Kanagawa University Economic Society
Discussion Paper No. 2017-2

Forecast Disagreement and Business Cycles

- Evidence from 12 years of ESP Forecast Surveys in Japan -

Nobuo Iizuka

September 27, 2017

Forecast disagreement and business cycles*

- Evidence from 12 years of ESP Forecast Surveys in Japan -

Nobuo Iizuka†

September 2017

Abstract

We examine the predictive power of forecast disagreement—the cross-sectional dispersion of individual forecasts—for key economic indicators. Using the ESP forecast survey, the first monthly forecast survey in Japan, we generate monthly disagreement and forecasting error series and test Granger causality between two series. Overall, disagreement does not have predictive power. Disagreement in the GDP growth rate does not Granger-cause forecasting error, though forecasting error Granger-causes GDP growth rate. Disagreement in the CPI inflation rate with a longer horizon Granger-causes forecasting error, which might be affected by the introduction of the quantitative qualitative monetary easing policy by the Bank of Japan. However, we need an additional forecast survey of CPI inflation to obtain clear evidence.

Key Words: Forecast disagreement; Granger causality; CPI inflation.

JEL Classification: C13; C32; E32; E58.

1 Introduction

Disagreement among forecasters, which is usually measured by the dispersion of point forecasts of a panel of respondents, is of considerable interest. Prior studies intensively analyze the drivers of disagreement. Mankiw *et al.*(2003) explain the movement of disagreement through a sticky-information model. Patton and Timmermann (2010) use a simple reduced-form state-space model to explain the cross-sectional dispersion of U.S. GDP growth and inflation forecasts. Dovern *et al.*(2012) and Siklos (2013) show the relationship between central bank transparency and disagreement.

*This research is supported by JSPS KAKENHI Grant Number 26380332.

†Faculty of Economics, Kanagawa University, 3-27-1 Rokkakubashi, Kanagawa-ku, YOKOHAMA, 221-8686, JAPAN, E-mail: nobuo-iizuka-0915@kanagawa-u.ac.jp

On the other hand, a recent study on disagreement stresses its forecast power. Legerstee and Franses (2015) examine the predictive power of disagreement and show that in some cases, using measures of disagreement in Markov regime-switching models yields a higher forecast accuracy. Atalla *et al.*(2016) examine quarterly oil price forecasts, showing show that average forecast error is positively correlated with disagreement and demonstrating the relationship between disagreement and oil price volatility. Our study is in line with these recent studies. We examine the forecast power of disagreement by testing whether disagreement Granger-causes forecast error using the Granger causality test.

As Atalla *et al.*(2016) state, we can show disagreement equal to the variance of forecast error. However, it is possible that the variance of the forecast error rises without an increase in disagreement, especially in a recessionary phase, when the actual differs substantially from the forecast. Even if the disagreement is small, the forecast error is large. Simultaneously, an increase in the forecast error possibly leads to the disagreement.

In this study, we construct monthly cross-sectional disagreement between forecasters and forecast error using the ESP Forecast (ESPF) survey in Japan along the lines of Dovern *et al.*(2012), who approximate fixed-horizon, 1 year ahead forecasts as a weighted average of the current and next calendar year [1]. The ESPF survey began in 2004 and has only 13 years of data, so we do not have a large enough sample for the same examination as Dovern *et al.*(2012). However, the ESPF survey provides not only annual but also quarterly forecasts, so we can generate monthly disagreement and forecast error with several forecast horizons. We understand that Granger causality is not necessarily true causality; however, we believe we can show the relationship between disagreement and forecast error that prior studies do not.

The remainder of the paper is organized into four sections. Section 2 explains the ESPF survey data and generating series, disagreement, and forecasting error. Section 3 examines the development of disagreement and forecasting error series. Section 4 shows the results of a Granger causality test. Section 5 provides the interpretation and conclusion.

2 Data

2.1 Data set

We use the ESPF survey data set compiled by the Japan Center for Economic Research (JCER) [2]. The ESPF survey is the first monthly survey of

[1] 'ESP' is the acronym of a public relations magazine of the Cabinet Office, 'Economy, Society, and Policy', and does not stand for extrasensory perception.(Komine *et al.*, 2009)

[2] The Association for Economic Planning conducts the ESPF from April 2004 and was taken over by the JCER from April 2012

macroeconomic forecasts conducted by professional forecasters in Japan. We can also obtain monthly forecast for Japan through Consensus Economics, the world's leading international economic survey organization, which is older than the ESPF. However, in Japan, the ESPF panel has twice the number of participants as the Consensus Economics panel does.

Each month since April 2004, the ESPF survey polls professional forecasters from private economic institutions in Japan for their predictions about the main macroeconomic indicators. Neither the BOJ nor the Japanese government participates in the survey. The ESPF requests respondents to provide annual and quarterly forecasts at the beginning of each month. Approximately 40 participants respond every month.

Our sample ranges from April 2004 to March 2017 and consists of 156 monthly observations. We focus on quarterly forecasts for three macroeconomic indicators: real GDP growth (quarter to quarter change, annualized), consumer price inflation (all items less fresh food, year-on-year change), and the unemployment rate. The ESPF survey included these three indicators since its launch. In our sample, we have 58 respondents, 25 of whom participated in 90 percent of the survey.

2.2 Measuring disagreement series

In the ESPF survey, we have h quarters ahead forecasts by forecaster i, and within each quarter, the survey is conducted monthly. Let a forecast target quarter q be dated by its last month; we denote the v-month ahead forecast surveyed at t, $y_{i,q(=t+v)|t}$, as follows:

- $v = 3(h+1)$ for the forecast if t is in the first month of a quarter;
- $v = 3(h+1)-1$ for the forecast if t is in the second month of a quarter; and
- $v = 3(h+1)-2$ for the forecast if t is in the third month of a quarter.

Table 1 provides an example of a forecast target and survey date. In this example, the forecast target is the fourth quarter of 2005 ($q = 200512$).

[Insert Table 1 about here]

As Siklos (2013) states, there is no universally agreed upon measure of forecast disagreement. However, most prior studies use the squared deviations among individual forecasts as a measure of forecast disagreement. In this study, we use the root mean squared deviations among individual forecasts as a measure of forecast disagreement because we use the Root Mean Square Error (RMSE) as a measure of forecast error.

$$D_{y,q|t} = \sqrt{\frac{1}{N}\Sigma(y_{i,q|t} - \overline{y}_{q|t})^2},$$

where $\overline{y}_{q|t}$ is the cross-sectional mean of the v-month ahead forecast for quarter q surveyed at time t. In this study, we use from 0-quarter-ahead (nowcasting) to 3-quarter ahead forecasts. Therefore, we can construct four monthly disagreement series as follows:

- Nowcast: $D_{y0,t} = D_{y,q|t}$ if $1 \leq v \leq 3$;
- one-quarter ahead forecast: $D_{y1,t} = D_{y,q|t}$ if $4 \leq v \leq 6$;
- two-quarter ahead forecast: $D_{y2,t} = D_{y,q|t}$ if $7 \leq v \leq 9$; and
- three-quarter ahead forecast: $D_{y3,t} = D_{y,q|t}$ if $10 \leq v \leq 12$.

Each series consists of 156 observations from April 2004 to March 2017.

2.3 Measuring forecasting error and other series

Following prior studies, we use the cross-sectional RMSE as a measure of forecast error.

$$RMSE_{y,q|t} = \sqrt{\frac{1}{N}\Sigma(A_q - y_{i,q|t})^2},$$

where A_q is real time actual data for quarter q. Similar to the disagreement series, we can construct four monthly forecasting error series as follows:

- Nowcast: $RMSE_{y0,t} = RMSE_{y,q|t}$ if $1 \leq v \leq 3$;
- one-quarter ahead forecast: $RMSE_{y1,t} = RMSE_{y,q|t}$ if $4 \leq v \leq 6$;
- two-quarter ahead forecast: $RMSE_{y2,t} = RMSE_{y,q|t}$ if $7 \leq v \leq 9$; and
- three-quarter ahead forecast: $RMSE_{y3,t} = RMSE_{y,q|t}$ if $10 \leq v \leq 12$.

In our sample, $RMSE_{y0,t}$ and $RMSE_{y1,t}$ consists of 156 observations from April 2004 to March 2017. $RMSE_{y2,t}$ consists of 153 observations from April 2004 to December 2016. $RMSE_{y3,t}$ consists of 150 observations from April 2004 to September 2016.

In our Granger causality test, we also include uncertainty since prior studies show a relationship between uncertainty and disagreement. Lahiri and Sheng (2010) establish a relationship connecting forecast uncertainty to disagreement, theoretically and empirically. Since the seminal work of Zarnowitz, and Lambros (1987), prior studies construct uncertainty measures using density forecasts or volatility in an actual series. Although the ESPF surveys density forecasts, it does not have a density forecasts database. Instead, we use the Volatility Index Japan (VXJ) as a proxy for uncertainty in the economy. The VXJ is a benchmark of future volatility in the Japanese stock market developed by the Center for the Study of Finance and Insurance, Osaka University. VXJ is daily data; therefore, we use VXJ at the beginning of each month.

2.4 Descriptive statistics

Before we analyze the Granger causality between disagreement and RMSE, we check the descriptive statistics. Figures 1 to 3 compare the nowcast and actual variables. Although the nowcast series is monthly, actual is quarterly. Thus, we place same actual data each month within same quarter. The two dotted lines are the mean forecast plus disagreement and mean forecast minus disagreement. The range of the two dotted lines show two instances of disagreement. The shaded areas denote recessions as identified by the Committee for Business Cycle Indicators in Japan. Table 2 summarizes the key statistics for disagreement, forecasting error and actual series. Three findings appear from these statistics. First, forecasts are more stable than the actual series. Second, forecasts are optimistic in a recession. Third, disagreement and RMSE become larger as the horizon become longer. These findings are in line with those of Dovern *et al.*(2012).

[Insert Table 2 about here]

However, for GDP growth, disagreement and RMSE develop differently from other two series. This is partly because the fluctuation of the actual series is very large. In our sample, from $2Q$ of 2004 to $2Q$ of 2017, the standard deviation of the GDP growth rate is 4.0 points, and the mean growth rate is 0.9 percent. In the same sample, the values are 2.0 points and 3.2 percent for the US, respectively.

Even when we nowcast, it is hard to forecast Japanese GDP growth. As we show in Figure 1 to 3, the actual series lie between the two dotted lines only in 44 months (26 percent) within the 156-month sample; the values for CPI inflation over 96 months is 61.5 percent and for the unemployment rate for 81 months is 51.9 percent. There is some uncertainty in estimating GDP growth: the average disagreements and RMSEs of the GDP growth rate are about the same with different horizons.

[Insert Figures 1 to 3 about here]

Therefore, we also use disagreement and RMSE of the average GDP growth rate, the two-quarter average for the one-horizon ahead forecast, three-quarter average for the two-horizon ahead forecast, and four-quarter average for the three-horizon ahead forecast.

3 Disagreement and forecast error series development

3.1 Disagreement and business cycles

Figures 4 to 6 compare the disagreement between the nowcast and the three-quarter ahead forecast. For GDP growth, we add the disagreement of the

four-quarter average growth. Each disagreement series moves together. For CPI inflation and the unemployment rate, the disagreement in the three-quarter ahead forecast is larger than those of the nowcast.

[Insert Figures 4 to 6 about here]

Prior studies show that disagreement becomes larger in a recessionary phase. In our sample, there were two recessionary phases. The first is from March 2008 to March 2009, and disagreement became larger in the latter half of this recession, during which the world was in the midst of financial crises. The second is from April 2012 to November 2012. However, disagreement stayed at the same level.

These findings show that disagreement for GDP growth is difficult to predict during a recession, and we need some shock to become larger.

3.2 Disagreement and other shocks

The Great East Japan Earthquake occurred on March 11th, 2011, and it falls within our sample. Disagreement of forecasts surveyed in April 2011, immediately after this earthquake spiked. Disagreement lasted longer, until November 2011. Economists needed some time to recognize the magnitude of the effect of the earthquake on the Japanese economy, check the development of the data and the government ' s econometric measures for recovery. Until November 2011, we could obtain actual data for the second and third quarters of 2011.

Additionally, the Japanese government raised consumption tax rate during our sample period, on April 2014. The disagreement in the GDP growth nowcast increased from December 2013 to December 2014, and in the four-quarter ahead forecast increased from January 2013 to December 2013. This is because economists were not sure about the magnitude of rushed product purchases ahead of the consumption tax rate hike and the slack in its reaction. Actually, the recovery of GDP after the tax rate was slower than expected. For CPI inflation and the unemployment rate, we also see an increase in disagreement around April 2014. However, the increase was not so large.

3.3 Disagreement and monetary policy

Prior studies show that monetary policy affects disagreement in the CPI inflation rate. In our sample, the BOJ introduced the quantitative qualitative monetary easing (QQE) policy in March 2013. In April 2013, immediately after introducing this measure, disagreement in the three-quarter ahead forecasts of the CPI inflation rate increased.

Disagreement also jumped on September 2014. Although the BOJ expanded QQE at the end of October 2014, this jump in disagreement seemed

unrelated to this policy change and was affected by a downward bearish forecast, the first interquartile rate. The CPI inflation rate in Japan is heavily affected by energy prices, and the WTI oil price began declining since June 2014. The decline in the oil price lowered the CPI inflation forecast by some forecasters, and disagreement increased.

3.4 Disagreement and forecasting error

Figures 7 to 9 compare disagreement and the RMSE of the nowcast. As we checked using the descriptive statistics, the change in RMSE (dotted line) is larger than those of disagreement, especially for the GDP growth rate. The RMSE of the GDP growth rate is large, even in the recovery phase. As we saw, this is partly due to high fluctuations in the GDP growth in Japan.

[Insert Figures 7-9 about here]

For the GDP growth rate, RMSE moves earlier than disagreement. The RMSE of CPI inflation rate and the unemployment rate move simultaneously with disagreement.

Table 3 shows that disagreement and RMSE correlate positively. The correlation of the CPI inflation rate and the unemployment does not decline as much as the forecast horizon becomes longer. On the other hand, the correlation of GDP growth decline as the forecast horizon becomes longer.

[Insert Table 3 about here]

4 Granger causality tests

4.1 Augmented Dicky-Fuller (ADF) test

We test whether monthly disagreement, forecast error, and the VXJ index has a unit root before conducting a Granger causality test. We specify the augmented Dickey—Fuller (ADF) test with up to twelve lags of the dependent variable and with a constant. We use the Akaike Information Criteria (AIC) to select the lag length.

Table 4 reports the result. We reject the null hypothesis of a unit root for disagreement and RMSE of the GDP growth rate and VXJ. On the other hand, we do not reject the null hypothesis of a unit root for disagreement related to CPI inflation and the unemployment rate. Thus, we use Toda and Yamamoto's (1995) methodology to estimate the VAR for CPI inflation and the unemployment rate.

[Insert Table 4 about here]

4.2 Granger causality between disagreement and RMSE: Same horizon

We apply two types of Granger causality test. First, we use disagreement and RMSE with the same horizon and VXJ. We estimate VAR with an earthquake dummy equal to one if the forecast date is from April 2011 to September 2011, and zero otherwise.

Second, we use nowcast RMSE in each VAR regression. In the second type, the horizon of disagreement and RMSE are not the same, except for the regression of nowcast disagreement and RMSE. We estimate the second type of regression to determine the nowcast RMSE before they know the one- to three-quarter ahead RMSE. Through the second test, we can see how forecasters react to the newest information. We also use the earthquake dummy in second test.

Figure 10 shows the result of first test. For the GDP growth rate, the nowcast RMSE Granger-causes disagreement. However, nowcast disagreement does not Granger-cause RMSE. For the GDP growth rate with a longer horizon, there is no Granger causality between disagreement and RMSE. However, RMSE Granger-causes disagreement in terms of the GDP growth rate with the two- to four-quarter average.

[Insert Figure 10 about here]

For the CPI inflation rate, all horizon RMSEs Granger-cause disagreement. On the other hand, the one- to two-quarter ahead disagreement Granger-causes RMSE. For unemployment, the result is mixed: the one-quarter ahead disagreement Granger-causes RMSE, while on the other hand, the two- to three-quarter ahead RMSE Granger-causes disagreement. The VXJ Granger-causes disagreement and RMSE most of cases.

4.3 Granger causality between disagreement and RMSE: All RMSE nowcast

Figure 11 shows the results of the second test. For the GDP growth rate, the nowcast RMSE Granger-causes all horizons except the three-quarter ahead disagreement. Disagreement does not Granger-cause RMSE, and VXJ causes only the nowcast disagreement. VXJ and RMSE Granger-cause each other.

[Insert Figure 11 about here]

For the CPI inflation rate, nowcast RMSE Granger-causes all horizon disagreements. The two- and three-quarters ahead disagreement Granger-causes nowcast RMSE; VXJ causes all horizons, except the two-quarter ahead disagreement; and VXJ and RMSE Granger-cause each other in most cases.

For the unemployment rate, nowcast RMSE does not Granger-cause any disagreement. The one- to three-quarters ahead disagreement Granger-causes nowcast RMSE.

5 Interpretation and conclusion

In this study, we investigate the Granger causality between disagreement and forecasting errors for three macroeconomic indicators. The forecasting errors measured by the cross-sectional RMSE of respondents are much larger than that for disagreement. Thus, diversification among forecasters ' views is not wide enough to foresee the business cycles.

Notably, the forecasting error of GDP growth is much larger than that for disagreement. The forecasting error Granger-causes disagreement, which tends to increase in recessionary phases, so it is hard to foresee a recession with disagreement. We can offer two explanations. One is forecasters ' herding behavior. They hesitate to change their forecasts without other forecasters changing their forecasts. After recognizing the forecast error, they begin to change their attitudes. The other is large and irregular fluctuations in Japan ' s GDP growth rate. Its standard deviation is twice that of the US; the mean growth rate of Japan is about one third that of the US. The fluctuation is too large to make forecasters provide different predictions from others. The release of GDP growth rate data has little power to affect financial markets, such as the stock market. Reform in the estimation process should proceed as the Japanese government investigates. We should use at least a four-quarter average forecast as a longer GDP growth rate expectation instead of the three-quarter ahead forecast.

For the CPI inflation rate, forecasting error also Granger-causes disagreement. Disagreement with a longer horizon causes forecasting error. When the QQE policy began, disagreement in the three-quarter ahead forecast increased and the range of disagreement between the nowcast and the three-quarter ahead forecasts widened. We can interpret this to mean that a change in the information expectation affects the future inflation rate. However, fluctuations in energy prices heavily affect the CPI inflation rate in this study. However, the BOJ considers this price important and provides predictions, we need to survey CPI inflation excluding food and energy as in the US SPF surveys.

While we investigate three macroeconomic indicators because they have been surveyed since the beginning of ESPF survey, the survey now asks for quarterly forecasts for eleven macroeconomic indicators. We should examine these indicators in a future study.

References

[1] Atalla, T., Joutz, F., and Pierru A. (2016). Does disagreement among oil price forecasters reflect volatility? Evidence from the ECB surveys. *International Journal of Forecasting*, 32(4), 1178-1192.

[2] Dovern. J., Fritsche, U., and Slacalek, J. (2012). Disagreement among forecasters in G7 countries. *The Review of Economics and Statistics*, 94(4), 1081-1096.

[3] Komine, T., Ban, K., Kawagoe, M. and Yoshida, H. (2009). What Have We Learned from a Survey of Japanese Professional Forecasters? – Taking Stock of Four Years of ESP Forecast Experience. *ESRI Discussion Paper Series*, 214, 1-26.

[4] Lahiri, K., Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, 25(4), 514-538.

[5] Legerstee, R. and Franses, H. P. (2015). Does disagreement amongst forecasters have predicted value? *Journal of Forecasting*, 34(4), 290-302.

[6] Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. *NBER Macroeconomics Annual*, 18, 209-248.

[7] Patton, A. J. and Timmermann, A. (2010). Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, 57(7), 803-820.

[8] Siklos, P. L. (2013). Sources of disagreement in inflation forecasts: An international empirical investigation. *Journal of International Economics*, 90(1), 218-231.

[9] Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1), 225-250.

[10] Zarnowitz, V. and Lambros, L .A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, 95(3), 591-621.

Table 1: Forecasting target and release date (Example Target=4Q of 2005)

		Release date
$h = 0$	$v = 1$	December 05
(Nowcast)	$v = 2$	November 05
	$v = 3$	October 05
$h = 1$	$v = 4$	September 05
	$v = 5$	August 05
	$v = 6$	July 05
$h = 2$	$v = 7$	June 05
	$v = 8$	May 05
	$v = 9$	April 05
$h = 3$	$v = 10$	March 05
	$v = 11$	February 05
	$v = 12$	January 05

Figure 1: Actual and nowcast: GDP growth rate

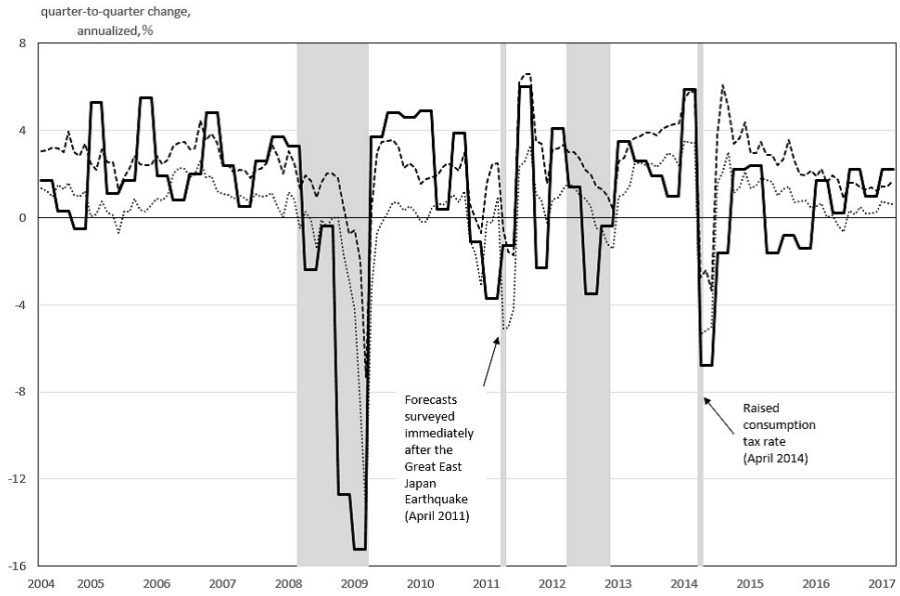

Figure 2: Actual and nowcast: CPI inflation rate

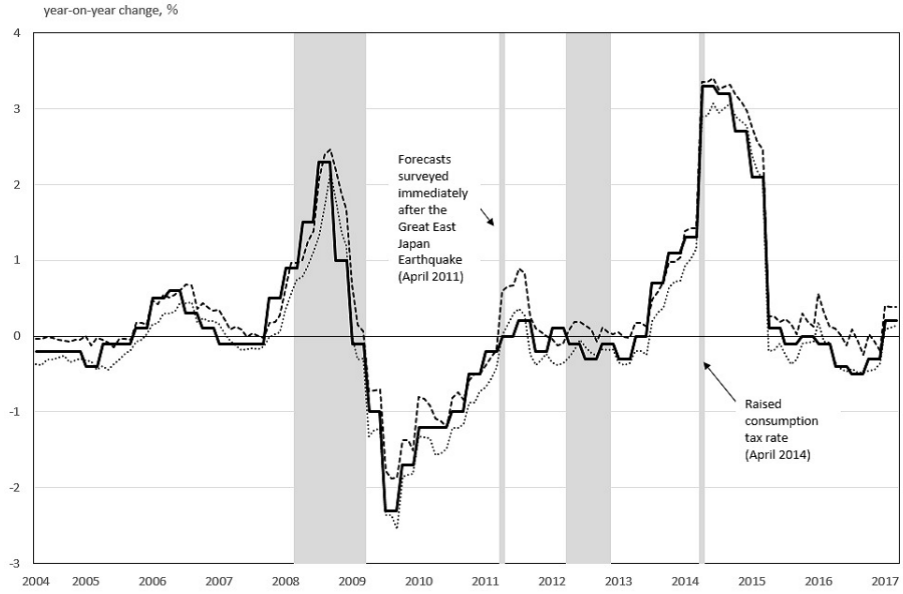

Figure 3: Actual and nowcast: Unemployment rate

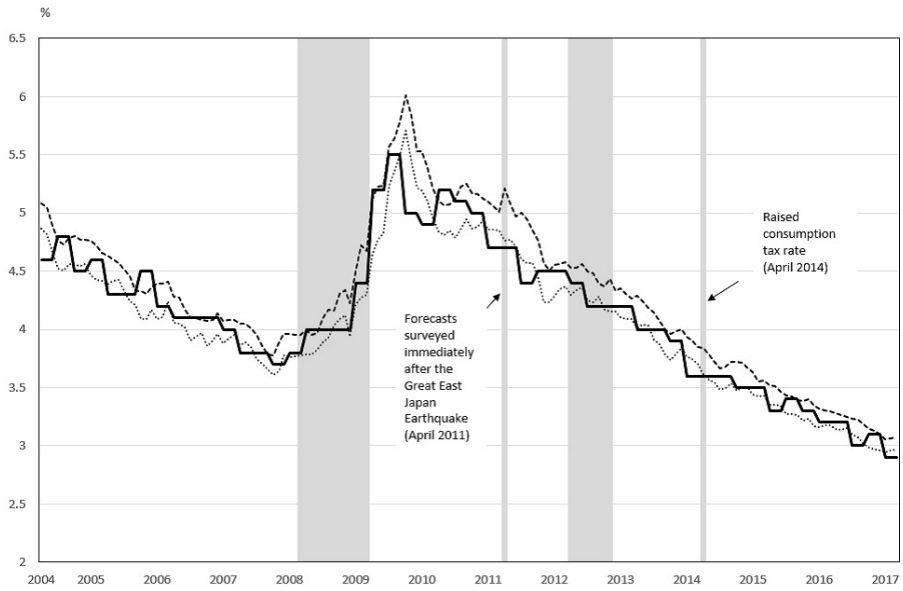

Table 2: Summary statistics

Statistic	Horizon	GDP growth	CPI inflation	Unemployment rate
Average disagreement	0	0.98	0.17	0.11
	1	0.86	0.22	0.14
	2	0.83	0.26	0.17
	3	0.82	0.29	0.19
Average RMSE	0	2.57	0.25	0.18
	1	2.84	0.42	0.22
	2	2.90	0.57	0.27
	3	2.88	0.70	0.32
Average level of actual		0.88	0.19	4.10
Standard deviation of actual		4.00	1.07	0.63

Figure 4: Disagreement in the GDP growth rate

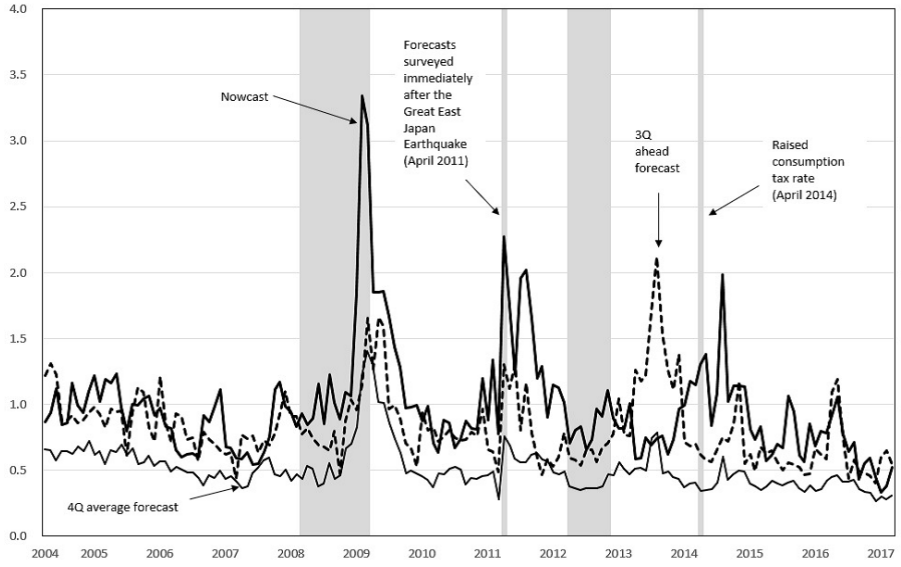

Figure 5: Disagreement in the CPI inflation rate

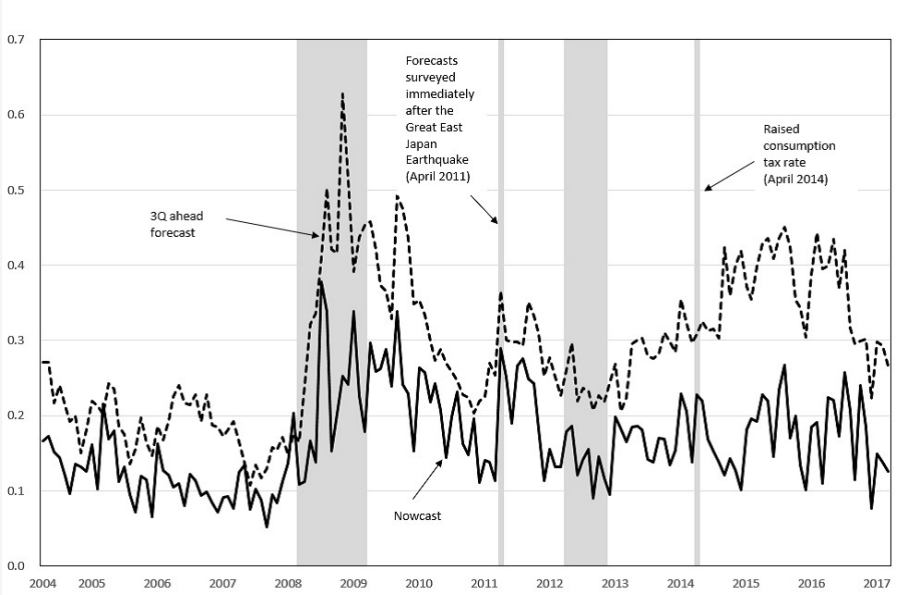

Figure 6: Disagreement of the Unemployment rate

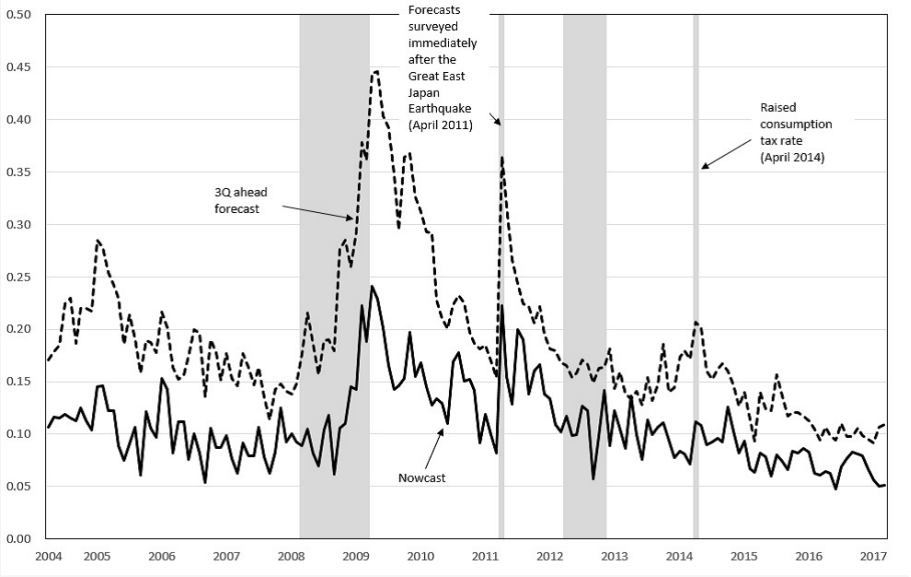

Figure 7: Disagreement and RMSE in the GDP growth rate

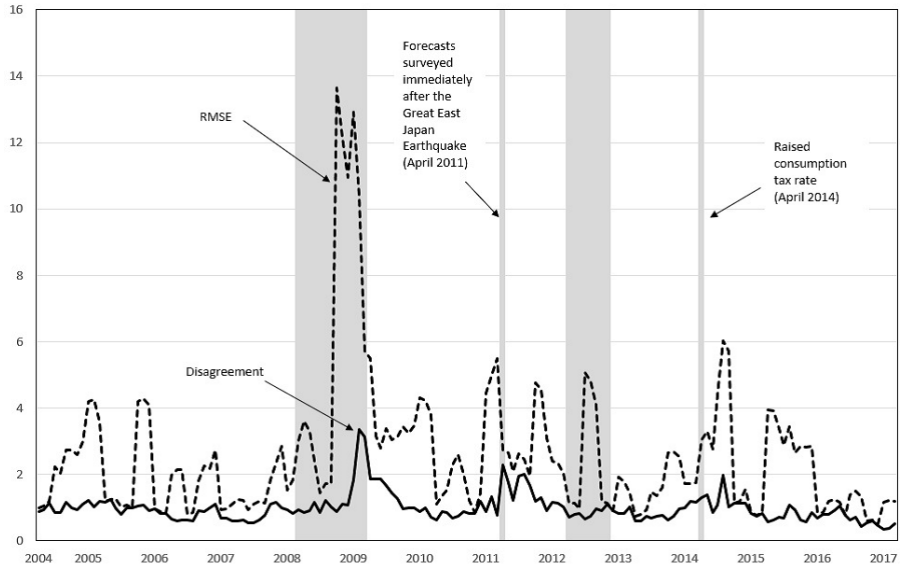

Figure 8: Disagreement and RMSE in the CPI inflation rate

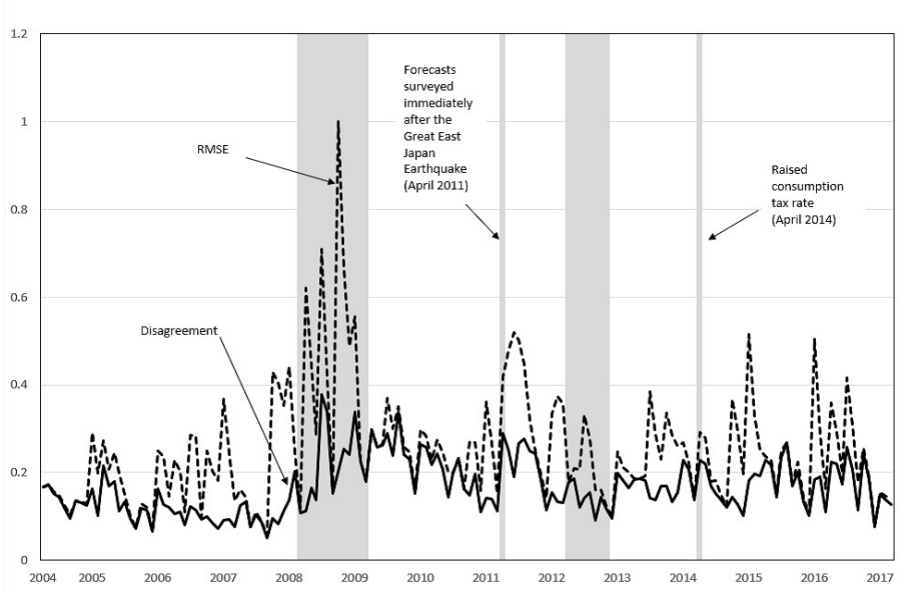

Figure 9: Disagreement and RMSE in the Unemployment rate

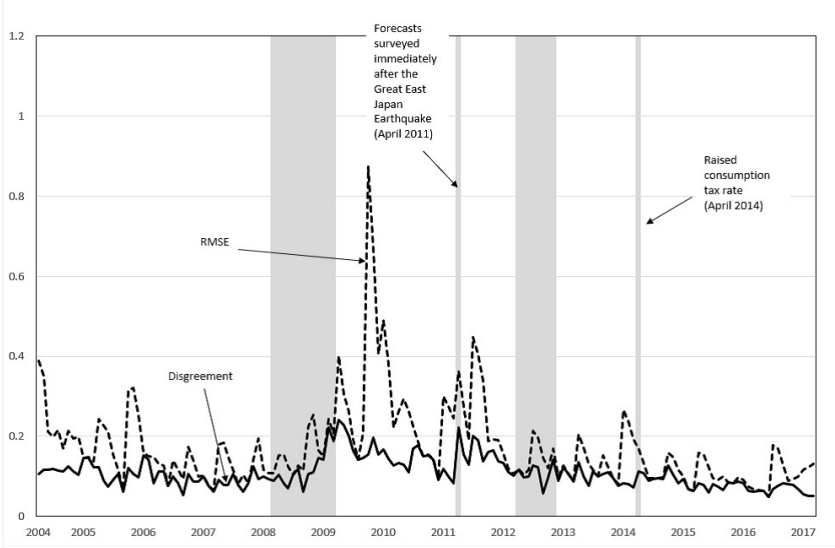

Table 3: Correlation between disagreement and RMSE

Horizon	GDP growth	CPI inflation	Unemployment rate
0	0.43***	0.52***	0.60***
1	0.21**	0.34***	0.70***
2	0.17**	0.35***	0.70***
3	0.06	0.43***	0.54***
$2q$ avg	0.17**		
$3q$ avg	0.08		
$4q$ avg	0.03		

Note: The null hypothesis of no correlation is rejected at the 10% (*), 5% (**), 1% (***) significance levels.

Table 4: ADF unit root tests for disagreement and RMSE

	GDP growth			CPI inflation			Unemployment		
	L	ADF	AIC	L	ADF	AIC	L	ADF	AIC
					Disagreement				
$H = 0$	0	-4.78***	0.378	8	-2.11	-3.268	9	-1.90	-4.613
$H = 1$	0	-4.10***	-0.377	11	-1.87	-3.410	7	-1.59	-4.272
$H = 2$	0	-4.37***	-0.472	4	-2.07	-3.524	3	-2.28	-4.241
$H = 3$	4	-4.50***	-0.349	2	-2.03	-3.391	3	-2.12	-4.119
$2q$ avg	4	-3.69***	-0.839						
$3q$ avg	0	-3.54***	-1.609						
$4q$ avg	0	-3.40**	-1.985						
					RMSE				
$H = 0$	1	-4.87***	3.532	6	-3.26**	-1.686	6	-2.68*	-2.226
$H = 1$	6	-2.85*	3.860	6	-3.18**	-0.766	1	-3.89***	-2.148
$H = 2$	6	-2.66*	4.000	5	-2.64*	-0.264	9	-2.19	-1.981
$H = 3$	6	-2.84*	4.019	6	-2.55	-0.175	0	-3.16**	-1.554
$2q$ avg	0	-4.02***	3.387						
$3q$ avg	6	-2.88**	2.825						
$4q$ avg	0	-2.91**	2.517						

Note: The null hypothesis is rejected at the 10% (*),5% (**), 1% (***) significance levels.

Figure 10: Granger causality test results: Same horizon

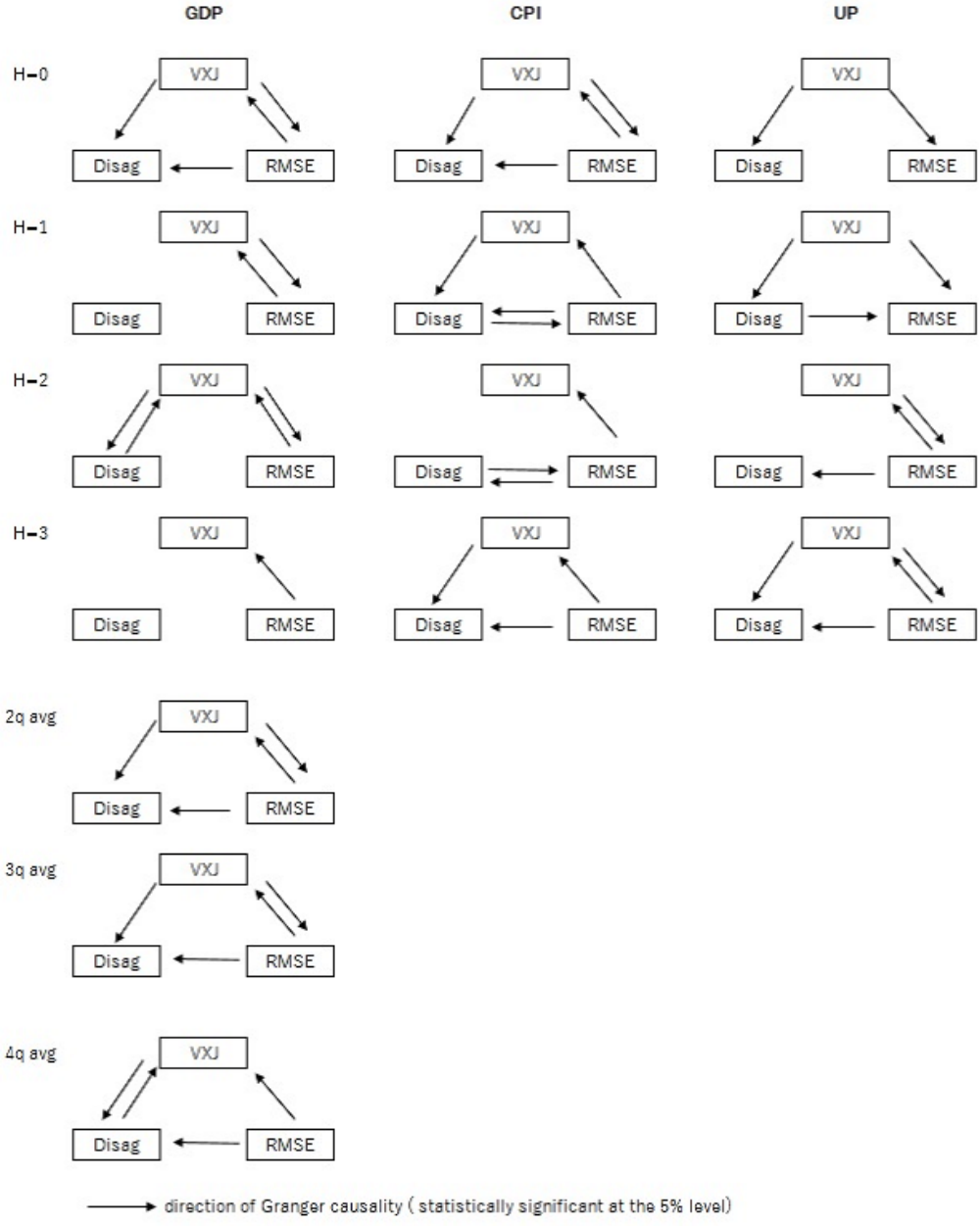

Figure 11: Granger causality test results: All RMSE nowcast