

June 2014

A comparative study of the forecasting performance of three international organizations

Pingfan Hong
*Zhibo Tan**

ABSTRACT

This article evaluates and compares the forecasting performance of three international organizations: the United Nations, the International Monetary Fund and the World Bank. The annual forecasts made by the United Nations in the period of 1981-2011 are found to be fairly robust, in terms of bias and efficiency. In comparison, the forecasting performance of the United Nations is found to be marginally better than the other two organizations during the period of 2000-2012. However, the forecasts of all these organizations missed the Great Recession of 2009 by a large margin.

JEL Classification: C30; C80

Keywords: evaluation of forecasts; forecasting errors; macroeconomic forecasting; financial crisis

* Pingfan Hong is the Chief of Global Economic Monitoring Unit in the United Nations Department of Economic and Social Affairs, hong@un.org. Zhibo Tan is a PhD student in Peking University and an intern at the United Nations, tzb0905@pku.edu.cn. We would like to thank Hung-yi Li for his comments. Views expressed here are those of the authors and they do not represent the views of the United Nations.

CONTENTS

I	Introduction	1
II	Technical setting for evaluating and comparing forecasts	2
III	Evaluation of the UN forecasts for the period of 1981–2012	3
IV	Comparing the forecasting performances of these organizations.	5
V	Discussions on the large forecasting errors of the Great Depression	11
VI	Conclusion	12
	References	13

UN/DESA Working Papers are preliminary documents circulated in a limited number of copies and posted on the DESA website at <http://www.un.org/en/development/desa/papers/> to stimulate discussion and critical comment. The views and opinions expressed herein are those of the author and do not necessarily reflect those of the United Nations Secretariat. The designations and terminology employed may not conform to United Nations practice and do not imply the expression of any opinion whatsoever on the part of the Organization.

Typesetter: *Nancy Settecasì*

UNITED NATIONS
Department of Economic and Social Affairs
UN Secretariat, 405 East 42nd Street
New York, N.Y. 10017, USA
e-mail: undesa@un.org
<http://www.un.org/en/development/desa/papers/>

I Introduction

The United Nations (UN), the International Monetary Fund (IMF) and the World Bank (WB) all make annual forecasts for the world economy in their publications of the *World Economic Situation and Prospects (WESP)* (UN, 2013), the *World Economic Outlook (WEO)* (IMF, 2013a), and the *Global Economic Prospects (GEP)* (World Bank, 2013a), respectively. While each of these publications would fulfill a specific institutional mandate within these organizations individually, together these forecasts also provide critical information for the general public, significantly influencing decision-making of the governments and businesses worldwide. This paper is intended to study the forecasting performance of these organizations.

A common feature among these international organizations is that they all make forecasts for the world economy based on the forecasts for individual countries, although they differ considerably in their approaches to making these forecasts.

Since the early 1970s, the UN forecasts for the world economy have been based on the cooperation with Project LINK, which is an international consortium for economic forecasting and policy analysis. Project LINK was initiated under the leadership of Lawrence Klein, consisting of about 100 experts worldwide.¹ The LINK modelling system for the world economy has evolved over the past decades. The original LINK country models were quite diverse, but since 2005, the models have been streamlined through a new system of the World Economic Forecasting Modelling (WEFM) (Altshuler et al., 2011).

The WEFM has maintained the bottom-up modelling approach and the international linkage mechanism of the original LINK system, but it has also made significant improvement on the original LINK system in several aspects. For example, in the WEFM, individual country models are developed by

following one of three different prototype models. Models for most developed countries and major developing countries follow a prototype model of full scale, with some 300 variables, and models for other developing countries follow a prototype model of much smaller size, with some 50 variables, of which there are nonoil and oil variants.

In the WEFM country models, behavioural equations are specified in a cointegration/error-correction framework, with the long-run relationship specified to follow a neo-classic theoretical framework while the short-run relationship constructed to follow a Keynesian framework and fit the data. The error correction mechanism ensures that the system moves towards the long-run path in the absence of shocks. Policy variables are modelled to follow rules according to country-specific situations, with flexible options for discretionary policy actions whenever necessary. As such, the same modelling system can be used for both policy analysis and forecasting.

The IMF forecasts draw primarily on the information the staff gathers through their consultations with member countries (IMF, 2013b). The forecasts are prepared by the country desk economists on the basis of internationally consistent assumptions about world economic activity, exchange rates, conditions in international financial and commodity markets, as well as the evolving situation in member countries. The methodology for individual country forecasts varies from country to country and series to series depending on many factors.

For approximately 50 of the largest economies, which account for 90 percent of world output, the IMF projections are updated for each WEO exercise. For other countries, the forecasts are updated and revised both during the WEO exercise and at the time of the IMF's regular Article IV consultations with member countries or in connection with the use of Fund resources, unless world economic developments necessitate more frequent updates. Composite data for country groups in the WEO are either sums or weighted averages of data for individual countries. All historical and projected data are recalculated to

¹ More information about Project LINK can be found at http://www.un.org/en/development/desa/policy/proj_link/index.shtml

reflect changes in country group composition in order to ensure that the economic indicators reflect changes in economic conditions and not changes in the country group composition.

The WB forecasts in the GEP are prepared by staff of the Development Prospects Group of the World Bank (World Bank, 2013b). The principal objective of the forecasting process is to illuminate the key forces acting in the global economy and its implications for developing countries, at present and over the medium term. Reflecting this objective, the main focus of attention is upon regional developments and impacts on groups of economically similar low- and middle-income countries.

In the WB, a large volume of high-frequency data, concentrating upon industrial output, inflation, trade, and international finance data for some 100 developing countries is used to help evaluate the international economic climate. Regional forecasts are constructed on the basis of 150 separate country-specific forecasts, prepared by the staff in conjunction with WB country experts. The country forecasts are aggregated and revised in a coherent model of the global economy.

The remainder of the paper is organized as follows. Section II defines a few technical terms to be used in evaluating the forecasting performance through the rest of the paper. Section III evaluates the forecasting performance of the UN during the periods of 1981-2012. Section IV compares the forecasting performance of the UN, IMF and WB at different aggregate and individual country levels. Section V provides some discussions of forecasting errors particularly for the period of the global financial crisis and the Great Recession, before the concluding section.

II Technical setting for evaluating and comparing forecasts

A number of papers have in the past made assessments of the forecasting performance of international

organizations. For example, Artis (1996) analysed the IMF short-term forecasts for developed and developing countries, showing that the IMF has over- or under-predicted the growth of advanced countries by about 1 percentage point, with the forecasting errors even larger for developing countries. In the same paper, an analysis of pre-1983 and post-1983 period shows that there is no significant difference in the accuracy between IMF forecasts and forecasts based on random walk models. The paper concluded that predicting the turning points for business cycles is the main weakness of the IMF forecasts.

Kreinin (2000) also finds that neither the Organization of Economic Cooperation and Development (OECD) nor the IMF succeeds in forecasting cyclical turning points; however, their projections appear fairly robust and certainly superior to those of a “naive” model. Pons (2000) investigates the size and nature of the IMF’s forecasting errors in GDP for G7 countries from 1971 to 1995, pointing out that the IMF forecasting accuracy has not improved over time.

Musso and Phillips (2002) are positive about the IMF forecasts, as they did not find a statistically significant biasness or inefficiency in the forecasting errors of the IMF. Also in contrast with Artis (1996) and Pons (2000), Záborský (2004), based on his evaluation of the forecasting performance of the IMF and the OECD during the period of 1973-2001, finds that the IMF forecast accuracy has improved since 1987, and the accuracy is better than the forecasts of naive models by about 30 percent.

Our paper differs from those earlier studies, as we focus on comparing the forecasting performance among the UN, IMF and WB. We compare the forecasts of these organizations not only at the global and regional aggregate levels, but also at the individual country level. We also give a special attention to the forecasting performance in the period of the global financial crisis.

The forecasts of these three organizations cover a large number of macroeconomic variables, but in our evaluation and comparison, we focus only on

one variable: GDP growth rate at the global, regional and individual country levels.

We define a number of technical terms, as follows, for evaluating the forecasting errors, which will be used through the rest of the paper.

f_t – forecast of GDP growth (%) for year t , as projected in $t-1$;²

g_t – GDP growth for year t , as officially reported in year $t+2$;³

$e_t = (100 + f_t) / (100 + g_t) - 1$ * 100, forecasting error.

Root mean square errors (RMSE):

$$RMSE = \sqrt{\sum_{t=1}^T e_t^2 / T}$$

Mean absolute errors (MAE):

$$MAE = \sum_{t=1}^T |e_t| / T$$

Mean absolute percentage errors (MAPE):

$$MAPE = \sum_{t=1}^T |(f_t - g_t) / g_t| / T$$

Another technical note is that the three international organizations have different approaches to calculating the aggregate GDP at the global and regional levels. For example, the IMF uses purchasing power parity (PPP) to calculate the aggregate GDP for the world, while the UN and the WB use official exchange rates, and as a result, the growth for the world

economy in the IMF forecasts is usually higher than the UN and the WB, because the PPP approach gives more weight to emerging economies, such as China and India, which have much higher growth than developed countries. The country grouping definition also varies: for example, although the memberships in the UN definition of developed countries, the IMF definition of advanced countries, and the WB definition of high-income countries are very close to each other, they are not identical. In order to avoid the difference in the forecasting errors among the three organizations caused by the differences in aggregating and country-grouping, when comparing the forecasting errors of these organizations, the errors are calculated according to their own respective aggregating and grouping definitions.

III Evaluation of the UN forecasts for the period of 1981-2012

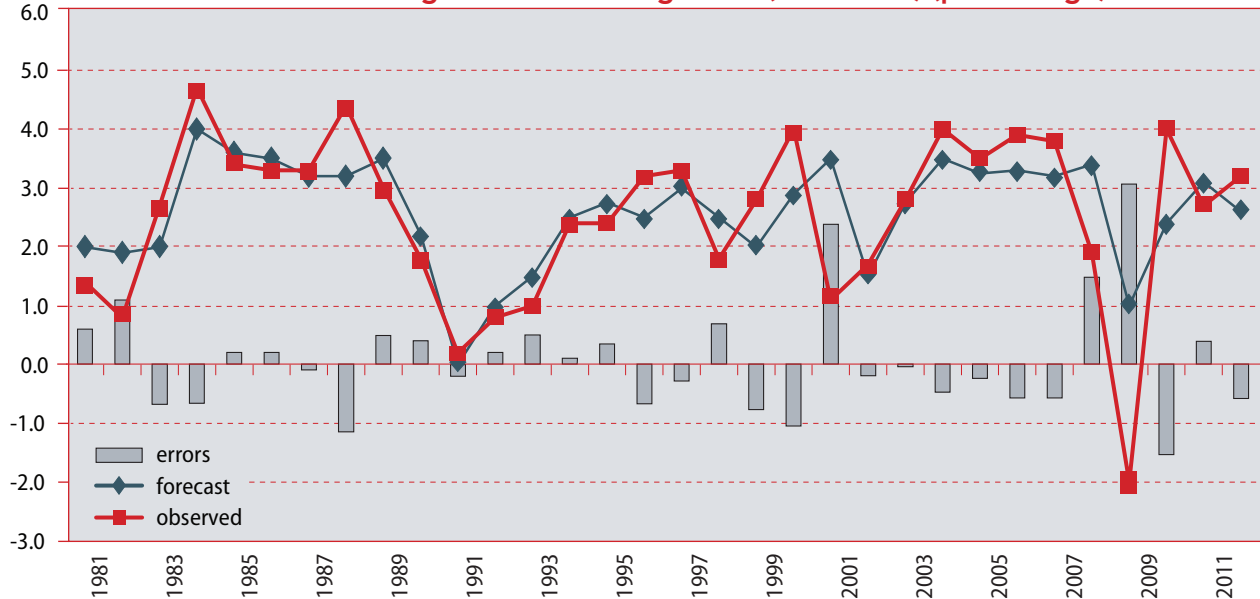
As shown in figure 1, the forecasting errors of the UN annual projection of the global economic growth rate are less than 1 percentage point for most years in the period of 1981 to 2012. However, the forecasting errors for the years of 1982, 2000, 2001, and 2008-2010 are egregiously large, as these years are all associated with the financial crises, namely, the Latin American debt crisis of the early 1980s, the burst of the high-tech bubble of the early 2000s, and the global financial crisis of 2008, as well as the Great Recession of 2009.

Table 1 provides the summary statistics of forecasting errors for the UN annual forecasts of the growth rate for the world economy, developed countries and developing countries respectively. The table shows that the mean value and the standard deviations of the forecasting errors are larger for the group of developing countries than for developed countries. The same is true as measured by other indicators of forecasting performances, such as the RMSE, MAE and MAPE, as well as the maximum and minimum values of the errors. One interesting point is that, by all these measures, the forecasting errors for the

² More precisely, we select the annual forecasts in the IMF *WEO* released in October, the UN *WESP* released in December and the WB *GEP* of December, respectively.

³ For most countries, the national statistic authorities revise their annual GDP statistics for the year of t three times within the next two years. However, many countries would also systematically revise their GDP statistics for a long period of time, once in a decade or so. For our purpose, we only take the official GDP statistics released within the period of $t+2$ as the final observation.

Figure 1
The UN forecasts for the global economic growth (1981-2012) (percentage)



world economy are smaller than the forecasting errors for either the group of developed countries or the group of developing countries, because the forecasting errors at the lower levels are offset each other when they are aggregated to the global level.

The statistics also indicate that the mean and median values of the forecasting errors are not significantly different from zero, suggesting that there is no systematic bias in the forecasts. Moreover, no significant serial correlation is found among the forecasting errors either, attesting that available information has been fully utilized and the forecasting is efficient.

Additionally, in comparing the UN forecasts with the forecasts produced by a naive random-walk model, no significant difference is found in the mean and median values of the forecasting errors between the UN forecasts and the random-walk forecasts, but the standard deviation of the forecasting errors of the UN is significantly smaller than that of the random-walk model, suggesting that the UN forecasting is superior to the random-walk forecasting.

A few “outliers”, namely, the years when the forecasting errors are exceptionally large, can be identified, using “two standard deviations” as the benchmark. At the global level, the forecasting errors in 2001 and 2009 are exceptionally large, over-predicting the growth of the world economy by 2.4 and 3.1 percentage points respectively. For the group of developed countries, exceptionally large forecasting errors are also found in 2001 and 2009, over-predicting the growth of developed countries by 2.3 and 3.1 percentage points respectively. For the group of developing countries, outliers are found in 1981 and 2001, when the growth of this group was over-predicted by 3.3 and 3.4 percentage points respectively, but the forecasting errors are also large for 1982, over-predicted by 2.5 percentage points; 1998, over-predicted by 2.4 percentage points; 2009, over-predicted by 2.2 percentage points; and 2010, under-predicted by 2.0 percentage points.

Obviously, without exceptions, all the egregiously large forecasting errors have occurred for the years when a virulent financial crisis wreaked havoc on the world economy, or a large number of economies.

IV Comparing the forecasting performance of these organizations

After evaluating the forecasting performance of the UN, we can adopt the same measures to evaluate the forecasts made by the IMF and WB, and compare the forecasting performance among these three organizations. We focus on the period of 2000-2012.

We start with comparing the forecasting errors of these three organizations at the global level. As shown in figure 2, the forecasting errors of these

three organizations for the growth of the world economy are very similar in terms of magnitude and patterns. In fact, these forecasting errors are highly correlated, with the correlation coefficient of 0.93 for UN and IMF, 0.92 for UN and WB, and 0.98 for IMF and WB, respectively.

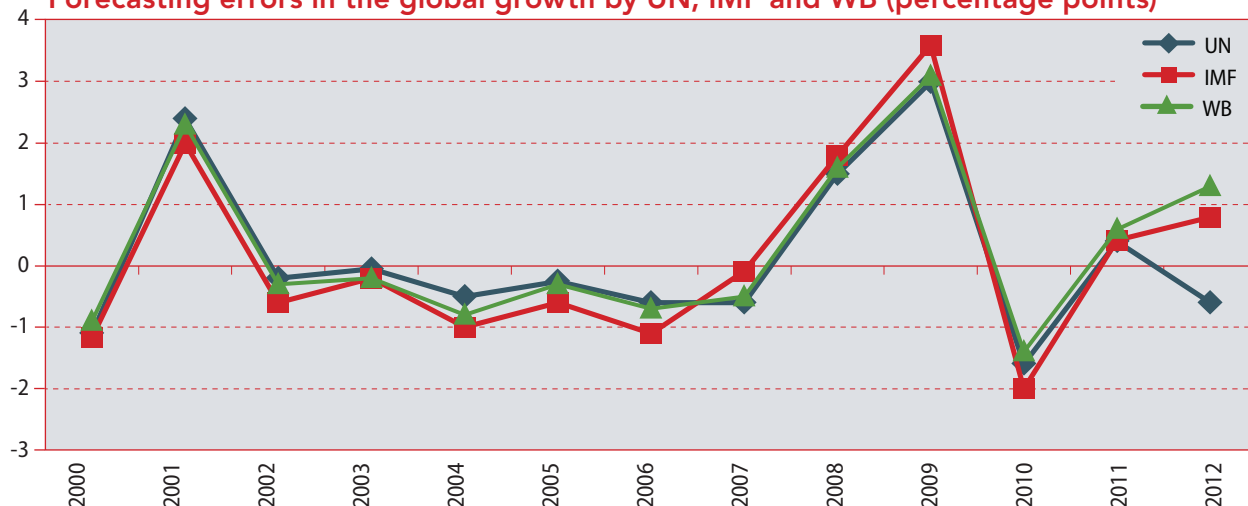
One exception is for the year of 2012, when the UN forecast deviated from the other two. The UN under-predicted the global growth in 2012 while the other two over-predicted. Meanwhile, they all made significantly large forecasting errors in 2001 and 2009.

Table 1

Summary statistics for the UN forecasting errors

	World		Developed countries		Developing countries	
	Model	Random walk	Model	Random walk	Model	Random walk
Mean	0.06	-0.05	0.08	-0.01	0.32	-0.14
p value	0.72	0.86	0.68	0.98	0.21	0.69
Median	-0.1	-0.1	-0.1	-0.1	0.1	-0.1
p value	0.74	0.96	0.9	0.82	0.39	0.5
Maximum	3	3.9	3	4	3.5	4.4
Minimum	-1.6	-6	-1.9	-6.2	-2.2	-5.1
Standard deviations	0.97	1.76	1.01	1.85	1.41	1.89
Skewness	1.15	-0.74	0.91	-0.87	0.56	0.2
Kurtosis	4.77	5.91	4.35	5.65	2.76	4.02
Jarque-Bera	11.26	13.80	6.81	12.98	1.72	1.56
Probability	0.00	0.00	0.03	0.00	0.42	0.46
Number of positive errors	15	13	15	15	17	13
Fraction of positive errors	0.47	0.42	0.47	0.47	0.53	0.42
Serial correlation	-0.16	-0.29	-0.07	-0.33	0.22	-0.34
p value	0.35	0.09	0.67	0.06	0.19	0.05
Root of mean square errors	0.94	1.69	0.99	1.8	1.39	1.79
Mean absolute errors	0.68	1.21	0.69	1.3	1.02	1.29
Mean absolute percentage errors	37.72	77.39	67.17	99.95	85.53	48.45
Mean equal test (model vs random walk)		p=0.63		p=0.73		p=0.13
Median equal test (model vs random walk)		p=1.00		p=0.93		p=0.58
Variance equal test (model vs random walk)		p=0.00		p=0.00		p=0.12

Figure 2
Forecasting errors in the global growth by UN, IMF and WB (percentage points)



As shown in table 2, the mean and median values of forecasting errors for these three organizations are all insignificantly different from zero, suggesting that their forecasts are unbiased, although the mean values of forecasting errors for the WB are relatively larger than that of the UN and the IMF. No serial correlations are found for the forecasting errors of these organizations, meaning that their forecasts are efficient. By the measures of RMSE, MAE and MAPE, the forecasts of these organizations are superior to the forecasts generated by a simple random-walk model. Also by the measures of RMSE, MAE and MAPE, the forecasts of the UN are marginally better than the other two.

We can also apply the same approach to comparing the forecasting errors for these three organizations in their forecasts for the group of developed countries and the group of developing countries respectively. Given the limited space, we cannot present the figures and statistic tables here, but they can be requested by the readers who are interested in them. The findings for the forecasting errors for these two country-groups are similar to those for the global economy as discussed above. In general, the forecasts are unbiased, efficient, and superior to random-walk models. Again, by the measures of RMSE, MAE and MAPE, the forecasts of the UN are marginally better than the other two organizations.

As the forecasts for the global economic growth and the growth of the country groups are the aggregate of the forecasts for the growth of individual countries, it is more interesting to compare the forecasting performance of these organizations in their forecasts for individual countries.

For simplicity, we only present in this paper the comparative statistics for the UN and the IMF in their forecasts for ten major individual countries, including five developed countries, namely, France, Germany, Japan, the United Kingdom and the United States, and five emerging economies, namely, Brazil, China, India, Russia and South Africa.

As shown in figure 3, which presents the forecasting errors of the UN and the IMF respectively for ten selected countries, among emerging economies, for example, for Brazil, both the UN and the IMF significantly (with errors larger than 2 percentage points) over-predicted the GDP growth for 2001 and 2009, while under-projected for 2004, 2007 and 2010. The forecasting errors of the IMF are larger than that of the UN for 2003 and 2012, but smaller for 2000, 2005, 2006 and 2007. For China, both the UN and the IMF under-predicted its GDP growth for a few years before the global financial crisis of 2008, but over-predicted for 2012, with the overall forecasting errors of the UN marginally smaller

Table 2

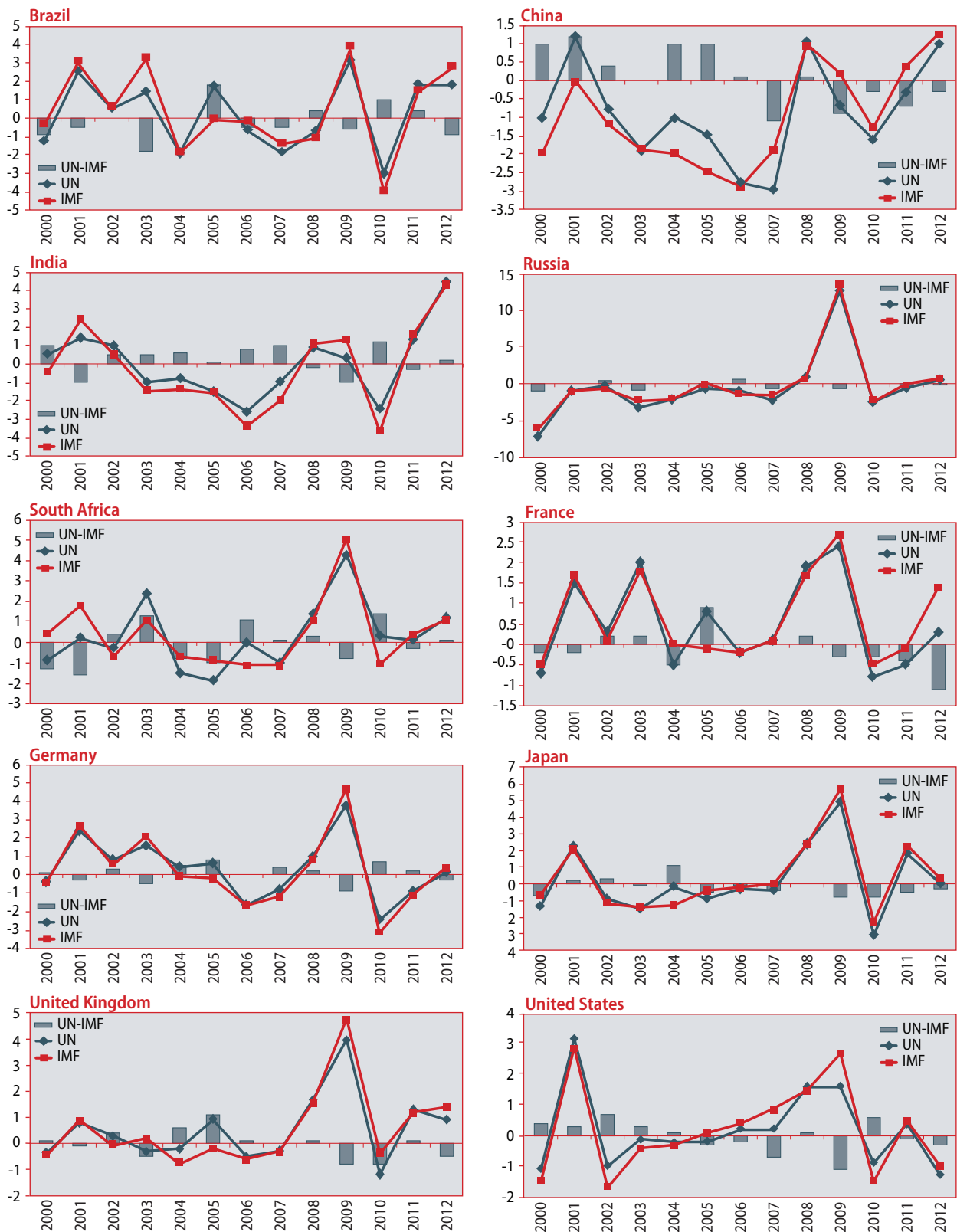
Summary statistics for the forecasting errors of UN, IMF and WB

	UN		IMF		WB	
	Model	Random walk	Model	Random walk	Model	Random walk
Mean	0.14	-0.02	0.14	0.02	0.29	0.04
p value	0.72	0.97	0.75	0.98	0.46	0.96
Median	-0.3	-0.4	-0.2	0.2	-0.3	0.2
p value	0.65	0.92	0.89	0.83	0.75	0.83
Maximum	3	3.9	3.6	3.6	3.1	3.9
Minimum	-1.6	-6.0	-2.0	-5.7	-1.4	-6.3
Standard deviations	1.36	2.41	1.56	2.29	1.38	2.45
Skewness	1	-0.76	0.83	-0.87	0.79	-1.01
Kurtosis	2.9	4.27	2.94	4.27	2.4	4.78
Jarque-Bera	2.16	2.13	1.51	2.53	1.53	3.94
Probability	0.34	0.34	0.47	0.28	0.47	0.14
Number of positive errors	4	6	5	7	5	7
Fraction of positive errors	0.31	0.46	0.38	0.54	0.38	0.54
serial correlation	-0.23	-0.4	-0.12	-0.37	-0.13	-0.4
p value	0.36	0.1	0.62	0.14	0.59	0.11
root of mean square errors	1.31	2.32	1.50	2.20	1.36	2.35
mean absolute errors	0.98	1.65	1.18	1.62	1.08	1.64
mean absolute percentage errors	47.12	71.62	71.58	82.10	51.15	70.35
Mean equal test		p=0.63		p=0.73		p=0.13
(model vs random walk)		p=0.58		p=0.71		p=0.55
Median equal test		p=0.00		p=0.00		p=0.12
(model vs random walk)		p=0.69		p=1.00		p=1.00
Variance equal test		p=0.06		p=0.20		p=0.06
(model vs random walk)		p=0.06		p=0.20		p=0.06
Statistics equal test						
(UN vs IMF)	Mean	p=1.00	Median	p=0.39	Variance	p=0.64
Statistics equal test						
(UN vs WB)	Mean	p=0.33	Median	p=1.00	Variance	p=0.96
Statistics equal test						
(IMF vs WB)	Mean	p=0.12	Median	p=0.15	Variance	p=0.67

than that of the IMF. The largest forecasting errors for both the UN and the IMF in their forecasts for individual countries are found for Russian Federation for 2009, over-predicting Russian GDP growth by more than 13 percentage points, as the Russian economy fell into a deep recession amid the global Great Recession.

Among developed countries, both the UN and the IMF registered egregiously large forecasting errors, at the size of about 5 percentage points, in their forecast for GDP growth of all developed countries for 2009, when these economies fell into the Great Recession. The forecasting errors of the UN for the United States in the period of 2007-2010 are

Figure 3
Forecasting errors in country growth by UN, IMF (percentage points)



relatively smaller than those of the IMF. No significant differences are found in the forecasting errors for other four developed economies between the UN and the IMF.

As shown in table 3, which summarizes the statistics of the forecasting errors for both the UN and the IMF in their forecasts of GDP growth for ten selected individual countries, among emerging economies, the measures of RMSE, MAE and MAPE for the forecasting errors in both the UN and the IMF forecasts for China and India are actually larger than those of random-walk models, indicating a significant room for improvement in the forecasting methods for these two organizations.

Among developed countries, the standard deviations, RMSE, MAE and MAPE of the IMF are the smallest for France. For the ten selected countries, the maximum forecasting errors of the IMF are relatively larger than those of the UN, except for India and the United States. The IMF forecasts tend to over-predict the GDP growth for these ten countries more than the UN forecasts.

The average mean values of the forecasting errors for the five emerging economies are -0.16 for the UN and -0.12 for the IMF, compared with the same terms for the five developed countries of 0.36 and 0.41. The average values of standard deviations in the forecasting errors for the five developing countries are 2.28 for

Table 3

Summary statistics of the forecasting errors for selected countries

In the UN forecasts										
	Brazil	China	France	Germany	India	Japan	Russia	South Africa	UK	USA
Mean	0.25	-0.87	0.51	0.34	0.04	0.21	-0.54	0.33	0.53	0.18
p value	0.65	0.04	0.12	0.48	0.94	0.73	0.67	0.49	0.17	0.62
Median	0.5	-1.0	0.3	0.4	0.3	-0.3	-0.9	0.1	0.3	-0.1
p value	0.7	0.07	0.23	0.55	0.86	0.89	0.08	0.65	0.22	0.81
Maximum	3.1	1.2	2.4	3.8	4.5	4.9	12.7	4.3	4.0	3.2
Minimum	-3.0	-3.0	-0.8	-2.5	-2.6	-3.1	-7.3	-1.9	-1.2	-1.3
Standard deviations	1.95	1.36	1.11	1.69	1.91	2.11	4.49	1.69	1.33	1.29
Skewness	-0.16	0.16	0.46	0.32	0.67	0.72	1.89	0.93	1.28	1
Kurtosis	1.7	2.14	1.78	2.79	3.38	3.05	7.31	3.46	4.56	3.32
Jarque-Bera	0.96	0.46	1.27	0.24	1.04	1.14	17.83	2	4.88	2.21
Probability	0.62	0.8	0.53	0.89	0.59	0.57	0	0.37	0.09	0.33
No. of positive errors	7	3	8	8	7	4	3	7	7	6
Fraction of positive errors	0.54	0.23	0.62	0.62	0.54	0.31	0.23	0.54	0.54	0.46
Serial correlation	-0.44	0.09	-0.16	-0.08	0.25	-0.26	-0.01	0.09	-0.15	-0.35
p value	0.07	0.72	0.53	0.76	0.31	0.29	0.98	0.73	0.53	0.16
RMSE	1.89	1.57	1.18	1.65	1.84	2.04	4.35	1.65	1.39	1.25
MAE	1.72	1.38	0.92	1.31	1.48	1.55	2.71	1.19	0.98	0.92
MAPE	192.16	14.30	106.65	219.93	26.52	149.39	39.28	46.41	92.20	132.92

(cont'd)

Table 3 (cont'd)

In the IMF forecasts										
	Brazil	China	France	Germany	India	Japan	Russia	South Africa	UK	USA
Mean	0.42	-0.98	0.62	0.25	-0.22	0.41	-0.24	0.42	0.55	0.2
p value	0.53	0.03	0.06	0.67	0.74	0.51	0.85	0.41	0.21	0.64
Median	-0.1	-1.3	0.1	-0.1	-0.5	-0.2	-1.0	0.4	-0.1	0.1
p value	0.62	0.04	0.23	0.89	0.65	0.86	0.13	0.57	0.40	0.78
Maximum	3.7	1.3	2.7	4.7	4.3	5.7	13.4	5.1	4.8	2.9
Minimum	-4.0	-2.9	-0.5	-3.2	-3.7	-2.3	-6.3	-1.1	-0.8	-1.7
Standard deviations	2.29	1.39	1.07	2.04	2.34	2.18	4.50	1.74	1.51	1.51
Skewness	-0.16	0.35	0.62	0.56	0.25	1.1	2.17	1.49	1.79	0.48
Kurtosis	2.2	1.72	1.92	3.13	2.27	3.62	7.93	4.99	5.8	2.22
Jarque-Bera	0.4	1.16	1.48	0.69	0.42	2.81	23.4	6.98	11.17	0.84
Probability	0.82	0.56	0.48	0.71	0.81	0.25	0	0.03	0	0.66
No. of positive errors	6	4	7	6	6	5	4	7	6	7
Fraction of positive errors	0.46	0.31	0.54	0.46	0.46	0.38	0.31	0.54	0.46	0.54
Serial correlation	-0.38	0.38	-0.15	-0.14	0.15	-0.11	-0.03	-0.06	0.09	-0.33
p value	0.12	0.12	0.55	0.58	0.55	0.65	0.92	0.8	0.72	0.18
RMSE	2.24	1.66	1.20	1.98	2.26	2.13	4.33	1.72	1.55	1.46
MAE	1.82	1.43	0.84	1.48	1.95	1.56	2.56	1.28	1.00	1.18
MAPE	333.71	15.07	96.81	251.06	32.65	174.78	37.49	52.05	101.9	130.05

Table 4

Summary statistics of pool forecasting errors for ten selected countries

	UN	IMF	
Mean	0.10	0.14	Mean equal test
p value	0.81	0.77	p=0.65
Median	-0.07	-0.40	Median equal test
p value	0.86	0.97	p=1.00
Maximum	3.64	4.43	Variance equal test
Minimum	-1.78	-2.05	p=0.60
Standard deviations	1.46	1.70	
Skewness	0.99	1.16	
Kurtosis	3.65	4.08	
Jarque-Bera	2.34	3.56	
Probability	0.31	0.17	
Number of positive errors	5.00	6.00	
Fraction of positive errors	0.38	0.46	
Serial correlation	-0.17	-0.19	
p value	0.48	0.45	
Root of mean square error	1.41	1.64	
Mean absolute error	1.08	1.27	

the UN and 2.45 for the IMF, compared with that for the five developed countries of 1.51 and 1.66.

As shown in table 4, which compares the statistics for the forecasting errors of the UN and IMF in their forecasts for the ten selected countries as a pool, the RMSE for the UN is 1.41 and for IMF is 1.64. By other measures, such as the standard deviations, MAE, and maximum and minimum forecasting errors, the UN forecasting performance is also moderately better than the IMF. Moreover, the null hypothesis about a normal distribution of forecasting errors cannot be rejected for the UN, but can be rejected for the IMF at the 0.20 significance level. In short, the UN forecasts for the selected individual countries are marginally better than the IMF forecasts, but the differences in the values of mean, median and variance for the forecasting errors of these two organizations are not statistically significant.

V Discussions on the large forecasting errors for the Great Recession

As mentioned above, one common egregiously large forecasting error shared by all the three international organizations is the failure to predict the Great Recession of 2009. By late 2008, all these organizations predicted a positive global growth for 2009, but as it turned out the world economy fell into the Great Recession in 2009, contracting 2 percent if measured by the aggregating approach of exchange rates, or 0.5 percent if measured by PPP.

In the hindsight, we may explain this failure by a number of factors: the increased complexity in the financial system, and the pervasive chicaneries and malodorous operations in high finance in the run up to the financial crisis; the unprecedentedly acute degree of the damages to financial intermediation when the crisis erupted; the extraordinarily large impact on the real economy; and the unparalleled international contagion and transmission of the financial shocks across the world economies.

To be fair, since 2006 and until 2008 when the financial crisis erupted, these international organizations had indeed increasingly issued warnings of the possible financial instability associated with such indicators as the rising house prices in developed countries and the widening global imbalances, for example, in the UN WESP 2006 (UN, 2006), and the IMF WEO 2006 (IMF, 2006).

However, the forecasters in these organizations could not grasp the whole picture of the ever growing complexity developed in the global financial system, featured by the rapid increases in the sub-prime loans, securitization of the mortgages, the credit default swaps and many other complicated financial derivatives and toxic assets. For example, securitization of mortgage-backed securities in the United States surged by 5 times in five years before collapsed in 2007 (Greenspan, 2010). It is also impossible for these forecasters to gather the information about many malodorous practices and chicaneries in high finance places worldwide, which were only revealed after the eruption of the financial crisis. Therefore, it is infeasible for the forecasters to build these factors in their models and take these factors into full account when they made their forecasts before the advent of the global financial crisis.

When the financial crisis erupted in 2008, as marked by the collapse of the Lehman Brothers, the damages to financial markets, particularly to the function of financial intermediation were unprecedented. As commented by Greenspan (2010), the evaporation of the global supply of short-term credits within hours or days of the Lehman failure is without historical precedent. A run on money market mutual funds was under way within hours of the announcement of Lehman's default. Even the fully collateralized repurchase agreement market encountered severe and unprecedented difficulties. In late 2008 and early 2009, the spreads on interbank lending and the credit default swaps all surged to historical highs. The forecasters were unable to assess the extreme severity of the damages the crisis brought to the financial system.

Moreover, even after observing the eruption of the financial crisis in late 2008, these forecasters still failed to fully ascertain the extraordinary impact of the financial shock on the real economy, particularly on business investment and trade. For example, business investment in the United States dropped by more than 20 percent in 2009. Such a magnitude of decline in real investment can hardly be simulated through any investment functions that are estimated on the data in the period before this financial crisis.⁴

Furthermore, the forecasters have also underestimated the magnitude of international contagion and international transmission of the financial crisis from the United States to the rest of the world through trade and financial channels. In the late 2008 and early 2009, exports of many countries plummeted more than 50 percent, partly caused by the withdrawal of trade credit. For 2009 as a whole, the world total exports dropped by 12 percent (World Trade Organization, 2011), but by late 2008, the international organizations still forecast a positive growth of world exports. Capital inflows to emerging economies dropped by 50 percent from the peak of 2007 to the trough of 2009 (International Institute of Finance, 2014).

It was the accumulation of all these factors that have attributed to the large forecasting errors of these international organizations.

In addition to the baseline forecasts, these organizations also produce alternative scenarios, or give the interval of possible outcomes. For example, when the UN made the forecast in 2008 for global growth of 2009, its baseline forecast was 1 percent, but its pessimistic scenario was -0.5 percent (UN, 2009). The IMF baseline forecast was 3 percent for 2009, but its forecasting interval as indicated by a fan

chart showed the possibility of the lower bound at 1 percent (IMF, 2009). The baseline forecasts of both the UN and IMF missed the actual growth by a large margin, by 3 percentage points, but their lower bound forecasts were much closer, missing the actual growth by about 1.5 percentage points.

VI Conclusion

The annual economic forecasts made by the United Nations in the period of 1981-2011 are found to be fairly robust in general, in terms of systematical bias and efficiency. The mean and median of forecasting errors are not significantly different from zero for the world economy, the groups of developed countries and developing countries, and individual countries. The forecasting performance of the UN is superior to random-walk forecasting models.

In comparison, the forecasting performance of the UN is found to be marginally better than that of the IMF and the WB at the global and country-group levels, and also than the IMF for ten selected large individual countries during the period of 2000-2011. Unfortunately, the forecasts made by these three organizations are all found to be poor for the years of financial crises, with particularly significant errors for the Great Recession of 2009.

In the hindsight, the failure to predict the Great Recession was due to a number of factors, including the increased complexity in the financial system, and the pervasive chicaneries and malodorous operations in high finance in the run up to the financial crisis; the underestimate of the unprecedentedly acute degree of the damages to financial intermediation when the crisis erupted; the extraordinarily large impact on the real economy; and the unparalleled international contagion and transmission of the financial shocks across the world economies. Improvement can be made in the forecasting performance for the future if efforts are taken to incorporate these issues into the forecasting models.

⁴ A review of various investment functions specified in the LINK country models in the past can be found in Hong (1995).

References

- Altshuler, Clive, Pingfan Hong and Hung-yi Li. 2011. "A World Economic Forecasting Model of the United Nations". Unpublished.
- Artis, M. J., 1996. How accurate are the IMF's short-term forecasts? Another examination of the World Economic Outlook. *IMF Working Paper No. 96/89*.
- Greenspan, Allen. 2010. *The Crisis*. Brookings Institutions.
http://www.brookings.edu/-/media/Projects/BPEA/Spring%202010/2010a_bpea_greenspan.PDF
- Hong, Pingfan. 1995. "Determinants of Investment and Policy Implications – With a Comparative Study of Investment Functions in the LINK Models". Working Paper, Social Science Research Network.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1758944
- International Institute of Finance. 2014. *Capital Flows to Emerging Market Economies*. <http://www.iif.com/emr/capflows201401.php>
- IMF. 2006. *World Economic Outlook*, September 2006. <http://www.imf.org/external/pubs/ft/weo/2006/02/>
- IMF. 2008. *World Economic Outlook*, October 2008. <http://www.imf.org/external/pubs/ft/weo/2008/02/pdf/text.pdf>
- IMF. 2013a. *World Economic Outlook*, issues from 2000 to 2013. <http://www.imf.org/external/ns/cs.aspx?id=29>
- IMF. 2013b. "Frequently asked questions: World Economic Outlook." <https://www.imf.org/external/pubs/ft/weo/faq.htm>
- Kreinin, M. E., 2000. Accuracy of OECD and IMF projection. *Journal of Policy Modelling*, Vol.22, Issue 1.
- Musso, A., and Phillips S., 2002. Comparing projections and outcomes of IMF supported Programs. *IMF Staff Papers*, Vol. 49, No. 1.
- Pons, J., 2000. "The accuracy of IMF and OECD forecasts for G7 countries." *Journal of Forecasting*, Vol. 19, Issue 1.
- United Nations. 2006. *World Economic Situation and Prospects 2006*.
http://www.un.org/en/development/desa/policy/wesp/wesp_archive/2006wesp.pdf
- United Nations. 2009. *World Economic Situation and Prospects 2009*.
http://www.un.org/en/development/desa/policy/wesp/wesp_archive/2009wesp.pdf
- United Nations. 2013. *World Economic Situation and Prospects*. Issues from 2004 to 2013, and *World Economic and Social Survey*, from 1970 to 2004. <http://www.un.org/en/development/desa/policy/wesp/archive.shtml>
- World Bank. 2013a. *Global Economic Prospects*. Issues from 2000 to 2013. <http://www.worldbank.org/en/publication/global-economic-prospects>
- World Bank. 2013b. "About the Global Economic Prospects." <http://www.worldbank.org/en/publication/global-economic-prospects/about>
- World Trade Organization. 2011. *World Trade Report 2011*.
http://www.wto.org/english/res_e/booksp_e/anrep_e/world_trade_report11_e.pdf
- Zámborský, P., 2004. "Evaluation of IMF and OECD Output Growth Forecasts for G7 Countries." *Economic Outlook*. Issue March, 2004.