Chapter 1

A Survey of Flash Data and Nowcasting in Japan

Nariyasu Yamasawa

Atomi University

1. Introduction

This study surveys two statistical approaches to making flash assessments of the real economy. The first approach involves making flash estimations using official statistics. The second approach involves making high-frequency forecasts using various econometric techniques.

A major problem with economic statistics is that the process of releasing them is slow, generally taking at least 1 month to gather, compile, and verify data. Compiling secondary statistics, such as gross domestic product (GDP) statistics, takes yet more time.

Delays in monitoring economic situations can negatively affect policymakers. In 2014, Japan's Prime Minister Shinzo Abe stated that the government would decide whether to raise consumption tax based on the performance of GDP in the July–September quarter. However, GDP data became available only on November 17, 2014, which meant that the government was able to first assess the economic situation 2 months after the relevant period.

Of course, the Statistics Bureau makes efforts to release data promptly, and data today are released faster than in the 1990s. However, there are limits to these flash releases, and to grasp the actual state of the economy sooner, it is necessary to forecast. A variety of research studies have been published about high-frequency forecasts using monthly and daily variables. Researchers have attempted to make monthly GDP forecasts using many different methods, including the current quarter model (CQM), the monthly data model, and the factor model. In addition, inflation forecasts, which are important from the viewpoint of monetary policy, are conducted by many institutions. In this chapter, we will examine these forecast techniques and how they are being used today.

2. Releasing Flash Data

2.1. Delays in obtaining economic data

Weather forecasts and economic forecasts are similar. However, one significant difference between the two is that it takes a long time to obtain economic data, whereas we can obtain weather data immediately. It is possible to gather data on temperature, humidity, and wind speed for each city in real time and use these data to make short-term forecasts. The website of Japan's Meteorological Agency shows radar "nowcasting" for rain, thunder, and tornados. This comprises forecasts of rain, thunder, and tornados from 1 hour before to 1 hour after the present time, in 10-minute intervals.

On the other hand, although market-related data can be obtained almost in real time, it takes longer for industrial production and sales data to be released. Some data are released only quarterly and some detailed data are released only yearly.

Economic policy has to be developed as soon as possible after the economic events. However, there are three kinds of lags in relation to executing economic policy: Recognition, execution, and effect lags. First, a "recognition lag" is a delayed recognition of changes in economic conditions. Second, an "execution lag" is a delayed development of economic policy because of the policymaking procedure. Third, an "effect lag" is a delay in the time taken for economic policy to produce actual outcomes.

The delayed release of economic data is one of the reasons why recognition lag occurs. Policymakers have to use past economic data when they develop economic policy and this is referred to, rather sardonically, as “driving a car by looking in the rearview mirror”.

2.2. Moving towards flash estimates

Recently, there has been a gradual increase in the release of flash estimates. GDP is a useful statistical measure for business cycle analysis; there is a strong need to release this information quickly, and the Statistics Bureau of Japan (part of the Ministry of Internal Affairs and Communications) has made efforts to do so.

Flash estimates are statistical data that are released quickly in substitution of statistics that are more accurate but released later. For example, the Statistics Bureau originally calculated GDP statistics by using data sources like the “Industry Statistics” of the Ministry of Economy, Trade and Industry (METI) and by studying the inflow and outflow of over 2000 commodities. Moreover, the bureau develops GDP statistics using three approaches: the production-, income-, and expenditure-side approaches. In addition, these statistics are designed to obey the three-sided equivalence principle, so the GDP figure that each approach gives must be the same.

Because the statistics are annual, the Cabinet Office compiles a flash quarterly estimate from only the expenditure side. This estimate is created by substituting quarterly statistics for annual statistics, such as the quarterly Household Survey, the Index of Industrial Production, and Financial Statements Statistics.

The Statistics Bureau made revisions in August 2002 that were central to the improvement of Japan’s flash GDP estimates. In that month, the bureau revised the way it made flash estimates by incorporating not only the demand-side but also supply-side statistics into its calculations. In addition, they published these estimates earlier than before. Specifically, before the revision, GDP statistics were released 2 months and 1 week after the relevant period, and after the revision, the first estimates were released 1 month and 2 weeks after, and the second estimates were released 2 months and 10 days after. Although there continues to be a problem in that the revisions that occurs between the first and second

estimates is often large, METI's effort to release the statistics earlier is commendable.

Japan is behind the rest of the world in that it does not yet release flash estimates of the supply side and income-side of its GDP. Most OECD countries make flash estimates of their supply-side and income-side GDPs according to the three-sided equivalence principle (Table 1), the three sides being production, distribution, and expenditure. Japan's Cabinet Office has noted that it is behind in releasing production and distribution estimates and is seeking to improve its estimates by including quarterly estimates of these aspects of the GDP after its next revision of the base year (planned for 2016).

2.3. Data revision problems

Flash estimates of GDP provide useful information about business conditions, but they also have problems. Because flash estimates are not made using the original data that are used to make final estimates, data revisions are unavoidable. Moreover, final estimates are not literally final and can be changed. Statistics change when statistics bureaux revise the base year or the deflator, or when they change from using a fixed price method to a chained price method to calculate the real GDP.

Table 1. Comparison of quarterly flash GDP estimates

	Production		Distribution		Expenditure	
	Real	Nominal	Real	Nominal	Real	Nominal
Japan	—	—	—	—	○	○
United Kingdom	○	Δ	Δ	○	○	○
Germany	Δ	○	—	○	Δ	○
France	○	○	—	○	○	○
United States	○	○	Δ	○	○	○
Canada	○	—	—	○	○	○
Australia	○	—	○	○	○	○

Note: ○ Available, Δ Partly Available, — Not Available

Source: The Cabinet Office (2014), "Research meeting on the next base year revision of National Accounts," 10th meeting, material 2

The data that can be obtained now is called real-time data, and it should be taken into account that policy judgments depend on real-time data. The OECD maintains a public database of real-time data and releases explanations of how it revises those data. For example, the OECD should theoretically only have one real GDP growth rate dataset for Japan in FY 1997, but this dataset was revised several times (Figure 1). The OECD first published Japan's FY1997 GDP growth as -0.8% , but in July 1998, it revised this figure to 0.5% and subsequently to almost 0% in 2014. In 1998, the fact that the GDP growth rate was negative made people believe that business conditions were poor, whereas it turned out that the growth rate was approximately zero. Thus, there is a risk in using real-time data to make judgments.

Researchers have attempted to investigate how data are revised from various viewpoints. A 2007 issue of the "National Account Quarterly" contained a special feature on this so-called "revision study". Kawagoe (2007) studied how Japan's real GDP was revised by dividing the period from the first quarter of 1994 to the third quarter of 2003 into two periods, namely, before and after the fourth quarter of 1998. It was found that the real GDP growth rate in the first period (based on chained 1995 prices) slowed from 1.5% to 0.9% , while the real GDP growth rate in the second period (based on chained 2000 prices) increased from 0.9% to 1.3% .

Figure 1. Change of the growth rate of real GDP in FY1997

Source: OECD Original Release Data and Revisions Database

2.4. Noise or news

The nature of the revision from flash estimates to final estimates has been discussed since the 1980s. The need for flash estimates is strong among market investors, but on the other hand, less credible data cause confusion. This problem is discussed as the problem of “noise or news” (Mankiw and Shapiro, 1986). Mankiw and Shapiro describe revisions that include measurement errors as “revisions based on noise”. They occur when fewer statistics are used to calculate a flash estimate than the amount used to calculate the final estimate, or when industrial classifications within those statistics are rougher than those of the final estimate. On the other hand, they call flash estimates that are designed as good forecasts of the final estimates “revisions based on news”. Further changes to these only occur when news emerges that could not have been known at the time the flash estimate was released. A revision based on news is a flash estimate that is considered to be the best and most desirable forecast of the final estimate possible at the time it was made. A revision based on noise is a flash estimate that is considered to have been produced from unreliable statistics.

Komaki (2005) studied the nature of GDP flash estimates in Japan, and concluded that GDP statistics are revised based on noise. He conducted rational expectation tests of flash estimates, such as bias and efficiency tests, and found that flash estimates of the GDP are different entities to final estimates. Komaki argued that flash estimates are insufficient for indicating the “true GDP” (which Komaki presupposes that final estimates indicate), so people should utilize other statistics, such as stock prices, to find this “true GDP”.

3. High-Frequency Model Forecasts

3.1. Complements of actual data

It takes time to gather and compile data for earlier release, in addition to effort on the part of the Statistics Bureau. In cases in which it is impossible to release data more quickly, researchers attempt to forecast the data.

The Statistics Bureau does, in a sense, forecast some part of the “final” statistics it releases, such as GDP statistics. Making “final flash estimates” is a similar process to refining forecast statistics. For example,

the main difference between the first and second flash GDP estimates is the way that capital investment is calculated. The Statistics Bureau makes the first flash estimate of capital investment based on METI's "Index of Industrial Shipment of Capital Goods", whereas it makes its second estimate based on the Ministry of Finance's "Financial Statements Statistics of Corporations". In a sense, the Statistics Bureau forecasting capital investment using the Index of Industrial Shipment of Capital Goods before the Financial Statements Statistics of Corporations are published is akin to using a proxy variable to forecast it.

For the first flash estimate, only 2 months of balance of payments data of the relevant quarter have been published. The data of the third month are forecast. For example, for the July–September estimate of GDP, the balance of payments data are available only for July and August, and the Statistics Bureau must forecast figures for September.

3.2. Quick estimation forecast

Quick estimation (QE) is a quarterly flash estimate of real GDP. QE is a very popular forecast and many organizations often make it public. The Japan Center for Economic Research gathers QE forecasts every month and publishes them as "Economy, Society, and Policy (ESP) forecasts".

Takao Komine of Hosei University, formerly chief economist of the Japan Center for Economic Research, coined the term "super short-term forecasts" for QE forecasts. Komine developed this term in the late 1980s in his book, "Introduction to New Business Cycle Observation Methods" (*Nihon Hyoronsha*), in which he classified economic forecasts into short-term, medium-term, and long-term forecasts. He defined short-term forecasts as forecasts of the economy 2 or 3 years ahead, medium-term forecasts as forecasts of the economy about 5 years ahead, and long-term forecasts as forecasts of the economy several decades ahead. His medium-term forecasts focus mainly on industrial structure forecasts and long-term forecasts focus mainly on population forecasts. Furthermore, he defined super short-term forecasts as those for the quarter ahead.

This classification of economic forecasts has been followed by other research institutes. They also consider medium-term forecast as those that focus mainly on industrial structure forecasts, long-term forecasts as those

that focus mainly on population forecasts, and short-term forecasts as those that focus on the turning points of business cycles. The only difference in terminology is that most institutions call forecasts of the upcoming quarter as “QE forecasts” instead of “super short-term forecasts”.

3.3. Bridge equation: Current quarter model

A popular method of making QE forecasts involves using bridge equations. GDP is quarterly data, but monthly data can be used when compiling some of the statistics needed to forecast GDP. If we change quarterly and monthly data into percentages of annual data, and make that quarterly data the dependent variable Q_t and monthly data the independent variable M_t , we can use the following bridge equation:

$$Q_t = \alpha + \beta \times M_t + e_t, \quad (1)$$

where α and β are coefficients, e_t is the residual, and M_t is an average of 3 months of data.

The CQM is a forecast method that involves using bridge equations. The method was developed by Lawrence Klein of University of Pennsylvania. If we have M_t for the January–March quarter, then we can estimate Q_t for that quarter. If we have data only for January, Klein suggests forecasting M_t for February and March using time-series analysis. After the release of February data, Klein recommends revising the dataset, forecasting M_t for March, and estimating the first-quarter GDP.

In practice, Klein’s forecast method takes into consideration not only the demand side but also the supply side and the income side. In addition, Klein uses many monthly datasets to estimate quarterly GDP data. Regarding the applications of the CQM, Yuzo Kumasaka of IT Economy Advisers LLC has released forecasts of the GDP growth rate of the United States using the CQM, while Yoshihisa Inada of Konan University has made weekly CQM forecasts of Japan’s GDP for 20 years (Figure 2).

Inada (2013) argues that CQM forecasts are quicker and more accurate than the market consensus.

Figure 2. CQM forecasts of United States real economic growth rates in October–December 2008 (seasonally adjusted quarterly growth of per annum rates)

3.4. Accuracy of super short-term forecasts

In this section, we will investigate the accuracy of super short-term, or QE, forecasts. Private forecasting companies have almost the same information, and each organization is aware of how to make super short-term forecasts of GDP to some degree. However, there is disparity between each company's forecast results because these companies do not have perfect information about how to make QE forecasts and they still rely heavily on guesswork. The super short-term forecasts these companies make are seldom accurate, and sometimes, their poor forecasts confuse the financial market.

According to Saito (2012), the margin of error of independent forecasts is very large. Saito found that the absolute average error of the real GDP growth rate was 1.08% from 2000 to 2012. Furthermore, Saito found that the probability of forecasts being within a 0.5% range of the actual GDP was only about 20%.

The announcement of the real GDP growth rate for the third quarter of 2014 was watched with keen interest because Prime Minister Abe indicated he would decide whether the government should raise the

consumption tax rate based on the GDP outcome for this period. A November 2014 survey of the Japan Center for Economic Research estimated that the government's first preliminary GDP growth estimate for the third quarter of 2014 would be 2.47%. However, the actual estimate was far worse than economists expected, pegging third quarter GDP growth at negative 1.6%. Independent forecast errors rose to about 4 percentage points in this period. Moreover, for the government's second flash GDP estimate, most economists expected an upward revision, but GDP growth was in fact revised downward. In April 2014, the consumption tax rate increased from 5 to 8%, and economists have become worried about business conditions after the tax hike.

The media have criticized the poor performance of economic forecasts. In the behavioral economics literature, some researchers argue that economic forecasting is not rational. Yamasawa (2011) summarized the various kinds of economic forecaster bias, such as herding and wishful expectations.

On the other hand, Iizuka (2014) argues that poor forecast performance is to at least some extent caused by the Statistics Bureau. According to Iizuka, at the time of the first preliminary estimate, forecasters do not know the material inventory and supply chain inventory statistics for that quarter. When it comes to the second preliminary estimate, the breakdown of estimate results of the demand side and the supply side are not open to the public. Thus, one cause of the forecast error is insufficient public information.

3.5. Monthly GDP

To know what Statistics Bureau's QE of the GDP will be ahead of time, several institutes calculate the monthly GDP before the QE for that quarter is released. The "monthly GDP" is quarterly GDP data that has been converted to monthly figures. Most basic GDP statistics, such as consumption and capital investment statistics, are calculated based on monthly data from sources like the Index of Industrial Production and trade statistics. Monthly GDP data provide information on business cycles. It is important that monthly GDP data is available to academic researchers, as researchers use this data to perform a variety of empirical studies.

One method to obtain monthly GDP is the bridge equation mentioned in Section 2.3. Equation (1) is estimated quarterly. If this relationship holds on a monthly basis, we can estimate monthly GDP series. We can obtain the monthly Q_t , by substituting monthly M_t into the equation rather than quarterly M_t .

The Japan Center for Economic Research has released monthly GDP data since January 1999. In addition, the NLI Research Institute began to release monthly GDP data from almost the same period. For more information about the methodology that the Japan Center for Economic Research uses to calculate monthly GDP, see Yamasawa (2003).

The Cabinet Office does not release monthly GDP data, but it does release integrated Private Consumption estimates on a monthly basis. These estimates are the monthly index of consumption, which is the largest component of GDP.

Hara and Yamane (2013) introduced another model for estimating monthly GDP using the bridge equation. They set quarterly GDP as a dependent variable and set the levels of the Index of Industrial Production and Index of Tertiary Industries Activity as the explanatory variables in their bridge equation. They also added data reduced to principal components to their explanatory variables, and attempted to fit various time series into their equation to create a new method of estimating monthly GDP.

It is appropriate that monthly GDP is not considered to be a forecast of GDP but rather is considered to be new real GDP data. Monthly GDP enable us to analyze the economic outlook in detail. For academic use, it can be a proxy for the overall economy. Monthly GDP data is desirable in empirical econometric studies because it gives analysts a larger sample size to work with, which improves the accuracy of estimations. In the past, analysts had to analyze monthly data from sources such as the Index of Industrial Production when analyzing production output, but if monthly GDP data becomes readily available, analysts can use it instead to make more detailed analyses.

However, monthly GDP fluctuates too much for analysis by growth rate. For this, it is better to use the real GDP gap, which is the difference between actual GDP and potential GDP (Yamasawa, 2003).

4. Nowcasting and the Dynamic Factor Model

4.1. Nowcasting

Recently, the concept of “nowcasting” — a portmanteau combination of “now” and “forecasting” — is spreading. According to Banbura *et al.* (2013), nowcasting means forecasting for the very near future or the very recent past. The word originates in meteorology. On the website of the Japanese Meteorological Agency, for example, a nowcasting page enables viewers to access approximately 1-hour forecasts for rain, tornados, and thunder.

“Nowcasting” is easy to understand if it is compared to earthquake forecasts. Ordinary economic forecasts are forecasts of “earthquakes that have not happened yet”, but nowcasting is a method of capturing “earthquakes that have happened already” as quickly as possible.

An example of nowcasting is given on the “nowcasting” website of the Federal Reserve Bank of Cleveland (Figure 3). This site displays the inflation rate “nowcast” of the current month, and updates this data daily. For example, it displayed the December 2014 inflation rate of the consumer price index in that same month, and provided daily updates.

According to the *International Journal of Forecasting*, which focused on nowcasting in 2013 (volume 29), many central banks around the world

Figure 3. Inflation nowcasting

Source: Federal Reserve Bank of Cleveland (<https://www.clevelandfed.org/>)

attempt to nowcast. Nowcasting has the same characteristics as QE forecasts in that both methods involve forecasting quarterly GDP. However, unlike QE forecasting, nowcasting involves gathering a large amount of data and updating forecasts when actual data are updated. Another point of difference is that nowcasting is mainly carried out using state space and factor models.

Nowcasting has become popular in the business world. “Nowcast.com” was established in 2011 in London to provide GDP forecasts of many countries, founded by Domenico Giannone and Lucrezia Reichlin, two pioneer researchers in the field.

The SWIFT Index is an index of GDP nowcasting of OECD countries. Named after the Society for Worldwide Interbank Financial Telecommunication (SWIFT), an organization formed in 1973 in Brussels and supported by financial institutions, the index provides nowcasting of the GDPs of the United States, the United Kingdom, Germany, and other OECD countries. It uses financial trade volumes to estimate GDP.

4.2. Dynamic factor model

In the methodology of nowcasting, the dynamic factor model is the most common method used to perform nowcasting. This section explains this model.

The dynamic factor model is expressed in the form of the state space model.

The state space model includes state variables that cannot be observed. Using this method, we incorporate quarterly GDP data from the Index of Industrial Production as the observable variables, and “business cycle” data series, for example, as unobservable variables. This is the dynamic factor.

In this method, the vector X_t is made up of n observable items, and the vector F_t is made up of q state variables. If we assume that F_t is decided by F_{t-1} , the state space model can be arranged as follows:

$$X_t = \mu + \Lambda F_t + e_t \text{ (observation equation)}$$

$$F_t = \mu + \Lambda F_{t-1} + B u_t \text{ (state equation),}$$

where e_t and u_t are error terms.

The merit of this model is that the number of time-series data is not constrained. We can integrate many series into one dynamic factor, and this dynamic factor gives an insight into business conditions.

A typical example of a dynamic factor model that analyses the business cycle is the Stock–Watson model, which assumes an unobservable variable that indicates “business conditions”. The movement of this unobservable variable is restricted by the state equation. In addition, the variables that express the business cycle in the Stock–Watson model are reflected in observable series, such as the Index of Industrial Production and the unemployment rate. The model develops one business condition index by integrating many monthly time series.

If we adopt the state equation, the business condition is expressed by F_t , although in the Stock–Watson model it is not a vector but rather a variable. In the observation equation X_t , a number of regression variables, including the Index of Industrial Production, the effective job openings-to-applications ratio, and large electric companies’ power consumption statistics, are considered to be a single vector.

The Nikkei BI is a Stock–Watson-type index (Figure 4). It has been released since June 2000 on a monthly basis by the Nikkei Inc. (literally the “Japan Economic Newspaper”). The BI estimates business

Figure 4. Nikkei business index

Source: Nihon Keizai Shimbun website (<http://www.nikkei.com/>)

conditions based on the Index of Industrial Production, retail and wholesale sales, extra labor hours, and the effective job openings-to-applications ratio.

4.3. State space model of monthly GDP

Mariano and Murasawa (2010) extended the dynamic factor model and developed monthly GDP forecasts using the state space model.

The following is the state space model again:

$$X_t = \mu + \Lambda F_t + e_t \text{ (observation equation)}$$

$$F_t = \mu + \Lambda F_{t-1} + B u_t \text{ (state equation),}$$

where e_t and u_t are error terms.

X_t is a vector of n observable variables. Most of these are monthly datasets, such as the Index of Industrial Production, but quarterly GDP is also included as a variable. This variable appears once a quarter but is missing for the rest of the period.

The relationship of quarterly GDP (seasonally adjusted) with monthly GDP (seasonally adjusted) is as follows:

$$\text{GDP in April–June quarter} = \text{average of monthly GDP in April, May, and June.}$$

In the Stock–Watson model, F_t is a singular variable that indexes business conditions, but here, F_t is monthly GDP. The basic consideration is that if the “business conditions” figure represents the movement of the economy, then it is identical to real monthly GDP. Therefore, the state equation should be constructed by setting one state variable that expresses “business conditions” as the monthly GDP.

The state space model can be solved by optimizing the likelihood function, but this includes-missing data as quarterly GDP data is only observable once every 3 months. However, the coefficient parameters can be estimated by adopting an Expectation-Maximization (EM) algorithm, which makes the missing data zero, and then inserts random data with a

mean of zero and a distribution of 1 into the equation. This method has been taken on by recent nowcasting analyses.

4.4. Applications of the dynamic factor model

In recent nowcasting analyses, applying an extended version of the dynamic factor model has become popular. In the Stock–Watson model and the state space expression of monthly GDP, only one dynamic factor can be used. However, this is not the case in all equations. The observation equation in the state space model indicates that n observable variables are explained by q dynamic factors. If $q = 1$, this is a model of one dynamic factor, and if q is greater than one, it is a model of multiple dynamic factors.

The dynamic factor model allows for a variety of time series to be used as the observable variable X_t . Some researchers have attempted to use financial indicators, such as stock prices and interest rates as their observable variable, while others have attempted to use survey data about business conditions.

Furthermore, real-time updates have become popular in the nowcasting field recently. These update the forecasts immediately by incorporating new data into the equation as it is released. For example, after the release of the Index of Industrial Production, such updates estimate GDP immediately.

When real-time updates are included in the dynamic factor model, some actual data goes missing as new data has a different time frame to old data. A commonly-used solution to this problem is to fill the missing data using the Kalman filter.

5. Expansion of the Forecast Field

5.1. From monthly to daily data

One possible future research trend could be the use of higher frequency data. Modern forecasters are shifting their focus to forecasting monthly data rather than quarterly data, and daily data rather than monthly data. Financial data, such as interest rates and stock market data, are easily

obtained daily data, and several researchers use this daily financial data to measure business conditions.

Some researchers are also utilizing sales data. The Nikkei Inc. has stored point of sales (POS) data and analyzed these data for 20 years.

The University of Tokyo (TU) Daily Price Index uses POS data as a price index. The TU daily price Index project is led by Tsutomu Watanabe of Tokyo University. TU's index calculates the daily price index using supermarket price and sales data stored on the POS system. The index covers 300 stores and more than 200,000 items. Daily series are released 3 days after the relevant period. There are a number of ways of integrating large volumes of price data, but TU's system adapts the Törnqvist index, which is a kind of weighted average method that considers both price and sales (Watanabe and Watanabe, 2014).

The TU daily price index gained the public's attention when it captured price movements better than the official monthly consumer price index (CPI) after the Great East Japan Earthquake. In addition, it allowed researchers to obtain price information faster than previously when the consumption tax rate was raised from 5–8% in April 2014 (Figure 5).

Figure 5. University of Tokyo daily price index

Note: Bottom line is the TU Daily Price Index, top line is the official Consumer Price Index

Source: TU Daily Price Project (http://www.cmdlab.co.jp/price_u-tokyo/media_e)

The weak point of this index is that it cannot cover all the items of the CPI. The official CPI is calculated using the prices of all the items that consumers purchase, so the weight of goods and services is each about 50%. However, the TU daily price index mainly covers daily goods sold by convenience stores and supermarkets, meaning that it doesn't properly represent consumption trends as a whole.

A common weakness of high-frequency data is that they tend to fluctuate substantially, which makes it difficult for analysts to establish basic trends using this data. Thus, the development of a statistical method to grasp these basic economic trends is necessary.

The TU project team's experimental development of a daily sales data index proves that forecasting daily movements of this data is possible. Converting annual GDP data into daily data is difficult at present, but if consumption data is recorded in real time in the future, forecasting the GDP daily will become a real possibility.

5.2. Using search engines

One form of high-frequency data that analysts can obtain is "big data". Choi and Varian (2012) explain how to use Google Trends and apply it to forecasting. Google Trends has provided daily and weekly reports on volumes of queries since 2004. It allows users to view graphs displaying the volume of any search term and download this data as a comma-separated value (CSV) file for free.

Choi and Varian (2012) introduced several studies that attempted to use the volumes of search engine queries related to certain GDP statistics to ascertain the movement of those statistics. These studies looked into, for example, auto sales, unemployment rates, tourism data, and consumer confidence index trends as indicated by Google search volumes.

Another example of business cycle analysis using search engines is a project by Yahoo! Japan, called "Is it possible to capture the 'present' business condition? (<http://event.yahoo.co.jp/bigdata/keiki/>)", first announced on the Yahoo! Japan homepage on April 13, 2013. This project investigated the correlation between actual business conditions and volumes of queries related to those business conditions. It also attempted to forecast a diffusion index of economic indicators based on related search

volumes. The study showed that the population of search terms input into Yahoo! Japan in 2012 was 7.5 billion, with about 600,000 of these terms used regularly, and 206 of the terms related to business conditions. Yahoo! Japan grouped terms with similar meanings and integrated them into one series using principal component analysis. It then regressed this principal component into the diffusion index to estimate GDP data. While the firm has not provided a recent update of these forecasts, the method represents one way of forecasting business conditions using search term data.

Shiroki *et al.* (2013), researchers from the Bank of Japan, tried to apply Google Trends to nowcasting by investigating the correlation between tourism sales before and after the Great East Japan Earthquake and trends in tourism-related search terms. They concluded that volumes of queries contained useful information for tourism sales.

5.3. Application to regional statistics

Several researchers have attempted to use nowcasting to estimate Gross Regional Product (GRP), to give the production output of each prefecture in Japan. The economic activity of prefectures can be ascertained from the Cabinet Office's prefectural GRP statistics, but the government takes a long time to release these data. In fact, the government releases GRP data for Japan's 47 prefectures 2 years and 3 months after the relevant period. As of March 2015, we have only annual GRP data for 2011 for the 47 prefectures.

Recently, examining business conditions by region has become very important because different regions have divergent business conditions. Many prefectures compile their own diffusion indexes of local business conditions, integrating various economic series. However, few prefectures have attempted to make flash estimates of their GRP, which is a better indicator than a diffusion index in terms of comprehensiveness.

Hyogo prefecture releases QE of its GRP (Ashiya, 2009). QE are flash estimates using a method that is also applied to national statistics. Using the least squares method, Ashiya (2009) estimated quarterly series by fewer indicators than those used for the final estimates. For example, Ashiya estimated final private consumption using the household survey, retail and wholesale sales, and department store sales data.

Gunma Prefectural Office (2006) estimates the GRP of Gunma prefecture. A unique aspect of their estimate is that it is a supply-side estimation, while flash GDP estimates is usually made using the demand side.

Yoshihisa Inada developed a method to estimate the GRP of Osaka prefecture. This estimate uses principal component analysis, the forecast series can be released faster than government data, and it has demonstrated accuracy (Ogawa and Inada, 2013).

The Cabinet Office also publishes monthly expenditure statistics for each prefecture called the Regional Domestic Expenditure Index (RDEI). These statistics give several components of GRP, such as private final consumption, private residential investment, private equipment investment, and public investment. We can estimate monthly GRP by estimating the rest of the GDP expenditure components. Yamasawa (2014) estimated final public consumption and net exports for 47 prefectures using panel data estimation. This monthly real GRP for 47 prefectures can be used to analyze business conditions (Figure 6).

Figure 6. Real monthly GRP (accumulation of 47 prefectures)

Source: Yamasawa (2014)

Real monthly prefectural GRP can be forecast at about the same rate as the quarterly GDP is forecast. If researchers had access to this monthly prefectural data, they would be able to use it in a variety of econometric studies. For this reason, it is important that a more comprehensive and timely method of forecasting the GRP for each of Japan's prefectures be developed.

References

- Ashiya, Tunenori (2009) "Present Conditions and Problems of Prefectural Accounting," *Statistics*, 96, pp. 54–71 (in Japanese).
- Banbura, Marta, Giannone, Domenico, Modugno, Michele and Reichlin, Lucrezia (2013) "Nowcasting and the Real-time Data Flow," ECB Working Paper, No. 1564.
- Choi, Hyunyoung and Varian, Hal (2012) "Predicting the Present with Google Trends," *The Economic Record*, 88, pp. 2–9.
- Gunma Prefectural Office (2006) "Practices and Problems of New Quarterly Preliminary Estimate of Prefectural Economic Account — Supply Side Approach," National Account Quarterly No. 132, Economic and Social Research Institute, Cabinet Office (in Japanese).
- Hara, Naoko and Yamane, Shotaro (2013) "New Monthly Estimation Approach for Nowcasting GDP Growth: The Case of Japan," Bank of Japan Working Paper Series No. 13-E-14.
- Iizuka, Nobuo (2014) "Making Hasty Information Disclosure of GDP Preliminary Estimates," Keizai Kyositsu, *Nihon Keizai Shimbun*, Morning Edition, December 26 (in Japanese).
- Inada, Yoshihisa (2013) "Super Short-Term Forecasting for Judging Business Conditions," Keizai Kyousitsu, Nikkei Inc., Morning Edition, September 20 (in Japanese).
- Kawagoe, Masaaki (2007) "A Consideration of Revision of the Growth Rate — Analysis of Real-Time Data of Real GDP," National Account Quarterly No. 134, Economic and Social Research Institute, Cabinet Office (in Japanese).
- Komaki, Yasuyuki (2005) "Measurement Errors or Forecast Errors: An Analysis of GDP Revision," *Kokumin Keizai Zasshi*, 191(1), pp. 1–15, Kobe University (in Japanese).
- Komine, Takao (2005) "Introduction to the New Business Cycle Observation," *Nihon Hyoron Sha* (in Japanese).

- Mankiw, N. Gregory and Shapiro, Matthew D. (1986) "News or Noise: An Analysis of GNP Revisions." *Survey of Current Business*, pp. 20–25.
- Mariano, S. Roberto and Murasawa, Yasutomo (2010) "A Coincident Index, Common Factors, and Monthly Real GDP," *Oxford Bulletin of Economics and Statistics*, 72(1), pp. 27–46.
- Ogawa, Ryo and Inada, Yoshihisa (2013) "The Development of Preliminary Estimates of Prefectural GDP that are Consistent, Quick, and Accurate," APIR Discussion Paper Series No. 33 2013/4 (in Japanese).
- Saito, Taro (2012) "An Inspection of the Accuracy of Forecasts of Preliminary GDP Estimates," Weekly Economist Letter, NLI Research Institute, May 11, (in Japanese).
- Shiroki, Noriyuki, Matsumura, Kohei, and Matsumoto, Azusa (2013) "An Availability of Search Term Data to Judge Business Conditions," BOJ Reports & Research Papers, January (in Japanese).
- Tsutomu, Watanabe and Kota, Watanabe (2014) "Estimating Daily Inflation Using Scanner Data: A Progress Report," CARF Working Paper Series, CARF-F-342, February.
- Yamasawa, Nariyasu (2003) "Monthly GDP as a Business Condition Indicator," in Asako, Kazumi, and Fukuda, Shinichi (eds.) "*Business Cycles and Business Forecasts*," Tokyo University Press, pp. 201–231 (in Japanese).
- Yamasawa, Nariyasu (2011) "New Essays on Economic Forecasts," *Nihon Hyoron Shya* (in Japanese).
- Yamasawa, Nariyasu (2014) "Measuring the Indirect Losses from Natural Disasters: The Case of the Great East Japan Earthquake," *61st Annual North American Meetings of the Regional Science Association International*, Washington.