An Evaluation of Chinese Economic Forecasts

H.O. Stekler[1]

Huixia Zhang[2]

Department of Economics

2115 G St. NW

George Washington University

Washington DC 20052

1Corresponding author, hstekler@gwu.edu, telephone: +202-994-6150, fax: +202-994-6147; [2]

summer.hx7@gmail.com

An Evaluation of Chinese Economic Forecasts

ABSTRACT

This paper analyses the real GDP growth and inflation forecasts prepared by the International Monetary Fund, the Organization for Economic Cooperation and Development, and the private forecasters between 1999 and 2010. The empirical results show that the longer term growth forecasts were inferior to a naïve model and were biased. The average of a number of short-term private forecasts dominated those made by the international organizations.

Key words: Chinese forecasts, IMF forecasts, OECD forecasts

An Evaluation of Chinese Economic Forecasts

Although many forecasts about the Chinese economy have been made, our knowledge about the accuracy of these predictions is limited. These forecasts have been prepared by international organizations such as the IMF, which publishes the *World Economic Outlook*, the OECD, private consulting firms, and Chinese economic forecasters. While the forecasts of the WEO and OECD have been analyzed extensively, the main focus of these studies has been on the predictions about the G-7 countries (Artis, 1988, 1997; Ash, Smyth and Heravi, 1990, 1998; Ashiya, 2006; Barrionuevo, 1993; Batchelor, 2001; Holden, Peel and Sandu, 1987; Kreinin, 2000; Loungani, 2001; Oller, 2000; Pons, 2000; Timmermann, 2007). None of these analyses had specifically focused on the forecasts about the Chinese economy nor included the years of the Great Recession. The comparisons of the international organizations forecasts with those made by the private sector were either in other contexts or part of a broader analysis.[1]

This study compares the accuracy of economic forecasts for China produced by (1) the International Monetary Fund (IMF), (2) the Organization for Economic Cooperation and Development (OECD), and (3) private sector economists from 1999 to 2010. The main contributions of the paper will be knowledge about the quality of the forecasts of the Chinese economy and the comparative accuracy of these private and international organization predictions.

The next section presents the data and section 2 describes the methodology. The following section contains a series of results about the quality of the forecasts. The concluding section relates these findings to the results of previous forecast evaluations.

---

[1] Batchelor (2001) and Loungani (2001) compared the forecasts of the international organizations with private sector predictions contained in *Consensus Forecasts*. While Batchelor only examined the G7 estimates, Loungani analyzed the forecasts of both advanced and emerging nations. Neither specifically focused on the China projections. Timmermann (2007) compared the Consensus and WEO forecasts in his analysis of *all* WEO projections.

1. Data

We examine the forecasts about the Chinese economy issued between 1999 and 2010 by (1) two international organizations, the IMF and OECD, and (2) a consensus obtained from Chinese individuals. The analysis will concentrate on the forecasts of two variables: the annual growth rate of real gross domestic product (real GDP) and the annual inflation rate (Consumer Price Index).

Both the IMF and OECD publish their forecasts twice a year. In each instance, two forecasts of the *annual* real growth rate (rate of inflation) are prepared: one for the current year; the other for the following year. The IMF issues its predictions in the *World Economic Outlook* which is usually published in April and September. In April the forecasts have leads of nine and twenty-one months relative to the end of the years to which they refer. The September predictions have four and sixteen month leads. Thus it is possible to calculate the relationship between forecast accuracy and the forecasting horizon.

The *OECD Economic Outlook* is the OECD's semi annual analysis of the major economic trends and projections for the next two years. It is published in June and December and forecasts the output, employment, prices and current balances for each member country and selected non-member countries. Although China is not a member of OECD, the OECD has consistently forecasted China's output since 1980. The June forecasts have a 7-month lead for the current year and a 19-month lead for the next year. Similarly, the leads for the December projections are one month and 13 months, respectively. We can also analyze the accuracy- lead relationship for these projections, but the IMF and OECD forecasts are not strictly comparable because they are not issued at the same time.

We also examine the consensus forecasts made by the private sector. We have the forecasts that were issued monthly for the period 1999 to 2008. It is thus possible to compare the international organizations' predictions with each other and also with the average (consensus) of those made by the private sector for this shorter period..

These forecasts are compared with the outcomes as measured by the first published numbers, not the historical revised data. These data are published in the April WEO reports. Thus forecasts for the current years are compared with the numbers published in April of the next years; the actual outcomes for the t+1 years come from the April t+2 publications.

2. Methodology

2.1 Accuracy Measures and Bias Test

We use conventional methodologies to determine the nature of the lead-accuracy relationships, whether the forecasts exhibit bias, and whether one set of forecasts encompasses another one. Both the mean absolute error (MAE) and the mean squared error (MSE) are used as measures of accuracy. They are then compared across forecasting horizons. As a benchmark and basis of comparison, we use Theil's U statistic, $U = \dfrac{\sum (A-F)^2}{\sum (A_t - A_{t-1})^2}$. In this case the benchmark is the naïve prediction (F^n), $F_t = \Delta A_{t-1}$, i.e. the same change that was observed last year is the forecast for this year.[2]

As a caveat, we should note that a precise quantitative comparison of the lead-accuracy relationship for the two international organizations is impossible because these forecasts were not prepared and issued at the same times. Moreover, the time periods for which all four sets of forecasts are available differ. However, it will still be possible to draw some qualitative

---

[2] In the tables we present another benchmark, Theil's coefficient U_1 which is based on the errors made in predicting that there will be no growth from year to year (i.e. no change). Given the growth in the Chinese economy over this period, this is not an appropriate benchmark.

conclusions about these relationships. *Thus the comparative results should only be considered indicative and not precise.*

We use two tests to determine whether the forecasts contain systematic errors. The first test uses the Mincer-Zarnowitz (1969) equation (1) to test for bias:

$$A_t = \beta_0 + \beta_1 F_t + \varepsilon_t , \qquad (1)$$

A_t and F_t are the forecast and actual values, respectively. The null of no bias is that $\beta_0 = 0$ *and* that $\beta_1 = 1$. While this equation provides the sufficient condition for bias, the necessary condition for no bias is that $\varepsilon_t = 0$ and is called the Holden-Peel test. (Holden-Peel, 1990).

2.2 Encompassing Tests

A comparison of the accuracy measures of two sets of forecasts is only one of several approaches for evaluating the different predictions. The use of encompassing tests is another approach, and it enables us to determine whether or not one set of forecasts contains information that is not included in the other predictions. One set of forecasts is said to encompass the other if it contains all the useful information about the outcomes and the other provides no additional information. If, however, neither forecast encompasses the other, a linear combination of the two would outperform either one. (Granger and Ramanathan, 1984).

Assuming that there are two forecasts, F_1 and F_2, two restricted tests are performed on the linear combinations of the two:

$$A_t = \beta_0 + \beta_1 F_{1,t} + \beta_2 F_{2,t} + \varepsilon_t , \qquad (2)$$

In the first case, if the null is that $\beta_1 = 1$ and $\beta_2 = 0$, then F_1 encompasses F_2 if the null is **not** rejected. On the other hand, if the null is that $\beta_1 = 0$ and $\beta_2 = 1$, then F_2 encompasses F_1 if that null is **not** rejected.

3. Results

    3.1 Accuracy Measures

Table 1.a presents the mean absolute errors (MAE) of the real GDP forecasts of the two

international organizations and the average of the individual Chinese forecasts.[3] We postpone the

discussion of the comparative performance of the various sets of forecasts and focus on the

horizon-accuracy relationship. In most cases the MAE become bigger as the forecast horizon

increases. For all forecast sets, the accuracy of the *current* year forecasts improved substantially

as the horizon decreased. The errors of the longer horizon forecasts, with leads of seven to nine

months, are more than double those made closer to the end of the current year.

    The same result was not observed in the year-ahead predictions. In each case, the

forecasts of all the year-ahead forecasts display similar size errors. For example, there is no

substantial change in the accuracy of the WEO (OECD) forecasts made with a lead of 16 (13)

months relative to those that were made 21 (19) months in advance. Similar results were

observed in the private sector consensus projections. Moreover, the errors of the inflation

forecasts (Table 1.b) show the same patterns.[4] All of these results are consistent with the findings

of Isiklar and Lahiri (2007) and are not surprising since there is very little increase in the

available information relating to next year.

    3.2 Bias

    The results from the Mincer-Zarnowitz and Holden-Peel tests for bias are presented in

Table 2. In virtually all cases, the null that the year-ahead growth forecasts are unbiased is

rejected at the 10% level or lower. The current year forecasts display mixed results, with all of

---

[3] All of the other descriptive statistics and error measures are also in Tables1.a and 1.b
[4] We did not have a consensus of private inflation forecasts.

the WEO forecasts biased. The consensus forecasts and the very short horizon OECD predictions were unbiased. We do not present the results for the inflation forecasts, but they were unbiased.

3.3 Comparative Performance

The forecasts from the various organizations were compared with each other and with the average of the private forecasts. As a caveat, we note that the forecast horizons of the two international organizations were different and, moreover, there were only 10-12 observations that could be used in this analysis. Our results should, therefore, be considered suggestive rather than definitive.

The comparative accuracy of the current year growth forecasts depended crucially on the length of the forecast horizon. See Table 1. For example the OECD predictions made one month before the end of the year were more accurate than the WEO projections made three months earlier. Similarly, those WEO projections showed smaller errors than the OECD estimates made with a seven months horizon, etc. In each case, the errors of the international organizations' current year forecasts were of the same order of magnitude as those obtained from averaging the individual private forecasts. The OECD real GDP year-ahead forecasts were somewhat more accurate than those of either the WEO or the consensus, which were comparable.

As a benchmark these forecasts can be compared with the naïve models. In every instance the forecasts were superior to the naïve no change projection. (See the entries for U_1 in Table 1.a). However, all the forecasts which were made at a horizon greater than seven months were inferior to the same change naïve forecast as measured by U_2 (Table 1.a).

The errors of the inflation forecasts also increase with longer horizons, but some of the other results are quite different. (Table 1.b). The WEO current year inflation forecasts are more accurate than those of the OECD even though the forecast horizon of the WEO is longer. The

year-ahead errors of both organizations are substantially the same and they were superior to both naïve forecasts.

On the basis of the evidence obtained from these descriptive statistics we are able to conclude: (1) that the errors made in forecasting both real GDP and inflation increase with the length of the forecasting horizon; (2) at each forecast horizon, the errors made by the international organizations were similar to those made by individuals; (3) however, we were not able to determine whether the forecasts of one organization was superior to those of the other; (4) the real GDP predictions made more than six months before the end of the current year are less accurate than those obtained by extrapolating a naïve model.

3.4 Encompassing Tests

The results of the encompassing tests are presented in Table 3.[5] The results relating to the short-term growth forecasts depend crucially on the forecast horizons that are used to make the comparison and demonstrate that real-time information is so crucial in making forecasts. Thus the OECD forecasts made at a one month horizon dominate those made by the WEO with a four month horizon. But the WEO estimates made with a four month horizon dominate the OECD predictions made with a seven month lead. The results of the other encompassing tests suggest that the OECD real GDP forecasts are generally superior. On the other hand, a comparison of the inflation forecasts shows that the WEO short-term projections encompass those made by the OECD. There were no definitive results for the year-ahead inflation estimates.

A comparison of the private consensus and the international organizations' forecasts yielded only one conclusive result. The consensus growth predictions made with the same forecasting leads dominated both organizations' very short-term growth forecasts. Otherwise

---

[5] The actual regressions may be obtained from the authors.

there were no significant results.[6] This finding differs somewhat from Batchelor's (2001) results that the consensus private G-7 forecasts generally dominated those of the international organizations. He showed that the private forecasts were more accurate and less biased. However, as indicated above, there were only 10 observations in our sample that we could compare.

4. Conclusions

The findings from our evaluation of these forecasts of the Chinese economy are consistent with those obtained from studies that examined the G7 predictions. (See Fildes and Stekler, 2002): (1) forecast errors increased with the length of the horizon; (2) the longer term growth forecasts were inferior to a naïve model; (3) the year-ahead growth forecasts rejected the null of no bias; and (4) the average of a number of short-term private forecasts dominated those made by the international organizations.

Bibliography

Artis, M.J. (1988). How accurate is the *World Economic Outlook*? A post mortem on short-term forecasting at the International Monetary Fund, *Staff Studies for the World Economic Outlook*, (International Monetary Fund), 1-49.

Artis, M.J. (1997). How accurate are the WEO's short-term forecasts? An examination of the *World Economic Outlook*, *Staff Studies for the World Economic Outlook*, (International Monetary Fund).

Ash, J.C.K., Smyth, D.J. and Heravi, S.M. (1990). The accuracy of OECD Forecasts of the international Economy: Demand, output, and prices. *International Journal of Forecasting*, 6, 379-392.

Ash, J.C.K., Smyth, D.J. and Heravi, S.M. (1998), Are OECD forecasts rational and useful? A directional analysis, *International Journal of Forecasting*, 14, 381-391.

---

[6] As a caveat, our conclusion that the average of the private growth predictions made with the same forecasting leads dominated both organizations' very short-term growth forecasts should be interpreted cautiously. This conclusion would not hold for all private predictions. One set of private forecasts was even inferior to the naïve model.

Ashiya, M. (2006). Testing the rationality of forecast revisions made by the IMF and OECD, Journal of Forecasting, 25. 175-186.

Barrionuevo, J.M. (1993). How accurate are the World Economic Outlook projections? *Staff Studies for the World Economic Outlook*, (International Monetary Fund).

Batchelor, R. (2001). How useful are the forecasts of international agencies? The IMF and OECD versus the consensus, *Applied Economics*, 33, 225-235.

Fildes, R. and Stekler, H. (2002). The state of macroeconomic forecasting, *Journal of Macroeconomics,* 24, 435-468.

Granger, C.W.J. and Ramanathan, R. (1984), "Improved Methods of Combining Forecasts", *Journal of Forecasting*, 3, 197-204.

Holden, K. and Peel. D. A. (1990). On testing for unbiasedness and efficiency of forecasts, *Manchester School*, 48, 120–127.

Holden, K., Peel. D. A. and Sandu, B. (1987). The accuracy of OECD forecasts, *Empirical Economics*, 12, 175-186.

Isiklar, G. and Lahiri, K. (2007). How far ahead can we forecast? Evidence from cross-country surveys, *International Journal of Forecasting*, 23, 167-187.

Kreinin, M. E. (2000). Accuracy of OECD and IMF projection, *Journal of Policy Modeling*, 22, 61-79.

Loungani, P. (2001). How accurate are private sector forecasts? Cross-Country evidence from consensus forecasts output growth, *International Journal of Forecasting*, 17, 419-432.

Mincer, J. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts in J. Mincer, ed. *Economic Forecasts and Expectations*, (National Bureau of Economic Research, New York.)

Oller L-E. and Barot, B. (2000). The accuracy of European growth and inflation forecasts, *International Journal of Forecasting*, 16, 293-315.

Pons, J. (2000). The accuracy of IMF and OECD forecasts for G7 countries, *Journal of Forecasting,* 19, 53-63.

Timmermann, A. (2007). An evaluation of the *World Economic Outlook* forecasts. *IMF Staff Papers*, 1-33.

Table1a	Descriptive Statistics - GDP															
	OECD_1	Consensus_1	WEO_4	Consensus_4	OECD_7	Consensus_7	WEO_9	Consensus_9	OECD_13	Consensus_13	WEO_16	Consensus_16	OECD_19	Consensus_19	WEO_21	Consensus_21
Obs	12	10	12	10	12	10	12	10	11	10	11	10	11	10	11	10
Period	1999-2010	1999-2008	1999-2010	1999-2008	1999-2010	1999-2008	1999-2010	1999-2008	2000-2010	2000-2009	2000-2010	2000-2009	2000-2010	2000-2009	2000-2010	2000-2009
Mean (of errors)	0.1152	0.0896	0.3569	0.2118	0.4152	0.5228	0.9708	0.7158	0.7711	1.0667	1.1731	1.0677	0.7620	1.1436	1.3646	1.2128
Median (of errors)	0.0500	0.0324	0.5000	0.2176	0.8000	0.6955	1.0000	0.9531	1.1000	1.3584	1.4000	1.0790	1.2000	1.0660	1.7000	1.1197
Std. Dev.	0.3405	0.2572	0.5983	0.5306	0.7532	0.8519	0.6725	0.8840	1.0427	1.3211	1.1851	1.5084	1.1949	1.4537	1.1846	1.4972
MAE	0.2573	0.2046	0.5482	0.4340	0.7573	0.8481	1.0083	0.9782	1.1262	1.4795	1.4400	1.5170	1.2562	1.4064	1.5790	1.4455
RMSE	0.3457	0.2600	0.6749	0.5461	0.8321	0.9626	1.1649	1.1025	1.2582	1.6458	1.6288	1.7854	1.3707	1.7916	1.7714	1.8678
U1	0.0189	0.0144	0.0374	0.0304	0.0464	0.0545	0.0669	0.0632	0.0701	0.0928	0.1000	0.1007	0.0765	0.1016	0.1099	0.1063
U2	0.3383	0.2601	0.6604	0.5465	0.8142	0.9632	1.1399	1.1033	1.2187	0.9755	1.5777	1.0751	1.3277	1.2767	1.7159	1.3532

Table 1b   Descriptive Statistics - Inflation

	OECD_1	WEO_4	OECD_7	WEO_9	OECD_13	WEO_16	OECD_19	WEO_21
Obs	12	12	12	12	11	11	11	11
Period	1999-2010	1999-2010	1999-2010	1999-2010	2000-2010	2000-2010	2000-2010	2000-2010
Mean (of errors)	0.1700	-0.2300	-0.2500	-0.1800	0.3300	-0.2780	-0.0800	-0.0240
Median (of errors)	0.1130	-0.1490	-0.0500	-0.4000	1.3520	-0.5000	-0.2000	-0.6900
Std. Dev.	1.1734	0.4349	1.5631	1.1606	2.4086	2.4824	2.7915	2.4863
MAE	1.1360	0.4750	1.5180	1.1260	2.3200	2.3830	2.6630	2.3710
RMSE	0.8040	0.3550	1.3010	0.8750	2.1290	2.0360	1.9920	2.0550
U1	0.1940	0.0830	0.2690	0.2120	0.4390	0.4350	0.4840	0.4670
U2	0.4104	0.1717	0.5483	0.4233	0.8024	0.8918	0.9208	0.8199

Table2   Probability that the Real GDP Forecasts were unbiased and/or efficient

Horizon	WEO (1999-2010)		OECD (1999-2010)		Consensus (1999-2008)	
	Holden-Peel	Mincer-Zarnowitz	Holden-Peel	Mincer-Zarnowitz	Holden-Peel	Mincer-Zarnowitz
1			0.2661	0.5137	0.2995	0.4458
4	0.0632	0.0746			0.2385	0.4972
7			0.0826	0.1582	0.0842	0.2373
9	0.0004018	0.002			0.0307	0.1117
13			0.0341	0.0398	0.031	0.0819
16	0.0057	0.0121			0.052	0.0814
19			0.0605	0.0809	0.0346	0.0784
21	0.0025	0.01			0.0306	0.0656

Table 3 Encompassing Results- Name Indicates the Organization that Encompasses

Table 3A

	OECD(1) vs. WEO(4)	OECD(7) vs. WEO(4)	OECD(7) vs. WEO(9)	OECD(13) vs. WEO(9)	OECD(13) vs. WEO(16)	OECD(19) vs. WEO(16)	OECD(19) vs. WEO(21)
Real GDP	OECD	WEO	OECD	?	OECD	OECD	OECD
Inflation	WEO	WEO	WEO	WEO	---	---	---

Table 3B

	OECD(1) vs. Consensus(1)	WEO(4) vs. Consensus(4)	OECD(7) vs. Consensus(7)	WEO(9) vs. Consensus(9)	OECD(13) vs. Consensus(13)	WEO(16) vs. Consensus(16)	OECD(19) vs. Consensus(19)	WEO(21) vs. Consensus(21)
Real GDP	Consensus	Consensus	----	?	?	?	OECD	?

Note: ---- indicates both forecasts encompass each other. ? Indicates neither forecast encompasses the other.