Bank of Japan Review 2022-E-1

Use of Alternative Data in the Bank of Japan's Research Activities

Research and Statistics Department
KAMEDA Seisaku

January 2022

The Bank of Japan (BOJ) has recently launched a new page on its website titled, "Alternative Data Analysis."[1] In light of the launch of this page, this paper outlines initiatives taken by the BOJ's research divisions (including but not limited to the Research and Statistics Department) in the field of alternative data analysis. Since the spread of COVID-19, the BOJ has been making active use of high-frequency data — such as mobility trends based on location data — in assessing economic conditions to conduct monetary policy. Moreover, in light of the lessons from the Global Financial Crisis of the late 2000s, the BOJ also has been continuing its efforts to strengthen the collection and use of various individual transaction data in the financial field. Such new forms of big data are called alternative data as opposed to traditional economic and financial statistics. The alternative data employed at the BOJ have been wide-ranging, including high-frequency data, textual data, and granular data; recently, the range of these data has been extending further to cover, for example, climate-related data.

What is Alternative Data?

In general, "alternative data" is a collective term for big data that have become available following the emergence of new sources or means of data collection due to recent technological innovations and advances in digitalization.[2] These alternative data are called thus in order to differentiate them from traditional statistical data—for example, macroeconomic indicators such as GDP statistics and disclosed data on listed firms' financial results—and as such are also referred to as nontraditional data. Typical examples of alternative data are mobility data (population data by geographical location and time), which are compiled using information on location collected from such sources as smartphones, and consumption data based on records of payments made by credit cards or mobile payment services.

High-Frequency Data Analysis during the COVID-19 Pandemic

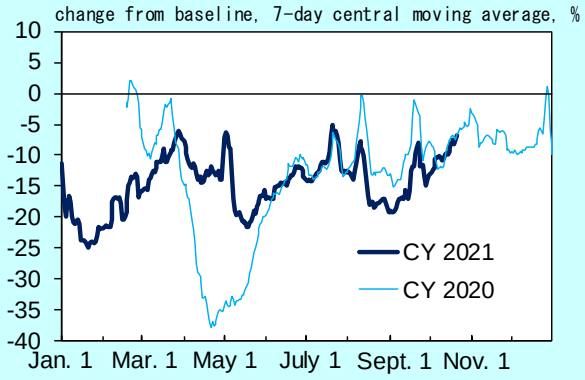
Benefits of using high-frequency data

Alternative data have become widely used in economic and financial research and analysis due in part to the COVID-19 pandemic. After the onset of the pandemic, consumers and firms practiced self-restraint as part of their response in guarding against COVID-19. In addition, strict public health measures were in place when the disease was spreading. These factors rapidly and adversely affected economic activity, particularly consumption in face-to-face services such as dining-out and accommodation, both at home and abroad. Moreover, short-term fluctuations in economic activity intensified depending on the spread of COVID-19. As these developments indicate, the pandemic has accelerated changes in economic conditions to a considerable degree, and this has increased the benefits of using alternative data that provide more up-to-date information than traditional statistical data are able to, particularly as they are updated on a daily, weekly, or semimonthly basis (called "high-frequency data") in the assessment of economic activity.[3] In light of the increased need for such alternative data, some private firms have sped up their efforts to provide digital information associated with their business operations either free of charge or for a fee while ensuring, for example, the anonymity of the information they disseminate.[4] In the meantime, improvement has been made in the processing capacity of information infrastructure that handles massive amounts of data, such as cloud servers and high performance devices. In addition, there has been progress in the spread of technologies and knowledge that assist technical data analysis, as exemplified by the increased use of artificial intelligence (AI), machine learning, and natural language processing. These factors also seem to have accelerated the use of alternative data in the economic and social context.

High-frequency data analysis conducted at the Bank of Japan (BOJ)

Under the circumstances of the growing need for and availability of high-frequency data, the BOJ has been making active use of such data since the relatively early stages of the COVID-19 pandemic, employing them as one of the resources for its assessment of economic activity and prices. Specifically, the BOJ has started incorporating the use of high-frequency data into its *Outlook for Economic Activity and Prices* (Outlook Report) in the April 2020 issue. A box in this issue examined the following high-frequency data to capture recent mobility trends and consumption: (1) the number of visitors at World Cultural Heritage Sites in Japan, and (2) the nighttime population of selected downtown areas in Tokyo. In subsequent issues, the BOJ employed not only these data but also other various high-frequency data and conducted associated analyses. Examples of these data and analyses are as follows: (3) mobility trends related to retail and recreation and empirical analyses that associate these trends with the number of confirmed new cases of COVID-19 by prefecture (Chart 1); (4) the number of visitors to restaurants, based on information collected through a reservation management system provided by a private firm; (5) consumption developments based on credit card spending; and (6) the number of people at airports in Japan.[5] In particular, data described in (3), (4), and (5) above have been regarded by the BOJ as indicators to be periodically monitored for the time being and thus were presented in the "The Background" section of the April 2021 issue and in subsequent issues of the Outlook Report.

[Chart 1] Mobility Trends Based on Location Data



Notes: 1. The baseline is the median on the corresponding day of the week during the 5-week period from January 3 to February 6, 2020.
2. Figures are mobility trends for places such as restaurants, shopping centers, and theme parks.
3. The latest figure is the average for October 18-24.

Source: Google LLC "Google COVID-19 Community Mobility Reports." https://www.google.com/covid19/mobility/.

Obviously, high-frequency data are also applicable to the analysis of overseas economic developments during the COVID-19 pandemic.[6] In fact, in the Outlook Report, boxes on overseas economic developments often incorporate charts and analyses using high-frequency data on selected economies.[7] In addition, papers published as the *Bank of Japan Review Series* include (1) a study that analyzes the impact of COVID-19 on U.S. consumption using high-frequency state-level data and (2) a study that develops a "responsiveness indicator" that comprehensively captures the impact that COVID-19 has on, for example, household behavior by way of modeling big data such as on public health measures in overseas countries and mobility trends, using machine learning.[8]

It should be noted that high-frequency alternative data had already been employed at the BOJ for analytical purposes even before the COVID-19 pandemic. For example, the BOJ released in 2013 a research paper on nowcasting (predicting the current state of) the transaction value of travel using aggregate data on the number of search queries for selected words performed on a major search engine.[9] The motivation of this paper derives from the significant changes in the transaction value of travel that were observed around the time of the Great East Japan Earthquake in 2011. In other words, the need for the analysis that makes use of new high-frequency data was brought about by increased fluctuations in economic activity due to an exogenous shock, as has recently been the case under the COVID-19 pandemic. It can also be said that, thanks to the earlier studies including the above and the continued discussions that they have stimulated, both within the BOJ and with external experts or researchers, the BOJ was able to promptly incorporate various high-frequency data that has become available during the COVID-19 pandemic into its research and analysis.

Caveats in using high-frequency data

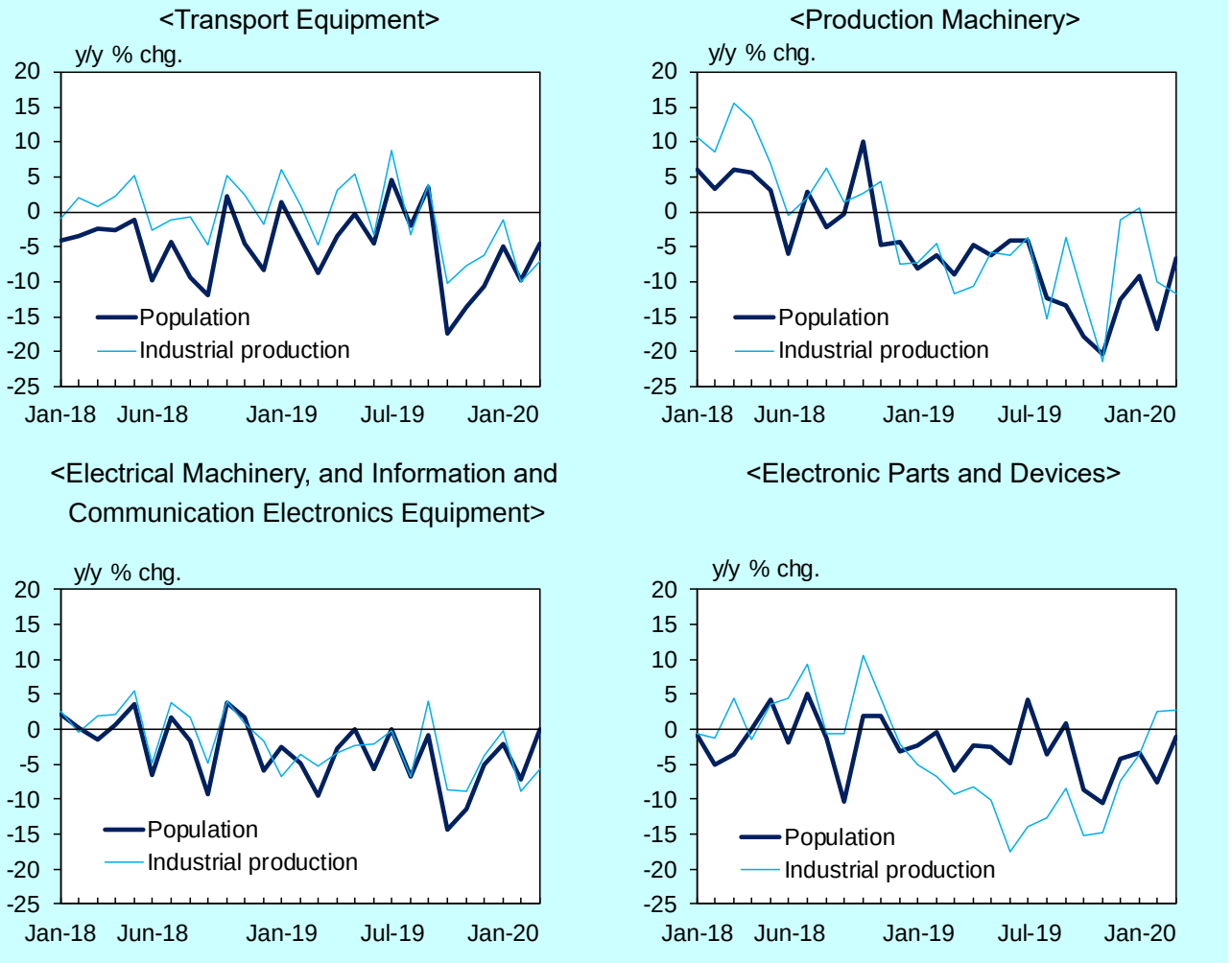
As explained, against the backdrop of increased short-term fluctuations in economic activity owing to the impact of COVID-19, high-frequency data have become increasingly useful. There are, however, specific caveats to their use, as follows.

First, many high-frequency data have begun to be made available around the time of the COVID-19 outbreak and thus have only a limited number of time series. One of the challenges associated with this is that making seasonal adjustments to such data often turns out to be difficult. Taking mobility data as an example, assume that there was an increase in mobility during Golden Week (the period from late April to early May

that includes several national holidays) in 2021. To accurately judge whether this merely reflects regular seasonal mobility patterns or can be regarded as evidence for the waning of the impact of COVID-19, it is desirable that several pre-pandemic years of mobility trends during Golden Week be examined. In addition, due to the short history of alternative data, not much has been revealed about their statistical tendencies; therefore, developments in these data tend to leave some margin of interpretation. This is in contrast to the situation concerning public statistical data, which have a longer history, since various bias regarding these data have been identified already, albeit to varying degrees; for example, the diffusion indices for employment conditions in the *Tankan* (Short-Term Economic Survey of Enterprises in Japan) have been known as tending to register temporary smaller negative figures for the June surveys, owing to the fact that these surveys are conducted just after new graduates enter the labor market. The limitation of alternative data in this regard is expected to be resolved eventually as more time-series data become available and more analysis and studies accumulate.

Second, with respect to many of high-frequency data, it has been difficult to clearly associate developments exhibited by them with specific economic activity. A typical example in this regard is when analyzing mobility data. For instance, people situated around a station in the evening as shown by these data might have been out for dinner in a near commercial district or about to take a train on their way home. Depending on whether these people were about to dine out or go home, the implications of their behavior for economic activity differ completely, including in terms of the extent of influence on private consumption. As this example suggests, interpreting high-frequency data through linking them with the value added from actual economic activity—such as consumption and production—requires some creative approaches. One such approach is to compare developments in high-frequency data with those in other economic data. For example, regarding the aforementioned data on mobility trends related to retail and recreation, as shown in Chart 1, the BOJ confirmed

[Chart 2] Production Indicator Based on the Hourly Population by Industry



Sources: Agoop; Ministry of Economy, Trade and Industry.

that the historical correlation coefficients between these data and selective expenditures according to the *Family Income and Expenditure Survey* conducted by the Ministry of Internal Affairs and Communications are very high—taking the values of 0.8-0.9—before incorporating them into its analyses in the Outlook Report and other occasions. Another example is drawn from a study conducted by BOJ economists in which they nowcast industry-level production as measured by the Indices of Industrial Production, using the number of people located at and around factories as a proxy variable for labor input.[10] As part of the study, the authors find that the accuracy of their nowcasting varies across industries; it is high for the industries of transport equipment, production machinery, and "electrical machinery, and information and communication electronics equipment," whereas it was low for the electronic parts and devices industry, which is less labor intensive (Chart 2).

Another approach to effectively interpret developments in high-frequency data is to extract and only analyze the datasets relevant for the purpose of analysis through the use of statistical techniques. In the study mentioned above, for example, the authors also address the nowcasting of sales in the eating and drinking industry. Though this is difficult using mobility data alone, primarily because the site areas of many restaurants are small, the authors succeed in improving the accuracy of their nowcasting of sales on that industry through the proper use of clustering techniques.

Sentiment Analysis Using Textual Data

Alternative data that can be used for economic analysis take a variety of forms besides high-frequency data. The following presents examples of how alternative data other than high-frequency data, such as textual data and granular data, are used in the BOJ's economic and financial analysis.[11]

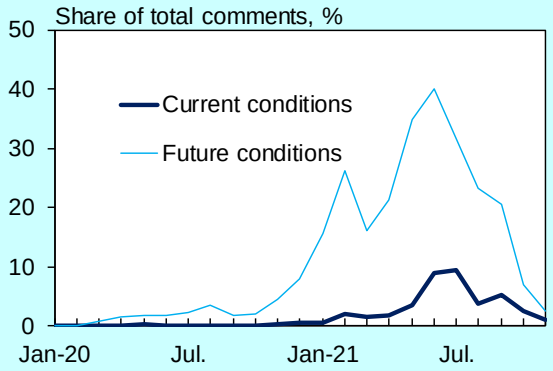
Textual data refers to words and sentences contained in, for example, public documents, reports, and media coverage. These data are regarded as a kind of alternative data because (1) they are completely different from quantitative data often used in conventional economic analysis, such as the economic growth rate and the inflation rate, and (2) many textual data are not provided by public entities but instead can be obtained through media, such as the internet and social networking sites.

In the case of the BOJ, several studies have been published by its staff on the analysis of textual data obtained from comments in the *Economy Watchers Survey*.[12] This survey, conducted by the Cabinet Office, provides suitable resources for text analysis since it has the following unique features: (1) its results are released in a timely manner; (2) it asks respondents, who hold jobs that enable them to closely watch developments in economic activity—for example, corporate managers and supermarket clerks—to give their assessments of current and future economic conditions; and (3) it also collects respondents' comments on the issues that they focus on when assessing economic conditions, which are then organized and published on the Cabinet Office's website.

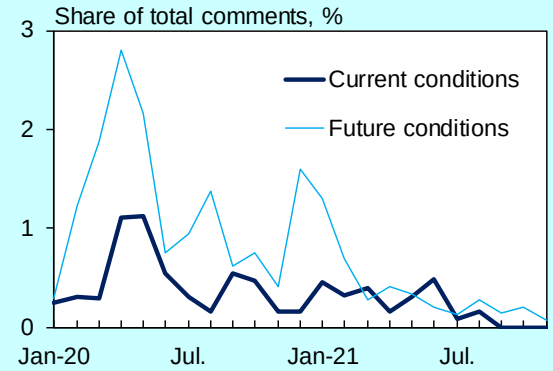
Turning to the specific content of the studies made at the BOJ, to explore the factors behind changes in the sentiment of *Economy Watchers Survey* respondents, some authors examine the frequency in appearance of certain words in respondents' comments—for example, "COVID-19," "vaccine," "bankruptcy (or business closure)," and "semiconductor" in the case of the recent study—and track changes in such frequency (Chart 3). In addition, the authors also focus on the frequency at

[Chart 3] Economy Watchers Survey: Share of comments including the term

< "Vaccine" >



< "Bankruptcy" >



Note: The shares are calculated from the Economy Watchers Survey.
Source: Cabinet Office.

[Chart 4] Co-occurrence Network Diagram

Note: The co-occurrence network diagram constructed from the Economy Watchers Survey (comments for the current conditions DI in June 2021).
Source: Cabinet Office.

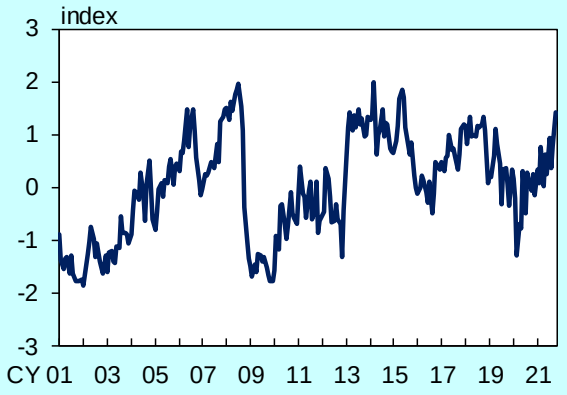
which several words are used in the same comment or context. In addition, some authors focus on the frequency at which particular words are used in the same comment or context. In this way, they construct a respective "co-occurrence network diagram" by extracting and visualizing the co-occurrence relationship between these words, using machine learning (Chart 4). The authors make use of the finding from the diagram to examine firm and consumer sentiment, their views on price changes, and economic developments lying behind these sectors' mindset.[13]

Furthermore, some of the authors construct a respective Price Sentiment Index (PSI), which is calculated as the share of comments implying inflation minus the share of comments implying deflation, by classifying *Economy Watchers Survey* comments into four types in terms implied price developments: inflation; deflation; zero inflation; and no mention of price developments (Chart 5).[14] This again uses machine learning. Examining developments in the PSI and comments that contribute to fluctuations in the index provides useful information concerning firms' perceptions of price developments as well as their short-term inflation expectations.

Financial Analysis Using Granular Data

Granular data often refer to microdata with particularly high granularity, such as those on loans from Bank A to Company B and on specific issues of government bonds. While granular data have been used in the analysis of the real economy, for example, in examining the financial conditions of individual firms, it is particularly of note that these data have also been actively employed in the research and analysis of financial markets and financial system stability, with their availability increasing in the fields related to financial markets and financial institutions in recent years.

[Chart 5] Price Sentiment Index (PSI)



Note: For more information on how the Price Sentiment Index (PSI) is constructed, please refer to the document in Footnote 14.

Source: Cabinet Office.

In the field of finance, the collection and use of granular data gained further momentum worldwide in the wake of Global Financial Crisis (GFC) in the late 2000s and the subsequent European debt crisis.[15] At the time of the GFC, banks exhibited excessive risk-taking behaviors in their lending and securities investments. Moreover, even nonbanks in their short and long term positions were highly leveraged with a small amount of cash reserves by actively engaging in, for example, repo and over-the-counter (OTC) derivative transactions. It has been pointed out that this factor exacerbated the crisis and also made the transmission channels of shocks complex. In light of such lessons, international organizations, governments, central banks, authorities in charge of financial and regulation supervision have cooperated and taken steps to strengthen the collection and use of data on various individual financial transactions. In addition, to restore confidence in the financial system after the GFC, a simultaneous stress testing of large banks was introduced in the United States in 2011. Since then, such testing has been increasingly implemented in major Western countries as a measure to examine financial institutions' profitability and capital adequacy relative to risk.[16] These developments have also contributed to a further spread in the use of granular data since conducting stress testing of that kind requires an accurate grasp of such factors as concentration risk for specific portfolios of individual banks. The BOJ also has accumulated experience in the use of granular data and its efforts in this regard have gradually begun to bear fruit after more than a decade since the outbreak of the GFC. Further progress has been made in financial analysis employing those data during the COVID-19 pandemic.

As an example of such progress, the BOJ has strengthened the collection of granular data on repo and OTC derivative transactions in cooperation with the Financial Services Agency, and the use of these data at the BOJ has provided insights into, for example, network structures of repo and OTC derivative transactions, developments in these transactions by major sector of market participants, and determinants of U.S. dollar funding rates.[17] Moreover, the use of granular data on repo transactions has also revealed that, when financial markets became unstable temporarily in March 2020, the initial stage of the COVID-19 outbreak, the functioning of the Japanese government bond (JGB) repo market declined due to a remarkable decrease in the number of JGB suppliers, mainly in reflection of an increase in demand for collateral. Regarding OTC derivative transactions, it has also been found that there were notable changes in the structure of the Japanese yen interest rate swaps market at the initial stage of the COVID-19 outbreak, as exemplified by a reversal of some market participants' net positions from normal times (in March 2020, foreign financial institutions were net fixed-rate payers whereas Japanese major banks were net fixed-rate receivers).

In addition, the BOJ has proceeded with the use of granular data in its analyses in the *Financial System Report*. For example, for conducting macro stress testing using the Financial Macro-econometric Model (FMM), the BOJ has refined the credit cost model in the FMM using granular data so as to examine risks to financial system stability posed by an increase in loans to middle-risk firms, which have been low-return borrowers for financial institutions.[18] This refinement enabled its macro stress testing using the FMM to take into account the heterogeneity in borrower firms' interest payment capacity. The BOJ has also published the results of (1) an analysis on increased connectedness between Japanese and overseas financial institutions using individual transaction data on the underwriting of syndicated loans and (2) a simulation analysis for the robustness of collateralized loan obligations (CLOs)—securitized products backed by multiple leveraged loans—using data on individual CLOs.[19] Moreover, starting with the October 2020 issue of the *Financial System Report*, the BOJ has been using financial data of around 800 thousand small and medium-sized enterprises (SMEs) as part of its analysis on domestic credit risk; more specifically, these data are used for the simulation analyses of the impacts that the spread of COVID-19 and measures to support corporate financing, such as effectively interest-free loans and cash payments, have on individual SMEs' balance sheets and default rates.

Other Kinds of Alternative Data Analysis

Other than the above, an increasing number of studies have been conducted to analyze data that have not much been the subject of focus in the conventional economic analysis. As an example of such studies, a research paper by BOJ economists investigates the relationship between innovation and productivity or research and development investment, using patent citation information as a proxy variable for the accumulation of innovation.[20] As another example, a research paper has been published on the physical risks posed by climate change; this paper measures the extent of the adverse impact that floods have on corporate profits, as well as the speed at which such impact subsides, by combining municipality-level data on flood damage recorded in *Flood Statistics* (compiled by the Ministry of Land, Infrastructure, Transport and Tourism) with firm-level financial data.[21]

Use of Alternative Data to Enhance the Existing Statistics

The previous sections have primarily focused on the use of alternative data in economic analysis. It should be noted here that some of these data can also be used for enhancing the existing macroeconomic statistics. Regarding the enhancement of the Corporate Goods Price Index (CGPI) compiled and published by the BOJ, there is a study on quality adjustments to price indices (i.e., the process of removing price changes arising from changes in quality of the products) using data on price developments from a price comparison website, employing machine learning.[22] Moreover, as part of the rebasing of the CGPI to 2020, which is scheduled to be effective in 2022, the BOJ plans to adopt web scraping (a technique in which a computer program automatically collects information from webpages) in collecting data employed when making quality adjustments to price indices using the hedonic quality adjustment method. As these examples suggest, alternative data have great potential for providing new insights into the compilation of traditional statistics.

Prospects for the Use of Alternative Data

It is expected that the availability of alternative data will continue increasing in terms of both amount and variety, mainly on the back of further progress in digitalization of the economy, enhancement in firms' information disclosure such as on their efforts to address climate change, and the global trend concerning the strengthening of data collection. While some view that high-frequency data will decrease in value as the economic impact of COVID-19 subsides, it should be noted that these data are also useful for analysis in other times of large short-term fluctuations in the economy; in fact, there have been an increasing number of such cases both at home and abroad arising from, for example, supply shocks for certain goods due to extreme weather events or supply-chain disruptions. Moreover, records of payments made via credit cards or mobile payment services, mentioned earlier in this paper, are highly valuable not only in nowcasting the current state of consumption but also as panel data with high coverage in analyzing consumer behavior by attribute such as age and gender.[23] Regarding text analysis, it is highly likely that progress will continue to be made in the diversification of media sources used for the analysis, as well as in the advancement of new methodologies including those concerning machine learning. It can be said that the use of alternative data in economic and financial analysis has only just started.

Meanwhile, it is important that the ways in which alternative data are released or provided be improved, as many of these data have been provided by private firms. At a forum on big data, co-hosted by the Research and Statistics Department of the BOJ and the Center for Advanced Research in Finance of the University of Tokyo in November 2020, many experts expressed the opinion that private firms' commitment to providing their data in a stable and consistent manner will significantly enhance the user-friendliness of alternative data, thereby leading to the growth of studies and businesses that employ these data.[24] In this regard, for many of the existing public statistics—including those compiled by the BOJ—rules concerning data publication and revision have been developed and made available to the public, and in some respects, these rules may be helpful to alternative data providers. To advance the economic and social use of alternative data to the extent suitable for the digital age, there need to be continued efforts among statistical data users, such as economists and researchers, private providers of alternative data, and public entities, including the government and the BOJ. It is desirable that all these entities proceed with knowledge and information sharing and active discussions toward further improvement in the collection, dissemination, and use of alternative data.

[1] The "Alternative Data Analysis" page is available at https://www.boj.or.jp/en/research/bigdata/index.htm. This page contains the links for all the papers and analyses that are cited in this paper and available on the Bank of Japan's website — for example, papers in the *Bank of Japan Working Paper Series* and the *Bank of Japan Review Series* and boxes in the *Outlook for Economic Activity and Prices*. Please note that some of the cited papers will be released on the website soon.

[2] There is no strict definition as to what kinds of data are classified as alternative data. In accordance with how they are described generally—as presented in this section—alternative data in this paper are broadly defined as those that are not traditional statistical data collected through surveys using questionnaires or the like, which have been frequently used in conventional economic analysis.

[3] High-frequency data are useful not only in the analysis of the real economy, on which this paper primarily focuses, but also in the analysis of financial markets. It should be noted in this regard that, unlike in the field of the real economy, a wide range of daily updated data have been available in the field of financial markets, as exemplified by stock prices and interest rates. For this reason, high-frequency data in the financial field generally refer to, for example, those updated every minute or second and cover individual intraday transactions and offers (tick-by-tick data). As examples of analyses conducted at the Bank of Japan using high-frequency financial data, the following papers employ transaction data in developing indicators of liquidity in the Japanese government bond (JGB) markets: Kurosaki, T., Kumano, Y., Okabe, K., and Nagano, T., "Liquidity in JGB Markets: An Evaluation from Transaction Data," *Bank of Japan Working Paper Series*, no. 15-E-2, May 2015; and Sakiyama, T. and Kobayashi, S., "Liquidity in the JGB Cash Market: An Evaluation from Detailed Transaction Data," *Bank of Japan Research Papers*, March 2018.

[4] In the public sector, the Cabinet Secretariat and the Cabinet Office have been operating the Regional Economy and Society Analyzing System (RESAS) and the V-RESAS, which collect and visualize various big data regarding the regional economy. The Bank of Japan has been making active use of data obtained from these systems in its assessment of economic activity and economic analysis.

[5] In more detail, examples of high-frequency data employed in the Outlook Report are as follows: data reported by Google on mobility trends for places it categorizes as "retail and recreation"; data from TableCheck Inc. on the number of visitors to restaurants that use this firm's reservation and customer management system; data provided by NTT DOCOMO, Inc. on the number of people at airports in Japan; and *JCB Consumption NOW*, a consumption indicator based on credit card spending developed by JCB, Co., Ltd. and Nowcast Inc.

[6] Overseas central banks also have been employing high-frequency data in their analyses. For example, at the Federal Reserve, analyses on the labor market have been made using high-frequency employment statistics provided by a private firm, the findings of which have been mentioned in its reports to the U.S. Congress. In the case of Europe, the Bank of England has presented analyses using data such as on services expenditures and internet searches in its *Monetary Policy Reports*.

[7] For example, Box 1 in the July 2021 Outlook Report employs data on mobility trends in the United States and Europe based on location data as one of the resources for examining the progress on the resumption of economy activity in these economies.

[8] Kobayashi, S., Nakahara, K., Oda, T., and Ueno, Y., "The Impact of COVID-19 on US Consumer Spending: Quantitative Analysis Using High-Frequency State-Level Data," *Bank of Japan Review Series*, no. 2020-E-7, October 2020; and Mori, I., Nakamura, T., and Norimasa, Y., "Gurōbaru ni mita kansenshō no kakei tō no kōdō e no eikyō: Kikai gakushū ni yoru apurōchi," *Bank of Japan Review Series*, no. 2021-J-5, May 2021 (available only in Japanese).

[9] Shiraki, N., Matsumura, K., and Matsumoto, A., "Potential of Search Data in Assessment of Current Economic Conditions," *Bank of Japan Research Papers* (Research and Statistics Department), April 2013.

[10] Oh, Y., Sugo, T., Takahashi, K., and Matsumura, K., "Nowcasting Economic Activity with Mobility Data," *Bank of Japan Working Paper Series*, no. 21-E-2, March 2021.

[11] While this paper explains alternative analyses by broadly categorizing alternative data into high-frequency data, granular data, and textual data, it should be noted that this categorization is only for simplicity. In practice, alternative data include those with several different characteristics, as exemplified by the fact that textual data include those updated with a high frequency.

[12] As another example of textual data analysis conducted at the BOJ, the following paper examines the relationship between the tone of the BOJ's economic assessments and financial market expectations of monetary policy: Kazato, M., Kurosaki, T., and Goshima, K., "Nippon Ginkō ni yoru keiki handan no tōn bunseki," *Institute for Monetary and Economic Studies (IMES) Discussion Paper Series*, no. 2019-J-16, November 2019 (available only in Japanese).

[13] Otaka, K. and Kan, K., "Kikai gakushū ni yoru keiki bunseki -- 'Keiki Wotchā Chōsa' no tekisutomainingu --" *Bank of Japan Working Paper Series*, no. 18-J-8, September 2018 (available only in Japanese); Okazaki, Y. and Tsuruga, T., " Biggudēta o mochiita keizai bukka bunseki ni tsuite -- kenkyū jirei no sābei to 'Keiki Wotchā Chōsa' no tekisuto bunseki no kokoromi --" *Bank of Japan Research Papers* (Research and Statistics Department), June 2015 (available only in Japanese); and Mikami, T., Yamagata, H., and Nakajima, J., "Text Analysis to Gauge the Reasons for Respondents' Assessment in the Economy Watcher Survey," *Bank of Japan Research Laboratory Series*, no. 21-J-2, December 2021.

[14] Nakajima, J., Yamagata, H., Okuda, T., Katsuki, S., and Shinohara, T., "Extracting Firms' Short-Term Inflation Expectations from the Economy Watchers Survey Using Text Analysis," *Bank of Japan Working Paper Series*, no. 21-E-12, October 2021.

[15] In light of the lessons from the GFC, members of and those concerned with the Financial Stability Board were requested to specify information gaps that need to be resolved in financial regulation and supervision and to take steps to collect necessary data at the Group of Twenty (G20) Leaders' Summit and the G20 Finance Ministers and Central Bank Governors Meeting held in 2009.

[16] For the frameworks and history of simultaneous stress testing, see also Financial System and Bank Examination Department of the Bank of Japan and Strategy Development and Management Bureau and Supervision Bureau of the Financial Services Agency, "Supervisory Simultaneous Stress Testing Based on Common Scenarios," *Bank of Japan Review Series*, no. 2020-E-9, December 2020.

[17] Sasamoto, K., Nakamura, A., Fujii, T., Semba, T., Suzuki, K., and Shinozaki, K., "New Initiatives to Improve the Transparency of Securities Financing Markets in Japan: Publication of *Statistics on Securities Financing Transactions in Japan*," *Bank of Japan Review Series*, no. 2020-E-1, February 2020; Horikawa, T., Matsui, Y., and Genma, Y., "A Network Analysis of the JGB Repo Market," *Bank of Japan Working Paper Series*, no. 21-E-14, December 2021; Financial Markets Department of the Bank of Japan and Policy and Markets Bureau of the Financial Services Agency, "Tentō deribatibu torihiki dēta tō no seibi to

katsuyō," *Bank of Japan Review Series*, no. 2021-J-6, June 2021 (available only in Japanese); Inoue, S., Miki, S., and Gemma, Y., "The Japanese Yen Interest Rate Swap Market Observed from OTC Derivative Transaction Data: the Impact of COVID-19," *Bank of Japan Review Series*, no. 2021-E-3, September 2021; Takizuka, Y. and Maruyama, R., "Foreign Exchange Option Market through the Lens of OTC Derivative Transaction Data: Recent Market Developments," *Bank of Japan Review Series*, no. 2021-E-7, December 2021; and Bank of Japan, *Financial System Report*, October 2021.

[18] Bank of Japan, *Financial System Report*, October 2018.

[19] Bank of Japan, *Financial System Report*, October 2019.

[20] Oh, Y. and Takahashi, K, "R&D and Innovation: Evidence from Patent Data," *Bank of Japan Working Paper Series*, no. 20-E-7, November 2020.

[21] Yamamoto, H. and Naka, T., "Quantitative Analysis of the Impact of Floods on Firms' Financial Conditions," *Bank of Japan Working Paper Series*, no. 21-E-10, July 2021.

[22] Abe, N. and Shinozaki, K., "Compilation of Experimental Price Indices Using Big Data and Machine Learning: A Comparative Analysis and Validity Verification of Quality Adjustments," *Bank of Japan Working Paper Series*, no. 18-E-13, August 2018.

[23] Nakajima, J., Takahashi, M., and Yagi, T., "An Assessment of Japan's Online Consumption Trends during the COVID-19 Pandemic," *Bank of Japan Working Paper Series*, forthcoming.

[24] A summary of discussion at the forum is available at https://www.boj.or.jp/research/brp/ron_2020/ron201225a.htm/ (in Japanese).

Bank of Japan Review is published by the Bank of Japan to explain recent economic and financial topics for a wide range of readers. This report, 2022-E-1, is a translation of the original Japanese version, 2021-J-16, published in December 2021. The views expressed in the Review are those of the author and do not necessarily represent those of the Bank of Japan. If you have comments or questions, please contact Research and Statistics Department (E-mail: post.rsd18@boj.or.jp). *Bank of Japan Review* and *Bank of Japan Working Papers* can be obtained on the Bank of Japan's website (https://www.boj.or.jp/en/index.htm).