Forecast Accuracy and Product Differentiation of Japanese Institutional Forecasters

Ashiya, Masahiro

(Citation)
神戸大学経済学研究科 Discussion Paper,428

(Issue Date)
2004-11

(Resource Type)
technical report

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/80200021

Forecast Accuracy and Product Differentiation of Japanese Institutional Forecasters

Masahiro ASHIYA [+]

November 2004

JEL Classification Codes: E37; C53; E17.

Keywords: Macroeconomic Forecast; Forecast evaluation; Analysis of variance.

[+] Faculty of Economics, Kobe University, 2-1 Rokko-dai, Nada, Kobe, 657-8501, Japan;
E-mail: ashiya@econ.kobe-u.ac.jp

Forecast Accuracy and Product Differentiation of Japanese Institutional Forecasters

This paper investigates whether some forecasters consistently outperform others using real GDP forecast data of 53 Japanese institution over the past 24 years. It finds that the accuracy rankings are not significantly different from those that might be expected when all institutions had equal forecasting ability. On the other hand, some institutions are relatively optimistic and others are pessimistic throughout the sample period. These results suggest that the macroeconomic forecasting business is competitive and each institution chooses the degree of "product differentiation" of its forecast so that accuracy and publicity are optimally balanced.

JEL Classification Codes: E37; C53; E17.

Keywords: Macroeconomic Forecast; Forecast evaluation; Analysis of variance.

1. Introduction

It is a regular year-end event in Japan that institutional forecasters release their economic forecasts for the ongoing fiscal year and for the next fiscal year. Some of these forecasts turn out to be accurate, and others not. A natural question is whether there are institutions that consistently outperform others. This is the first issue we address using the track records of real GDP forecasts over the past 24 years.

We select the data of 53 institutions, which participate in ten or more surveys (See Section 2 for details). Forecast accuracy varies considerably among these institutions. Table 1 shows the descriptive statistics of the mean absolute forecast error (MAFE) of each institution for the current fiscal year and for the next fiscal year. As for the current-year forecast, MAFE of the best institution is 0.373%, and that of the worst institution is 0.700%. The average of MAFE is 0.529% and the standard deviation is 0.069. As for the year-ahead forecast, the best is 1.110% and the worst is 1.833%.

Are these differences in forecast accuracy large enough to suggest that there are differences in forecasting ability among these institutions? To answer this question, we must take it into account that some years are more difficult to forecast than others. The variance of absolute forecast errors tends to be larger in these difficult-to-forecast years. It follows that the level of MAFE is mainly determined by the performance in the difficult-to-forecast years. Therefore MAFE is not an appropriate measure of forecast accuracy.

We consider accuracy ranking of the institutions instead. We employ the non-parametric test of ranking developed by Skillings and Mack (1981), which is robust to changes in the variance of the forecast variables. The result in Section 3 shows that the accuracy rankings are not significantly different from those that might be expected when all institutions had equal forecasting ability. Namely we cannot reject the hypothesis that all forecasters are equal. Our result strengthens those of Batchelor (1990), Batchelor and Dua (1990a, b), and Kolb and Stekler (1996), since our data cover 24 years while their data contain at most 11 years.

Next we investigate whether some institutions are consistently more optimistic than others. When the institutions are ranked according to the relative levels of their forecasts, we find this ranking significantly different from random one. Namely, some institutions are consistently optimistic and others are consistently pessimistic relative to the industry average. [1] This result is remarkable in that the ranking was stable throughout the sample period, which covers four business cycles

What is the economic rationale behind these results? As Batchelor and Dua (1990b) and Laster et al. (1999) suggest, institutional forecasters pursue two objectives: accuracy and publicity. A conventional forecast never comes to public notice, but an extreme forecast attracts public attention if it turns out to be true. Hence forecasters have an incentive to be extreme. It is unwise, however, to be too extreme if the macroeconomic forecasting business is competitive. An extreme forecast is less accurate than the consensus forecast (as Zarnowitz and Braun (1993) document), and inaccurate forecasters will be driven out of business if the forecasting industry is competitive. Consequently each institution chooses the degree of extremeness of its forecast so that accuracy and publicity are optimally balanced.

The paper is organized as follows. Section 2 explains the data. Section 3 evaluates the variance of forecast accuracy, and Section 4 evaluates the variance of forecast level. Section 5 concludes.

2. Data

Toyo Keizai Inc. has published the forecasts of about 70 Japanese institutions in the February or March issue of "Monthly Statistics (Tokei Geppo)" since 1970s. [2] In every December, institution i releases forecasts of the Japanese real GDP growth rate for the ongoing fiscal year and for the next fiscal year. We call the former $f_{t,t}^{i}$ and the latter $f_{t,t+1}^{i}$. For example, February 2004 issue contains forecasts for fiscal year 2003 (from April 2003 to March 2004) and for fiscal year 2004 (from April 2004 to March 2005). We treat the former as $f_{2003,2003}^{i}$ and the latter as $f_{2003,2004}^{i}$.

To avoid the effect of the second Oil Shock, we use the forecasts published from February 1981 on. That is, we use $f_{t,t}^{i}$ for the fiscal years 1980 through 2003 and $f_{t,t+1}^{i}$ for the fiscal years 1981 through 2003. We exclude institutions that participate in less than 10 surveys, leaving 53 institutions. The average number of observations per institution is 18.42 for current-year forecast ($f_{t,t}^{i}$) and 18.21 for year-ahead forecast

($f^{i}_{t,t+1}$).

As for the actual growth rate g_t, Keane and Runkle (1990) argue that the revised data introduces a systematic bias because the extent of revision is unpredictable for the forecasters (See also Stark and Croushore (2002)). For this reason we use the initial announcement of the Japanese government usually released in June. [3] The Japanese economy experienced four business cycles in our sample period: the peaks were 1984, 1990, 1996, and 2000, and the troughs were 1981, 1986, 1993, 1998, and 2001. [4]

3. Forecast accuracy

To test whether all institutional forecasters were equally able in forecasting, we consider accuracy ranking of the institutions. Skillings and Mack (1981) generalize Friedman's (1937) distribution-free test and develop the following non-parametric test applicable to unbalanced panels.

Suppose the panel data consists of *N* institutions and *M* periods. Let N_t ($\leq N$) be the number of institutions that release forecasts in year *t*. Let $r_t^i \in \{1, \cdots, N_t\}$ be the rank of the absolute forecast error of institution *i* in year *t*. If ties occur, we use average ranks. If institution *i* does not participate in year *t*, we assume $r_t^i = 0.5(1+N_t)$. We define the adjusted rank of institution *i* in year *t*, A_t^i, as

$$A_t^i \equiv \left(\frac{12}{1+N_t}\right)^{0.5} \left(r_t^i - \frac{1+N_t}{2}\right).$$

The first term of A_t^i compensates for the difference in observations. The second term measures relative performance. A negative (positive) A_t^i indicates that the rank of institution *i* in year *t* is above (below) the middle. $A^i \equiv \sum_{t=1}^{M} A_t^i$ denotes the sum of the adjusted rank. If A^i is close to zero, forecast accuracy of institution *i* is on average similar to other institutions. If A^i is significantly smaller (larger) than zero, forecast accuracy of institution *i* is on average better (worse) than other institutions. Let $A \equiv \left(A^1, \cdots, A^N\right)'$.

Next we consider the covariance matrix, V, of the random vector A. Define m_{ij} as the number of years containing forecasts from institution i and j. Then the elements of V, σ_{ij}, are defined as

$$\sigma_{ij} = \begin{cases} -m_{ij} & \text{if} \quad i \neq j \\ \sum_{k \neq i} m_{ik} & \text{if} \quad i = j \end{cases}.$$

Let V_{11} be the upper left $N-1$ by $N-1$ submatrix of V, and let V_{11}^{-1} be the inverse of V_{11}. Define $\hat{A} \equiv \left(A^1, \cdots, A^{N-1}\right)'$. Skillings and Mack (1981) show that, under the null hypothesis that there is no difference in forecasting ability, the statistic

$$S \equiv \hat{A}' V_{11}^{-1} \hat{A}$$

has an asymptotic chi-squared distribution with $N-1$ degrees of freedom. Significantly large S indicates that institutions were not equal in forecasting ability.

The first row of Table 2 presents the values of S-statistic calculated from the absolute forecast errors of 53 institutions over 24 years. We obtain $S = 48.28$ for current-year forecasts and $S = 40.87$ for year-ahead forecasts. Since neither statistic is significant at 0.10 level (their p-values are larger than 0.60), we find no significant difference in the forecasters' abilities.

To confirm the above result, we estimate the following fixed-effects model used by O'Brien (1990). Let $AFE_{t,t}^{i} \equiv \left| f_{t,t}^{i} - g_t \right|$ be the absolute forecast error of the current-year forecast made by institution i in year t. $dum^i(j)$ ($j = 1, \cdots, 52$) denotes the individual dummy and

$$dum^i(j) = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases}.$$

$dum_t(s)$ ($s = 1980, \cdots, 2002$) denotes the year dummy and

$$dum_t(s) = \begin{cases} 1 & \text{if} \quad s = t \\ 0 & \text{otherwise} \end{cases}.$$

The regression we consider is

$$AFE_{t,t}^{i} = \alpha + \sum_{j=1}^{52} \beta_j \cdot dum^i(j) + \sum_{s=1980}^{2002} \gamma_s \cdot dum_t(s) + u_{t,t}^{i}.$$

If β_j is significantly smaller (larger) than zero, absolute forecast error of institution j is smaller (larger) than other institutions conditional on year effects. The null hypothesis is $\beta_1 = \cdots = \beta_{52} = 0$, i.e., institutions are homogeneous in average forecast accuracy.

The second row of Table 2 presents the result of *F*-test on the coefficients of the individual dummies. It shows that the individual effect is not significant at 0.10 level. Hence there is no evidence that institutions differ systematically in forecast accuracy.

4. Forecast level

This section examines whether some institutions consistently release optimistic (or pessimistic) forecasts. To address this question, the observed distribution of their forecasts is compared with the distribution expected if their relative forecast levels each year were purely random.

First we employ the ranking-based test of Skillings and Mack (1981) explained in Section 3. The first row of Table 3 shows the values of *S*-statistic. We obtain $S = 81.39$ for current-year forecasts and $S = 196.14$ for year-ahead forecasts, both of which are significant at 0.01 level. It indicates that some forecasters were relatively optimistic while others were relatively pessimistic during the sample period.

We obtain the same result when we consider the following fixed-effects model:

$$f_{t,t}^{i} = \alpha + \sum_{j=1}^{52} \beta_j \cdot dum^i(j) + \sum_{s=1980}^{2002} \gamma_s \cdot dum_t(s) + u_{t,t}^{i}.$$

The null hypothesis is $\beta_1 = \cdots = \beta_{52} = 0$, i.e., institutions are homogeneous in average forecast level conditional on year effects. The second row of Table 3 shows the result of *F*-test. The null hypothesis is clearly rejected for both current-year and year-ahead forecasts.

One remaining question is whether each institution chooses the same relative forecast level for current-year forecasts and year-ahead forecasts. Let $A_t^i\left(f_{t,t}^i\right)$ be the adjusted rank of forecast level of $f_{t,t}^i$ (defined in Section 3), and let $A_t^i\left(f_{t,t+1}^i\right)$ be the adjusted rank of forecast level of $f_{t,t+1}^i$. Then the correlation between $A_t^i\left(f_{t,t}^i\right)$ and $A_t^i\left(f_{t,t+1}^i\right)$ is 0.605, which indicates that an institution optimistic (pessimistic) for the current year is also optimistic (pessimistic) for the next year.

5. Conclusions

This paper analyzes the real GDP forecast record of 53 Japanese institutions over 24 years. First we evaluate relative forecast accuracy among these institutions. When performance is measured by the mean absolute forecast error, there appears to be significant difference in accuracy among them (Table 1). The difference vanishes, however, when either a ranking-based test is employed or year effects are explicitly considered (Table 2). Namely we find that all institutions are equal in forecasting ability. We next evaluate relative forecast level among these institutions. Our data provide clear evidence that some institutions are consistently optimistic while others are consistently pessimistic relative to the industry average (Table 3).

These findings are in accordance with the notion that the forecasting industry is competitive and each institution seeks publicity. The competitive pressure eliminates inaccurate forecasters, and consequently all remaining institutions are equal in ability. These institutions try to draw public attention without compromising forecast accuracy by modestly differentiating their forecasts from others.

Notes

1. Batchelor and Dua (1990b) examine the Blue Chip forecasters for 11 years and find similar results.
2. Ashiya (2002, 2003, forthcoming) also uses the data from "Monthly Statistics (Tokei Geppo)".
3. We obtain the same results by using the revised data of g_t released in June of year $t+2$.
4. The initial announcements of the actual growth rates for fiscal years 1980 to 2003 were 3.8%, 2.7%, 3.3%, 3.7%, 5.7%, 4.2%, 2.6%, 4.9%, 5.1%, 5.0%, 5.7%, 3.5%, 0.8%, 0.0%, 0.6%, 2.3%, 3.0%, -0.7%, -2.0%, 0.5%, 0.9%, -1.3%, 1.6%, and 3.2%.

References

Ashiya, Masahiro, 2002. "Accuracy and Rationality of Japanese Institutional Forecasters," Japan and the World Economy, 14 (2), April, pp.203-213.

Ashiya, Masahiro, 2003. "Testing the Rationality of Japanese GDP Forecasts: The Sign of Forecast Revision Matters." Journal of Economic Behavior and Organization, 50, 263-269.

Ashiya, Masahiro, forthcoming. "Twenty-two Years of the Japanese Institutional Forecasts." Applied Financial Economics Letters.

Batchelor. Roy A., 1990. "All Forecasters Are Equal," Journal of Business and Economic Statistics, 8(1), 143-144.

Batchelor, Roy A., Dua, Pami, 1990a. "Forecaster Ideology, Forecasting Technique, and the Accuracy of Economic Forecasts," International Journal of Forecasting, 6, 3-10.

Batchelor, Roy A., Dua, Pami, 1990b. "Product Differentiation in the Economic Forecasting Industry," International Journal of Forecasting, 6, 311-316.

Friedman, Milton, 1937. "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," Journal of the American Statistical Association. 32, 675-701.

Keane, M.P., Runkle, D.E., 1990. "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." American Economic Review, 80, 714-735.

Kolb, R.A., Stekler, H.O., 1996. "How Well Do Analysts Forecast Interest Rates?" Journal of Forecasting, 15, 385-394.

Laster, David, Bennett, Paul, and Geoum, In Sun, 1999. "Rational Bias in Macroeconomic Forecasts," Quarterly Journal of Economics, 114, 293-318.

O'Brien, Patricia C., 1990. "Forecast Accuracy of Individual Analysts in Nine Industries," Journal of Accounting Research, 28(2), 286-304.

Skillings, John H., Mack, Gregory A., 1981. "On the Use of a Friedman-Type Statistic in Balanced and Unbalanced Block Designs," Technometrics, 23(2), 171-177.

Stark, Tom, Croushore, Dean, 2002, "Forecasting with a Real-time Data Set for Macroeconomists." Journal of Macroeconomics, 24, 507-531.

Zarnowitz, Victor, Braun, Phillip, 1993. "Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance." in Stock, James H., Watson, Mark W. eds. *Business cycles, indicators, and forecasting. NBER Studies in Business Cycles, vol. 28.* Chicago and London.

Table 1: Mean absolute forecast error of 53 institutions

	Current year	Year-ahead
Min.	0.373	1.110
Max.	0.700	1.833
Avg.	0.529	1.403
STD	0.069	0.144

Table 2: Variance tests for the absolute forecast error

	Current year	Year-ahead
S-statistic	48.28	40.87
(*p*-value)	(0.621)	(0.867)
F-test	1.163	0.813
(*p*-value)	(0.204)	(0.824)

Table 3: Variance tests for the forecast level

	Current year	Year-ahead
S-statistic	81.39	196.14
(*p*-value)	(0.006)	(0.000)
F-test	1.453	4.372
(*p*-value)	(0.022)	(0.000)