations” segment measures the new-observations effect from the “hard” (non-survey-based) variables as a whole. The “Hard, data revisions” is the data-revisions effect from the “hard” variables as a whole. “Soft” is the sum of the new-observations and data-revisions effects from the five survey-based indicators. “Parameter revision” is the parameter-revisions effect through all 19 indicator variables. “Consensus” refers to the ESP forecast converted to log differences. The conversion formula is 100 \times \log(1 + z/100) where z is the annualized rate of change in percent as reported in the ESP report.

indicators as of April 1 show negative growth over several consecutive months since the end of the previous year. Having seen this, our dynamic factor model decides that the underlying common factors must have been negative and persistent with no turnarounds during 2009:Q1. The newly found persistence in the common factors forces the model to lower the nowcast of the 2008:Q4 to 2009:Q1 growth. However, at the shortest end of the horizon (at \( k = -1 \) or \( v = \text{May 12} \)), the nowcast almost matches the consensus, as data releases after April 6 start reporting positive changes.

The data-revisions effect represented by the “Hard, data revisions” segment is relatively minor, except at the April 6 to April 16 interval. Over this interval, there were data releases on Capacity Utilization and IIP with substantial data revisions.

\[ \phi^{(1)}_f, \phi^{(2)}_f \] in (5), is \((-0.29, -0.37)\) as of January 1, 2009, and \((0.32, -0.37)\) as of April 1, 2009.
8.3 The earthquake of March 11, 2011

The next figure (Fig. 4) covers the earthquake of March 11, 2011, and the ensuing nuclear accident at the Fukushima Daiichi. The target growth $y_t$ is from 2011:Q1 to Q2, or $t = \text{June } 2011$.

The figure’s striking feature is the nowcast’s large U-shaped movement. As can be seen from the 4-way makeup of the bars between April 28 and June 1, 2011, the main driver is the hard indicator variables. This segment during this period was dominated by data releases on IIP and Passenger Car Production that report sharp February to March drops and quick rebounds for April. For example, Passenger Car Production showed a February to March decline of 62% (according to the 2011/4/28 release) followed by a 40% increase from March to April (according to the 2011/5/31 release).

The gyration for automobile-related indicators is widely attributed to the well-publicized disruption of the domestic supply chain. The same is probably true for many other components of IIP. The nowcast, unaware that the gyration was a responses to the temporary shock due to the earthquake, obliges by fully responding to it. In contrast, the consensus hardly budged, most likely because the forecasters dismissed those monthly movements as temporary, with no implication for the second-quarter GDP. However, as in the Lehman episode, with subsequent data releases of positive changes coming in, the nowcast caught up with the consensus at short horizons.

![Revision Analysis, Target is 2011:Q1 to Q2 Growth](image)

Fig. 4 Revision analysis, target is 2011:Q1 to Q2 growth. Notes: The target real GDP growth is from 2011:Q1 to Q2. See note to Fig. 3 for other details
8.4 The COVID-19 voluntary lockdown of 2020

The residents in Japan were requested (but not required) to stay home during the second quarter of 2020. The flash estimate of the 2020:Q1 to Q2 growth was, as seen from Table 1, the jaw-dropping \(-32.56\%\) per year. Figure 5 shows that while both the nowcast and the consensus took a precipitous fall, the nowcast lagged behind. As seen from the “Soft” segment of the bar graph, the soft indicators provided timely red flags, but the nowcast failed to respond in full because it took a month or more for the hard indicators to reflect the devastating impact of the lockdown. In particular, for the interval from March 19 to March 31, the negative effect from the soft variables is nearly offset by the hard variables. This is mainly due to the IIP release on 2020/3/31 reporting a positive change from January to February. However, as in the earlier two episodes, the nowcast started narrowing the difference with the consensus, this time because the hard indicators ceased sending conflicting signs.

At the shortest horizon \(k = -1\) or at \(v =\) August 7, 2020, the consensus’s forecast error is smaller than the nowcast’s if the GDP growth measure is the flash. It is possible, however, that the consensus’s superiority will vanish if the growth measure is the final (to be released in August 2022).

8.5 The mean data-revisions effect

The “Hard, data revisions” segment of the bar graph gets thin if data revisions are infrequent for many of the hard variables. To adjust for the dilution due to the infrequency, we here consider, for each hard variable, the mean data-revisions effect, defined as the absolute value of the data-revisions effect averaged over those intervals where there is a data revision for the variable. Define the mean new-observations effect similarly. Table 6 displays the ratio of the mean data-revisions effect to the mean new-observations effect. The entry “n.a.” indicates that there are no data revisions in any of the 15 intervals of the episode.

The takeaway from the table is that the data-revisions effect cannot be ignored for GDP, Capacity Utilization, and other staple indicators. It can even be a dominant factor. The ratio for GDP for the Lehman crisis means the data-revisions effect is 6.47 times as large as the new-observations effect. The 15 intervals for this episode covers the period from December 2, 2008, to May 12, 2009 (see Fig. 3). There are three GDP

---

27 This phenomenon of soft indicators providing timely warnings followed belatedly by hard indicators during the 2020 COVID crisis is also observed in the context of nowcasting Euro-area GDP. See Figure 10 of Cascali-Garcia et al. (2021).

28 To be more precise, let \(r_{jv}\) be the data-revisions effect for variable \(j\) over interval \(v = 1, 2, \ldots, 15\). For example, for the COVID episode, the first interval is from March 10 to March 19, 2020, and the last of the 15 intervals is from July 27 to August 7, 2020. Let \(\mathcal{V}_j\) be the set of intervals for which there is a data revision for variable \(j\) and let \(#\mathcal{V}_j\) be its cardinality. This \(#\mathcal{V}_j\) is an integer between 0 and 15. Then, the mean data-revisions effect for variable \(j\) for the episode in question is

\[
\frac{1}{#\mathcal{V}_j} \sum_{v \in \mathcal{V}_j} r_{jv}.
\]

If \(#\mathcal{V}_j = 0\), the value is missing.
releases during this period. One of them (released on 2009/2/16) has a flash estimate (for December 2008, or the 2008:Q3 to Q4 growth) and all three of them include quite substantial data revisions.

9 Concluding remarks

Our conclusion is that the model-based nowcasting is a useful complement to market forecasts. Thanks to its transparency and daily updatability, it provides a real-time stringent check on existing forecasts. The check is stringent because our nowcast performs as well as the market consensus at most forecasting horizons. It is true that the consensus is a better predictor of the flash GDP growth estimate at very short horizons, but we find it significant that the consensus’s advantage disappears after a series of revisions about GDP.

The results from our revision analysis of the three disasters point to various directions along which the nowcast’s performance might be improved.

- Since Banbura et al. (2013), which found higher frequency indicators not helping to improve forecast accuracy, most nowcasting studies, including ours, chose the unit time interval to be a month. Our nowcast’s delayed response to the onset of COVID, shown in Fig. 5, should be mitigated if higher-frequency indicators (such as the Japanese equivalent of the weekly initial claims data from the U.S. Bureau of Labor Statistics) or even their monthly aggregates are included.
- Our nowcast’s over-reaction to the gyration in car production in the aftermath of the March 2011 earthquake arises because our dynamic factor model is linear.
Table 6  The mean data-revisions effect for “Hard” variables

| Variable name            | Episodes                                | March 2011 quake, target growth is 2011:Q1 to Q2 | COVID lockdown, target growth is 2020:Q1 to Q2 |
|--------------------------|-----------------------------------------|-----------------------------------------------|-----------------------------------------------|
|                          | Lehman crisis, target growth is 2008:Q4 to 2009:Q1 |                                               |                                               |
| Monthly earnings         | 0.25                                    | 0.06                                          | 0.15                                          |
| Capacity utilization     | 1.96                                    | 0.15                                          | 0.33                                          |
| Construction orders      | 0.09                                    | n.a.                                         | n.a.                                          |
| Housing starts           | n.a.                                    | n.a.                                         | n.a.                                          |
| Machinery orders         | 0.00                                    | 0.66                                          | 0.04                                          |
| Unemployment             | 0.30                                    | 0.04                                          | n.a.                                          |
| Passenger car production | 0.03                                    | 0.00                                          | 0.04                                          |
| Vehicle sales            | 0.38                                    | n.a.                                         | n.a.                                          |
| Exports                  | 0.00                                    | 0.01                                          | 0.01                                          |
| Imports                  | 0.01                                    | 0.00                                          | 0.01                                          |
| IIP                      | 0.05                                    | 0.05                                          | 0.10                                          |
| Tertiary Industry Index  | 0.11                                    | 0.11                                          | 0.38                                          |
| Retail Sales Index       | 0.34                                    | 0.11                                          | 0.05                                          |
| Real GDP                 | 6.47                                    | 0.39                                          | 0.11                                          |

The ratio of the mean data-revisions effect to the mean new-observations effect. The mean data-revisions effect is defined as the mean absolute value of the data-revisions effect conditional on data revisions (see footnote 28 for a precise definition). The mean new-observations effect is defined similarly, as the mean absolute value of the new-observations effect conditional on new observations.
Very recently, several studies for the US GDP and the Euro-area GDP have shown that nowcasting performance improves if the model incorporates nonlinearities such as the stochastic volatility, Markov regime switching, fat tails, or combinations thereof. Our nowcast’s over-reaction would not have occurred if the model assumed the specific factor for car production to be more volatile or less persistent during the crisis. It remains an open question whether those nonlinearities can trick the model into behaving this way.

- The large and concentrated parameter-revisions effect, shown in Fig. 3, would not have occurred and the nowcast’s evolution would have been less jagged with Markov switching and perhaps with the stochastic volatility.

Extending the dynamic factor model along these lines is left for future research.

**Funding** The research reported here was supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science, and Technology of the Japanese government (Grant Number 16K03595(C)).

**Declarations**

**Conflict of interest** The authors have no conflict of interest to report.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

### Appendix 1. System matrices with temporal aggregation

In this appendix, we represent the temporally aggregated Mariano–Murasawa model as the state-space model (1) and (2) by choosing the system matrices appropriately.

Because the common factor appears in all observation equations, it is useful to define $\tilde{\omega}_j$, whose length doesn’t depend on $j$, as

$$
\tilde{\omega}_j^{\ell} \equiv \begin{pmatrix}
\omega_j^{\ell}, & 0, \ldots, 0
\end{pmatrix}, \quad \ell_{\text{max}} \equiv \max\{\ell_1, \ldots, \ell_p\}. \quad (A1.1)
$$

Then, (12) can be rewritten as

$$
y_{jt} = \mu_j + \beta_j^{(1)} \tilde{\omega}_j^{\ell} f_{t-1}^{(1)} + \beta_j^{(2)} \tilde{\omega}_j^{\ell} f_{t-1}^{(2)} + \omega_j^{(j)} u_{j, t-\ell_j}, \quad j = 1, 2, \ldots, p. \quad (A1.2)
$$

29 See, e.g., Marcellino et al. (2016), for a dynamic factor model featuring the stochastic volatility. For Markov-switching models, see, e.g., Barnett et al. (2016), Camacho et al. (2018), and more recently, Caldara et al. (2021). Antolin-Diaz et al. (2021) finds that forecast errors during the COVID lockdown for the USA are greatly reduced if the model features both the stochastic volatility and fat tails.
where

\[
\begin{bmatrix} f_t^{(1)} \\ f_t^{(1)} \\ \vdots \\ f_t^{(1)} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)} \quad \text{and} \quad \begin{bmatrix} f_t^{(2)} \\ f_t^{(2)} \\ \vdots \\ f_t^{(2)} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)}, \quad \begin{bmatrix} u_t^{(1)} \\ u_t^{(1)} \\ \vdots \\ u_t^{(1)} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)} \quad \text{and} \quad \begin{bmatrix} u_t^{(2)} \\ u_t^{(2)} \\ \vdots \\ u_t^{(2)} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)}.
\]

(A1.3)

The model can be represented as the state-space model (1) and (2). The \( m \), the \( \alpha_t \), the \( r \), and the \( \eta_t \) are

\[
m = 2(1 + \ell_{\text{max}}) + \sum_{j=1}^{p} (1 + \ell_j), \quad \alpha_t \equiv \begin{bmatrix} f_t^{(1)} \\ f_t^{(2)} \\ \vdots \\ u_t^{(1)} \\ u_t^{(2)} \end{bmatrix}_{(m \times 1)}, \quad r = 2 + p, \quad \eta_t \equiv \begin{bmatrix} \xi_t^{(1)} \\ \xi_t^{(2)} \\ \vdots \\ \xi_t^{(p)} \end{bmatrix}_{(r \times 1)}.
\]

(A1.4)

The system matrices are as follows. Start with the observation equations (1). We can write the \( p \) temporally aggregated equations (A1.2) as (1) if we choose

\[
\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}_{(p \times 1)}, \quad \begin{bmatrix} \beta_{1,1}^{(1)} \\ \beta_{1,2}^{(1)} \\ \vdots \\ \beta_{p,1}^{(1)} \end{bmatrix}_{(p \times m)} \quad \text{and} \quad \begin{bmatrix} \beta_{1,1}^{(2)} \\ \beta_{1,2}^{(2)} \\ \vdots \\ \beta_{p,2}^{(2)} \end{bmatrix}_{(p \times m)} \quad \text{and} \quad \begin{bmatrix} \omega_1' \omega_2' \vdots \omega_p' \end{bmatrix}_{(p \times \sum_{j=1}^{p} (1+\ell_j))}.
\]

(A1.5)

Next, turn to the state equations (2). The augmented dynamics of \( f_t^{(i)} \equiv (f_t^{(i)}, \ldots, f_{t-\ell_{\text{max}}}^{(i)})' \) \( (i = 1, 2) \) implied by (5) can be written as

\[
\begin{bmatrix} f_t^{(1)} \\ f_t^{(2)} \\ \vdots \\ u_t^{(1)} \\ u_t^{(2)} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)} = \begin{bmatrix} \mathbf{Y}_1(\phi_f^{(i)}, \ell_{\text{max}}) f_{t-1}^{(i)} + \begin{bmatrix} \xi_t^{(i)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{\ell_{\text{max}} \times 1} \end{bmatrix}_{((1+\ell_{\text{max}}) \times 1)}
\]

(A1.6)
where
\[
\Upsilon_1(\theta, \ell) = \begin{bmatrix} \theta & 0 & 0 & \cdots & 0 \\ \text{1+\ell elements} & \mathbf{I}_\ell & \mathbf{0} \end{bmatrix}_{(1+\ell) \times (1+\ell)}, \quad \Upsilon_1(\theta, \ell) = \theta \text{ if } \ell = 0. \tag{A1.7}
\]

Similarly, the augmented dynamics of \(u^{(j)}_t = (u_{jt}, \ldots, u_{j,t-\ell_j})'\) implied by (6) can be written as
\[
u^{(j)}_t \equiv \begin{bmatrix} \xi_{jt} \\ \vdots \\ 0 \end{bmatrix}
\]
with \(\ell_j\) elements, \(j=1,2,\ldots,p\).
\[
(A1.8)
\]

Therefore, the dynamics of \(\alpha_t\), which collects \(f^{(1)}_t, f^{(2)}_t\), and \(u^{(j)}_t (j = 1, 2, \ldots, p)\), can be written as (2) with
\[
T \equiv \text{diag} \left( \Upsilon_1(\phi_1, \ell_{\text{max}}), \Upsilon_1(\phi_2, \ell_{\text{max}}), \Upsilon_1(\phi_3, \ell_1), \ldots, \Upsilon_1(\phi_p, \ell_p) \right), \quad r = 2 + p, \quad R \equiv \text{diag}(\mathbf{I}_2 \otimes \mathbf{e}_{\ell_{\text{max}}}, \mathbf{e}_{\ell_1}, \ldots, \mathbf{e}_{\ell_p}), \quad Q \equiv \text{diag}(\sigma_1^2, \ldots, \sigma_p^2),
\]
\[
(A1.9)
\]

where
\[
\mathbf{e}_\ell \equiv \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{\(\ell\) elements.} \quad \mathbf{e}_\ell = 1 \text{ if } \ell = 0. \tag{A1.10}
\]

**Appendix 2. Adaptation to missing values**

This appendix shows how state-space models can be modified to accommodate missing values. Standard textbooks (Section 3.4.7 of Harvey (1990) and Section 4.10 of Durbin and Koopman (2012)) provide a procedure for adapting Kalman filtering/smoothing and the likelihood function to missing values. However, (as far as we are aware) there is available no full derivation of the procedure. Mariano and Murasawa (2003) provide only a sketch of it. The description in Camacho and Perez-Quiros (2010) is more detailed, but it does not provide a proof of the validity of the procedure. This appendix provides a proof.
2.1. The case of no missing values

The material for the case of no missing values is entirely standard. We will present relevant formulas without derivation. Those formulas will be modified later on to accommodate missing values.

The state-space model is

\[
\begin{align*}
\text{observation equation} & \quad y_t = d_t + Z_t \alpha_t + \epsilon_t, \{\epsilon_t\} \\
\text{state equation} & \quad \alpha_t = T_t \alpha_{t-1} + R_t \eta_t, \{\eta_t\}
\end{align*}
\]

where \( \epsilon_t \) is the vector of observation errors. The system matrices, \((d_t, Z_t, H_t, T_t, R_t, Q_t)\), are subscripted by \( t \) to allow them to possibly depend deterministically on time. For the unobservable state vector \( \alpha_t \), define its prediction \( \alpha_{t|\tau} \) and its prediction error variance \( P_{t|\tau} \) as \(^{30}\)

\[
\begin{align*}
\alpha_{t|\tau} & \equiv \text{E}(\alpha_t | y_1, y_2, \ldots, y_{\tau}) \text{ so, for example, } \alpha_{t|t-1} \equiv \text{E}(\alpha_t | y_1, y_2, \ldots, y_{t-1}), \\
P_{t|\tau} & \equiv \text{Var}(\alpha_t | y_1, y_2, \ldots, y_{\tau}) \text{ so, for example, } P_{t|t-1} \equiv \text{Var}(\alpha_t | y_1, y_2, \ldots, y_{t-1}).
\end{align*}
\]

The Kalman filtering recursion is

\[
\begin{align*}
\alpha_{t|t} = & \quad \alpha_{t|t-1} + P_{t|t-1} Z_t' F_{t}^{-1} v_t, \quad (m \times 1) \\
P_{t|t} = & \quad P_{t|t-1} - P_{t|t-1} Z_t' F_{t}^{-1} Z_t, \quad (m \times m)
\end{align*}
\]

where

\[
\begin{align*}
v_t & \equiv y_t - (d_t + Z_t \alpha_{t|t-1}), \quad (p \times 1) \\
F_t & \equiv Z_t P_{t|t-1} Z_t' + H_t, \quad (p \times p)
\end{align*}
\]

\(^{30}\) In Durbin and Koopman (2012), \( a_{t|\tau} \equiv \alpha_{t|\tau}, \alpha_t \equiv \alpha_{t|t-1}, \hat{\alpha}_t \equiv \alpha_{t|t}, P_t \equiv P_{t|t-1}, \) and \( V_t \equiv P_{t|t} - \).
A by-product of this recursion is \( \{\alpha_t|t, P_t|t\}_{t=1}^n \) which will be useful in Kalman smoothing below. To kick off the recursion, the initial value, \( (\alpha_{1|0}, P_{1|0}) \), needs to be provided. A popular choice, if the matrices \( (T, R_t) \) are time-invariant, is the unconditional mean and variance of the state vector:\footnote{The “vec” operator in (A2.9) stacks the columns of the matrix with the columns ordered from left to right.}

\[
\alpha_{1|0} = 0, \quad \text{vec}(P_{1|0}) = (I - T \otimes T)^{-1} \text{vec}(RQR').
\]  

The likelihood of the sample \( (y_1, y_2, \ldots, y_n) \) can be readily derived by what is called the prediction error formula:

\[
\log L = -\frac{1}{2} \sum_{t=1}^{n} \left[ p \log(2\pi) + \log |F_t| + v_t^T F_t^{-1} v_t \right].
\]  

Kalman smoothing is a recursion for generating \( \{\alpha_t|n, P_t|n\} \) \( t = n-1, n-2, \ldots, 1 \).

\[
\begin{align*}
\alpha_t|n &= \alpha_t|t + P_t|t^T (\alpha_{t+1|n} - T_{t+1|t} \alpha_t|t), \quad t = n-1, n-2, \ldots, 1. \\
\text{where} \quad P_t|n &= P_t|t + P_t|t^T (P_{t+1|n} - P_{t+1|t} P_t|t)^{-1} P_t|t, \quad t = n-1, n-2, \ldots, 1.
\end{align*}
\]  

The recursion proceeds backward from \( t = n \), taking the sequence \( \{\alpha_t|t, P_t|t, P_t|t-1\}_{t=1}^n \) (which are, as noted above, available from Kalman filtering) as given.

### 2.2. Accommodating missing values

We now show how the above formulas can be modified to accommodate missing values in \( y_t \). Data for nowcasting necessarily include missing cases due to mixed frequencies. The textbook procedure for accommodating missing values, which we follow in this paper, has (to our knowledge) not been shown to produce the correct likelihood function.\footnote{See Section 3.4.7 of Harvey (1990) or Section 4.10 of Durbin and Koopman (2012) for an exposition of the procedure. Neither of them provides a proof that the procedure yields the correct likelihood function.} The key idea is to replace missing values by some noise that is totally unrelated to the system. Clearly, the state equation (A2.2), which is about the unobservable state vector, is not affected. What is affected is the observation equation (A2.1).

Allow some elements of \( y_t \) to be unobservable. Let \( p_t (\leq p) \) be the number of observed elements and \( S_t \) be the selection matrix. Here is an example. \( p = 3 \) and the
first and third elements of \( y_t \) are missing. Then \( p_t = 1 \) and

\[
S_t y_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{so} \quad S_t y_t = \begin{bmatrix} 0 \\ y_{2t} \\ 0 \end{bmatrix}.
\]

\( S_t y_t \) is observable, but is not jointly normal. Consider the observation vector, \( \tilde{y}_t \), in which zeros are replaced by some noise that is normally distributed:

\[
\tilde{y}_t = S_t y_t + (I_p - S_t) \vec{e}_t, \quad \vec{e}_t \sim \mathcal{N}(0, \sigma_e^2 \cdot I_p)
\]

(A2.12)

The noise \( \vec{e}_t \) is assumed to be independent of the model’s shocks (\( \vec{e}_s, \eta_s \)) for all \((t, s)\). In the above example,

\[
S_t y_t = \begin{bmatrix} 0 \\ y_{2t} \\ 0 \end{bmatrix}, \quad \text{but} \quad S_t y_t + (I - S_t) \vec{e}_t = \begin{bmatrix} e_{1t} \\ y_{2t} \\ e_{3t} \end{bmatrix}.
\]

Substitute (A2.1) into (A2.12) to obtain the observation equation with missing values:

\[
\text{(observation equation)} \quad \tilde{y}_t = \tilde{d}_t + \tilde{Z}_t + \alpha_t + \tilde{\epsilon}_t, \quad \{\tilde{\epsilon}_t\}
\]

white noise with \( \tilde{\epsilon}_t \sim \mathcal{N}(0, \tilde{H}_t) \),

(A2.13)

where

\[
\tilde{d}_t \equiv S_t d_t, \quad \tilde{Z}_t \equiv S_t Z_t, \quad \tilde{\epsilon}_t \equiv S_t \vec{e}_t + (I_p - S_t) \vec{e}_t, \quad \tilde{H}_t \equiv S_t H_t S_t + \sigma_e^2 \cdot (I_p - S_t).
\]

(A2.14)

The observation equation is now (A2.13) rather than (A2.1). In the above Kalman filtering equations, the prediction equations, (A2.8a) and (A2.8b), are not affected by missing values. To adapt the updating equations, (A2.5a) and (A2.5b), to missing values, (i) replace “\( y_t \)” by “\( \tilde{y}_t \)”, “\( d_t \)” by “\( \tilde{d}_t \)”, “\( Z_t \)” by “\( \tilde{Z}_t \)”, and “\( H_t \)” by “\( \tilde{H}_t \)” in (A2.5a) and (A2.5b), and then (ii) substitute (A2.12) and (A2.14) for \( \tilde{y}_t, \tilde{d}_t, \tilde{Z}_t, \) and \( \tilde{H}_t \). This results in

\[
\text{(updating)} \quad \begin{cases}
\alpha_{t|t} = \alpha_{t|t-1} + P_{t|t-1} Z_t, & \quad S_t F_t^{-1} v_t, \\
P_{t|t} = P_{t|t-1} - P_{t|t-1} Z_t S_t F_t^{-1} S_t Z_t P_{t|t-1},
\end{cases}
\]

(A2.15a)

(A2.15b)
where

\[ v_t = S_t y_t + (I - S_t) e_t - S_t (d_t + Z_t \alpha_{t|t-1}) \]  

(A2.16)

\[ F_t = S_t Z_t P_{t|t-1} Z_t' S_t + S_t H_t S_t + \sigma_e^2 (I_p - S_t) \]  

(A2.17)

We show that \( e_t \) and \( \sigma_e^2 \) can be eliminated from Kalman filtering. We do this in three steps.

- Given \( S_t \), choose a permutation matrix \( C_t \), an identity matrix with its rows resequenced in some order, such that

\[
\begin{bmatrix}
W_t \\
0
\end{bmatrix}_{(p \times p)} \begin{bmatrix}
0 \\
U_t
\end{bmatrix}_{((p - p_t) \times p)} = \begin{bmatrix}
0 \\
U_t
\end{bmatrix}_{((p - p_t) \times p)}
\]

where \( W_t \) is a subset of rows from \( I_p \). Permutation matrices are unitary matrices (so \( C_t C_t' = I \)). Other well-known properties are:

(a) \[
C_t (I - S_t) = \begin{bmatrix}
0 \\
U_t
\end{bmatrix}_{((p - p_t) \times p)},
\]

where \( U_t \) is the remaining subset of rows from \( I_p \) not included in \( W_t \).

(b) (A2.18) and (a) immediately imply

\[
C_t = \begin{bmatrix}
W_t' \\
U_t'
\end{bmatrix}.
\]

(c) \( W_t W_t' = I \), \( W_t U_t' = 0 \) \( (p_t \times (p - p_t)) \), and \( U_t U_t' = I \), because \( C_t \) is a unitary matrix.

The above matrix \( W_t \) is a “compression matrix” because it extracts from \( y_t \) only non-missing elements. In the above example with \( S_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \),

\[
C_t = \begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}, \quad W_t = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \quad U_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_t y_t = [y_{2t}].
\]
Structure of $F_t$. From (A2.17),

$$C_t F_t C_t' = C_t S_t (Z_t P_{t|t-1} Z_t' + H_t) S_t C_t' + \alpha^2_t C_t (I - S_t) C_t'.$$

\begin{align*}
(A) &= \begin{bmatrix} W_t & 0 \end{bmatrix} \begin{bmatrix} Z_t & P_{t|t-1} Z_t' + H_t \end{bmatrix} \begin{bmatrix} W_t' & 0 \end{bmatrix} \\
&= \begin{bmatrix} \bar{F}_t & 0 \\
(p_t \times p_t) & (p_t \times (p_t - p_t))
\end{bmatrix},
\end{align*}

where

$$\bar{F}_t \equiv W_t Z_t P_{t|t-1} Z_t' W_t' + W_t H_t W_t'. \quad \text{(A2.19)}$$

\begin{align*}
(B) &= \begin{bmatrix} 0 & U_t \end{bmatrix} \begin{bmatrix} W_t' & U_t' \end{bmatrix} \\
&= \begin{bmatrix} 0 & (p_t \times p_t) \\
(p_t \times (p_t - p_t)) & (p_t \times (p_t - p_t))
\end{bmatrix},
\end{align*}

by (a) and (b).

Therefore,

$$C_t F_t C_t' = D_t \quad \text{with} \quad D_t \equiv \begin{bmatrix} \bar{F}_t & 0 \\
(p_t \times p_t) & (p_t \times (p_t - p_t))
\end{bmatrix},$$

so

$$F_t^{-1} = C_t' D_t^{-1} C_t \quad \text{with} \quad D_t^{-1} = \begin{bmatrix} \bar{F}_t^{-1} & 0 \\
(p_t \times p_t) & (p_t \times (p_t - p_t))
\end{bmatrix} \begin{bmatrix} 1 & \alpha^2_t I_{p_t - p_t} \\
((p_t - p_t) \times p_t) & ((p_t - p_t) \times (p_t - p_t))
\end{bmatrix}. \quad \text{(A2.20)}$$

Given this expression (A2.20) about the structure of $F_t$, and using (A2.18) and (a), we can easily derive the following:

\begin{align*}
S_t F_t^{-1} S_t &= \begin{bmatrix} W_t' & \bar{F}_t^{-1} & W_t \end{bmatrix} \begin{bmatrix} p_t \times p_t & p_t \times (p_t - p_t) & p_t \times p_t \\
(p_t \times p_t) & (p_t \times (p_t - p_t)) & (p_t \times p_t)
\end{bmatrix} \\
S_t F_t^{-1} (I - S_t) &= \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} p_t \times p_t \\
(p_t \times p_t)
\end{bmatrix}. \quad \text{(A2.21a, b)}
\end{align*}
\[
(\mathbf{I} - \mathbf{S}_t) \mathbf{F}_t^{-1} (\mathbf{I} - \mathbf{S}_t) = \frac{1}{\sigma^2} \sum_{t} \mathbf{U}_t' \mathbf{U}_t.
\] (A2.21c)

- We show below that the updating equations (A2.15a) and (A2.15b) with (A2.16) and (A2.17) become

\[
\begin{aligned}
\alpha_{t|t} &= \alpha_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{W}'_t \mathbf{F}_t^{-1} \mathbf{v}_t, \\
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{W}'_t \mathbf{F}_t^{-1} \mathbf{W}_t \mathbf{Z}_t \mathbf{P}_{t|t-1},
\end{aligned}
\] (A2.22a)

with

\[
\mathbf{v}_t \equiv \mathbf{W}_t \mathbf{y}_t - \mathbf{W}_t \mathbf{d}_t - \mathbf{W}_t \mathbf{Z}_t \alpha_{t|t-1}
\] (A2.23)

and \(\mathbf{F}_t\) given by (A2.19).

**Proof** (A2.15b) becomes (A2.22b) because of (A2.21a). (A2.15a) becomes (A2.22a) because

\[
\begin{aligned}
\mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{S}_t \mathbf{F}_t^{-1} \mathbf{v}_t &= \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{S}_t \mathbf{F}_t^{-1} [\mathbf{S}_t \mathbf{y}_t + (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t - \mathbf{S}_t (\mathbf{d}_t + \mathbf{Z}_t \alpha_{t|t-1})] \\
&= \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{S}_t \mathbf{F}_t^{-1} \mathbf{S}_t (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \alpha_{t|t-1}) + \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{S}_t \mathbf{F}_t^{-1} (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \\
&= \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{S}_t \mathbf{F}_t^{-1} \mathbf{S}_t (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \alpha_{t|t-1}) \\
&= \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{W}'_t \mathbf{F}_t^{-1} \mathbf{W}_t (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \alpha_{t|t-1}) \\
&= \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{W}'_t \mathbf{F}_t^{-1} \mathbf{v}_t.
\end{aligned}
\]

- Mechanically, to modify the original updating equations, (A2.5a) and (A2.5b) with (A2.6) and (A2.7), to missing values, all we need to do is replace “\(\mathbf{y}_t\)” by “\(\mathbf{W}_t \mathbf{y}_t\)”, “\(\mathbf{d}_t\)” by “\(\mathbf{W}_t \mathbf{d}_t\)”, “\(\mathbf{Z}_t\)” by “\(\mathbf{W}_t \mathbf{Z}_t\)”, and “\(\mathbf{H}_t\)” by “\(\mathbf{W}_t \mathbf{H}_t \mathbf{W}'_t\).” This point is made, without proof, in Section 3.4.7 of Harvey (1990) and Section 4.10 of Durbin and Koopman (2012).

The likelihood function, (A2.10), is written in terms of \(\mathbf{v}_t\) and \(\mathbf{F}_t\). Below we rewrite it in terms of \(\mathbf{v}_t\) and \(\mathbf{F}_t\).

- Show: \(\log |\mathbf{F}_t| = \log |\mathbf{\bar{F}}_t| + (p - p_t) \log (\sigma_e^2)\).

**Proof**

\[
\log |\mathbf{F}_t| = -\log |\mathbf{F}_t^{-1}|
\]
\[
= -\log |\mathbf{C}'_t \text{diag}(\mathbf{\bar{F}}_t^{-1}, \frac{1}{\sigma^2} \mathbf{I}_{p - p_t}) \mathbf{C}_t| \quad \text{(by A2.20)}
\]
\[
= -\log |\text{diag}(\mathbf{\bar{F}}_t^{-1}, \frac{1}{\sigma^2} \mathbf{I}_{p - p_t})| \quad |\mathbf{C}'_t \mathbf{C}_t|
\]
Nowcasting Japan’s GDP

\[= -\log \left| \text{diag} \left( \mathbf{F}_t^{-1}, \frac{1}{\sigma^2_e} \mathbf{I}_{p-p_t} \right) \right| \]  (because \( \mathbf{C}_t \) is a unitary matrix)

\[= -\log |\mathbf{F}_t^{-1}| + (p - p_t) \log(\sigma^2_e) \]

\[= \log |\mathbf{F}_t| + (p - p_t) \log(\sigma^2_e). \]

- Rewrite \( \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t \) in (A2.10).

\[\mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t = \left[ \mathbf{S}_t \mathbf{y}_t - \mathbf{S}_t \mathbf{d}_t - \mathbf{S}_t \mathbf{Z}_t \mathbf{a}_t \right] + (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \]

\[\mathbf{F}_t^{-1} \left[ \mathbf{S}_t \mathbf{y}_t - \mathbf{S}_t \mathbf{d}_t - \mathbf{S}_t \mathbf{Z}_t \mathbf{a}_t \right] + (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \]  (by (A2.16))

\[= \mathbf{F}_t^{-1} \left[ \mathbf{S}_t (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \mathbf{a}_t) + (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \right] \]

\[= \mathbf{F}_t^{-1} \left[ \mathbf{S}_t (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \mathbf{a}_t) + (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \right] \]

\[+ 2 \left( \mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \mathbf{a}_t \right)' \mathbf{F}_t^{-1} (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t + \epsilon_t' (\mathbf{I} - \mathbf{S}_t) \mathbf{F}_t^{-1} (\mathbf{I} - \mathbf{S}_t) \mathbf{e}_t \]  \hspace{1cm} (1)

By (A2.21a),

\[= (\mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \mathbf{a}_t)' \mathbf{W}_t' \mathbf{F}_t^{-1} \mathbf{W}_t \left( \mathbf{y}_t - \mathbf{d}_t - \mathbf{Z}_t \mathbf{a}_t \right) = \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t, \]  \hspace{1cm} (A2.24)

where \( \mathbf{v}_t \) is defined in (A2.23). By (A2.21b), the term (2) vanishes. By (A2.21c),

\[= \frac{1}{\sigma^2_e} \epsilon_t' \mathbf{U}_t' \mathbf{U}_t \epsilon_t. \]  \hspace{1cm} (A2.25)

- So the likelihood function (A2.10) becomes

\[\log L = -\frac{1}{2} \sum_{t=1}^{n} \left[ p \log(2\pi) + \log |\mathbf{F}_t| + \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t \right] \]

\[\text{depends on model parameters but not on } (\sigma^2_e, \epsilon_t) \]  \hspace{1cm} (A2.26)

\[= -\frac{1}{2} \sum_{t=1}^{n} \left[ (p - p_t) \log(\sigma^2_e) + \frac{1}{\sigma^2_e} \epsilon_t' \mathbf{U}_t' \mathbf{U}_t \epsilon_t \right] \]

\[\text{depends on } (\sigma^2_e, \epsilon_t) \text{ but not on model parameters} \]

Since the second summation does not depend on the model parameters, the maximum likelihood estimate of the model parameters can be obtained by maximizing the first summation.

- Again, mechanically, all we need to do to adapt the likelihood function (A2.10) to missing cases is replace “\( \mathbf{v}_t \)” by “\( \mathbf{v}_t \)” and “\( \mathbf{F}_t \)” by “\( \mathbf{F}_t \)”.
Appendix 3. Details on numerical maximization of the likelihood function

A detailed description of our estimation procedure is as follows. The model parameters are: \((\mu^*, \beta^{(1)}, \beta^{(2)}, \phi^{(1)}, \phi^{(2)}, \phi_1, \phi_2, \ldots, \phi_p, \sigma^2_1, \sigma^2_2, \ldots, \sigma^2_p)\) where \(\mu^* = (\mu^*_1, \ldots, \mu^*_p)\) and \(\beta^{(i)} = (\beta^{(i)}_1, \ldots, \beta^{(i)}_p)\), \(i = 1, 2\). \(\mu^*\) is estimated as follows. We first estimate \(\mu_j (j = 1, 2, \ldots, p)\) by the sample mean of \(y_{jt}\). \(\mu^*\) is then recovered by the relation (13). The remaining parameters are estimated by maximum likelihood. The procedure for obtaining the initial parameter value, similar to the one in Giannone et al. (2008), is as follows. (i) Using monthly series only, extract two leading factors by the principal component analysis. (ii) The common factors’ AR(1) parameters \((\phi^{(1)}, \phi^{(2)})\) are estimated from the factors thus obtained, \((\beta^{(1)}, \beta^{(2)}_j) (j = 1, 2, \ldots, p)\) from regressing \(y_{jt}\) on the common factors, and the remaining parameters from the OLS residuals. (iii) Estimate the model’s common factors from the whole sample including both monthly and quarterly series by Kalman filtering and smoothing. (iv) Redo (ii), this time with the estimate of the common factors obtained from (iii). This last step produces our initial parameter values.

Maximization of the likelihood function starting with this initial parameter value is done on the programming language R using its “ucminf” package. The package stops when the line length of its line search is close to zero. The parameter estimate is then passed on to the BFGS option of R’s “optim” package, which for all cases resulted in convergence.

References

Anderson RG (2006) Replicability, real-time data, and the science of economic research: FRED, ALFRED, and VDC. Federal Reserve Bank Saint Louis Rev 88:81
Andreou E, Gagliardini P, Ghysels E, Rubin M (2019) Inference in group factor models with an application to mixed-frequency data. Econometrica 87:1267–1305
Antolin-Diaz J, Drechsel T, Petrella I (2021) Advances in nowcasting economic activity: secular trends, large shocks and new data. CEPR Discussion Paper No. DP15926
Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70:191–221
Baibura M, Modugno M (2014) Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. J Appl Econom 29:133–160
Baibura M, Giannone D, Modugno M, Reichlin L (2013) Now-casting and the real-time data flow. In: Handbook of economic forecasting, vol 2. Elsevier, pp 195–237
Barnett WA, Chauvet M, Leiva-Leon D (2016) Real-time nowcasting of nominal GDP with structural breaks. J Econom 191:312–324
Bok B, Caratelli D, Giannone D, Sbordone AM, Tambalotti A (2018) Macroeconomic nowcasting and forecasting with big data. Annu Rev Econ 10:615–643. https://doi.org/10.1146/annurev-economics-080217-053214
Bragoli D (2017) Now-casting the Japanese economy. Int J Forecast 33:390–402
Caldara D, Cascaldi-Garcia D, Cuba-Borda P, Loria F (2021) Understanding growth-at-risk: a Markov-switching approach. SSRN. https://ssrn.com/abstract=3992793
Camacho M, Perez-Quirios G (2010) Introducing the euro-sting: short-term indicator of euro area growth. J Appl Econom 25:663–694
Camacho M, Perez-Quirios G, Poncela P (2018) Markov-switching dynamic factor models in real time. Int J Forecast 34:598–611
Cascaldi-Garcia D, Ferreira Thiago RT, Giannone D, Modugno M (2021) Back to the present: learning about the euro area through a now-casting model. International Finance Discussion Paper
Castle JL, Ellis C (2002) Building a real-time database for GDP (E). Bank of England Quarterly Bulletin, vol 42
Chikamatsu K, Hirakata N, Kido Y, Otaka K (2021) Mixed-frequency approaches to nowcasting GDP: an application to Japan. Jpn. World Econ., vol 57(C)
Croushore D (2011) Frontiers of real-time data analysis. J Econ Lit 49:72–100
Croushore D, Stark T (2001) A real-time data set for macroeconomists. J Econom 105:111–130
Diebold F, Mariano R (1995) Comparing predictive accuracy. J Bus Econ Stat 20:253–263
Durbin J, Koopman SJ (2012) Time series analysis by state space methods. Oxford University Press
Giannone D, Reichlin L, Small D (2008) Nowcasting: the real-time informational content of macroeconomic data. J Monet Econ 55:665–676
Giannone D, Henry J, Lalik M, Modugno M (2012) An area-wide real-time database for the euro area. Rev Econ Stud 94:1000–1013
Hara N, Yamane S (2013) New monthly estimation approach for Nowcasting GDP growth: the case of Japan. Bank of Japan Working Paper Series 13-E14, Bank of Japan
Harvey AC (1990) Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge
Hayashi F, Yuta T (2021) The nowcast revision analysis extended. Econ Lett. https://doi.org/10.1016/j.econlet.2021.110112
Higgins PC (2014) GDPNow: a model for GDP’Nowcasting. FRB Atlanta Working Paper 2014-7, Federal Reserve Bank of Atlanta
Liebmann J (2014) Real-time nowcasting of GDP: a factor model vs. professional forecasters. Oxf Bull Econ Stat 76:783–811
Marcellino M, Porqueddu M, Venditti F (2016) Short-term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility. J Bus Econ Stat 34:118–127
Mariano RS, Murasawa Y (2003) A new coincident index of business cycles based on monthly and quarterly series. J Appl Econom 18:427–443
Sims CA (2002) The role of models and probabilities in the monetary policy process. Brook Pap Econ Act 2002:1–40
Urasawa S (2014) Real-time GDP forecasting for Japan: a dynamic factor model approach. J Jpn Int Econ 34:116–134
Vermeulen P (2012) Quantifying the qualitative responses of the output purchasing managers index in the US and the Euro area. Working Paper Series No. 1417, European Central Bank
Zwijnenburg J (2015) Revisions of quarterly GDP in selected OECD countries. OECD Stat Brief 22:12

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.