International Journal of Forecasting 32 (2016) 154–167

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

# How accurate are professional forecasts in Asia? Evidence from ten countries

Qiwei Chen [a], Mauro Costantini [a,*], Bruno Deschamps [b]

[a] Brunel University, Department of Economics and Finance, United Kingdom

[b] Nottingham University Business School China, China

ARTICLE INFO

*Keywords:*
Professional forecasts
Forecast efficiency
Forecast bias
Asia

ABSTRACT

This paper assesses the performances of professional GDP growth and inflation forecasts for ten Asian economies for the period 1995–2012. We evaluate the accuracy of the forecasts, and test for unbiasedness and efficiency. Our results show that (i) forecast errors are large for most of the countries, but there are big differences between countries; (ii) forecasts improve slowly as the forecast horizon shortens, which helps to explain the magnitudes of the forecast errors; (iii) GDP growth forecasts underreact to economic news but inflation forecasts are mostly efficient; (iv) the sizes of forecast biases vary widely between countries, with a tendency for inflation to be overestimated; and (v) forecasts have value in predicting the direction of change.



## 1. Introduction

There has been intensive study of the performances of professional macroeconomic forecasts. Using various data sets and methodologies, the empirical literature has analyzed the issues of forecast accuracy, unbiasedness and efficiency extensively, and shed light on the way in which forecasters form their expectations. One limitation of the literature is that it has focused mainly on large advanced countries, such as the US and other G-7 countries (see, e.g., Ager, Kappler, & Osterloh, 2009; Clements & Taylor, 2001; Dovern & Weisser, 2011; Isiklar, Lahiri, & Loungani, 2006). Only recently have some studies paid specific attention to emerging countries (e.g., Krkoska & Teksoz, 2009, for transition countries; Carvalho & Minella, 2012, for Brazil; and Capistrán & López-Moctezuma, 2014, for Mexico). However, little is known about the performances of professional macroeconomic forecasts in Asia, with the notable exception of a small number of studies that have focused on individual countries (see Ashiya, 2005, for Japan; Lahiri & Isiklar, 2009, for India; and Deschamps & Bianchi, 2012, for China).[1]

In this paper, we use the *Asian-Pacific Consensus Forecasts* to provide the first comprehensive evaluation of the macroeconomic forecasts produced for ten Asian economies, namely China, Hong Kong, India, Indonesia, Japan, Korea, Malaysia, Singapore, Taiwan, and Thailand. We assess the accuracy, unbiasedness and efficiency of forecasts of GDP growth and inflation, two key variables for macroeconomic analysis (see Costantini & Kunst, 2011; Golinelli & Parigi, 2008, 2014).

Several studies have found differences in forecast performances between advanced and emerging economies,

* Correspondence to: Brunel University, Department of Economics and Finance, London, Uxbridge, UB8 3PH, UK. Tel.: +44 0 1895 267958; fax: +44 1895 269770.

*E-mail address:* mauro.costantini@brunel.ac.uk (M. Costantini).

[1] Ashiya (2005) and Lahiri and Isiklar (2009) use different techniques from those used in this paper, and Deschamps and Bianchi (2012) do not assess the directional forecast accuracy.

http://dx.doi.org/10.1016/j.ijforecast.2015.05.004

especially in terms of accuracy, information rigidities and the efficient use of information (Dovern, Fritsche, Loungani, & Tamirisa, 2015; Loungani, 2001; Loungani, Stekler, & Tamirisa, 2013). After several decades of fast growth, some Asian economies have recently acquired the status of advanced economies, while some others, though still emerging, are growing rapidly. In this respect, it is worth investigating the performances of forecasts in these newly-advanced economies and comparing them with those observed in previous studies for advanced and emerging countries. In addition, it is also important to examine whether any progress has been made in improving forecast performances over the years, since the economies of many countries have transitioned from low/middle income to middle/high income.

Another important point with Asian economies is that they have all experienced large economic fluctuations: while recessions have tended to be more severe and longer-lasting than those in developed countries (Hong, Lee, & Tang, 2010), sharp economic recoveries have also occurred. Furthermore, Asia has made remarkable progress in fighting against inflation (Filardo & Genberg, 2010), and it is of interest to examine the performances of forecasters in such a volatile and fast-changing environment.

We analyze professional Asian macroeconomic forecasts over the period 1995–2012. The data set includes a large number of forecasters, and fixed-event forecasts are reported for horizons of up to 24 months. To evaluate the accuracy of the professional forecasts, we use the RMSE and a recent directional measure proposed by Blaskowitz and Herwartz (2009). While accuracy, as measured by quantitative errors, is important, predicting the direction of change of crucial variables correctly may also be important. This is the case for GDP growth and inflation, which are the most important macroeconomic goals for policy makers (a central bank can raise/lower the interest rate if inflation rises/falls, to stabilize the economy). To test for forecast unbiasedness and efficiency, we use the econometric approach that was initially developed by Davies and Lahiri (1995), and later extended by Clements, Joutz, and Stekler (2007), Ager et al. (2009) and Dovern and Weisser (2011). We choose to analyze individual forecasts rather than consensus forecasts in order to shed light on the individual heterogeneity across forecasters and avoid any problem of aggregation bias.

It should be noted that Loungani (2001), Loungani et al. (2013) and Dovern et al. (2015) use a larger data set that includes ours. However, our paper differs from theirs in several respects. First, they do not analyze inflation forecasts. Second, we focus on individual countries, whereas they pool across all countries (Asian and non-Asian).[2] Third, we analyze individual forecasts, whereas Loungani (2001) and Loungani et al. (2013) study consensus forecasts. Finally, we also address various other issues, such as directional accuracy, long-term predictability, and the acquisition of information.

[2] Using a different methodology, Dovern et al. (2015) only report results for individual countries in the case of efficiency.

Our analysis shows large forecast errors for both GDP growth and inflation series in most cases, with considerable differences in terms of accuracy both across countries (especially for inflation) and across forecasters. We find that the forecasts improve very slowly from long to short horizons, which may help to explain the large magnitude of the forecast errors. However, there is no evidence that the forecasts have improved over the years. On the other hand, we find that the forecasts are rather accurate in terms of directional changes. The findings also show that the GDP growth forecasts are unbiased for about half of the countries. For the inflation series, we often find a tendency to overpredict. Asia has experienced a decline in inflation over the past two decades, and the forecasters have failed to adjust fully to this trend of slowing inflation, causing an overprediction bias. As for the efficiency of the forecasts, evidence of a moderate underreaction is found for GDP growth, but not for inflation.

The paper is organized as follows. Section 2 presents the data. In Section 3, we assess the accuracy of the forecasts, and in particular the RMSE. In Section 4 we test for forecast unbiasedness, and in Section 5 we test for forecast efficiency. Section 6 investigates the sources of forecast accuracy disparities in Asia. Section 7 evaluates the directional forecast accuracy, and Section 8 concludes.

## 2. Data

In this study we use the *Asia Pacific Consensus Forecasts* data set, provided by Consensus Economics, which consists of monthly predictions made by a panel of professional forecasting institutions. We consider GDP growth and inflation forecasts for ten Asian economies, namely China, India, Indonesia, Taiwan, Hong Kong, Korea, Malaysia, Japan, Thailand, and Singapore. The forecast institutions surveyed are typically large and reputable financial institutions (e.g., commercial banks, investment banks, and insurance companies), industrial corporations, consulting firms, and research institutes. The sample includes both non-Asian forecasters (e.g., Goldman Sachs and Morgan Stanley) and local ones (e.g., Hyundai for Korea, Mizuho for Japan and Tata for India).

The structure of the data is as follows. Every month, each panelist forecasts GDP growth and inflation (i.e., the consumer price index inflation) for both the current year and the next year. Thus, each forecaster releases up to 24 forecasts for each target year. For instance, the first set of forecasts for the year 2012 is made in January 2011, and the final set is made in December 2012. Thus, the data set has a three-dimension panel structure, with 18 target years $t$ ($t = 1, \ldots, T$, with $T = 18$), 24 forecast horizons $h$ ($h = 1, \ldots, H$, with $H = 24$), and $N$ forecasters $i$ ($i = 1, \ldots, N$, where $N$ varies across countries). The horizon $h$ denotes the number of months ahead that the forecast is made. For instance, when a forecast is made in February 2012 for the target year 2012, the horizon is 11 months, i.e., $h = 11$. Therefore, horizons $h = 1$ to $h = 12$ correspond to current-year forecasts, and $h = 13$ to $h = 24$ correspond to forecasts for the following year, i.e., the forecasts made 13 to 24 months ahead. Our sample includes

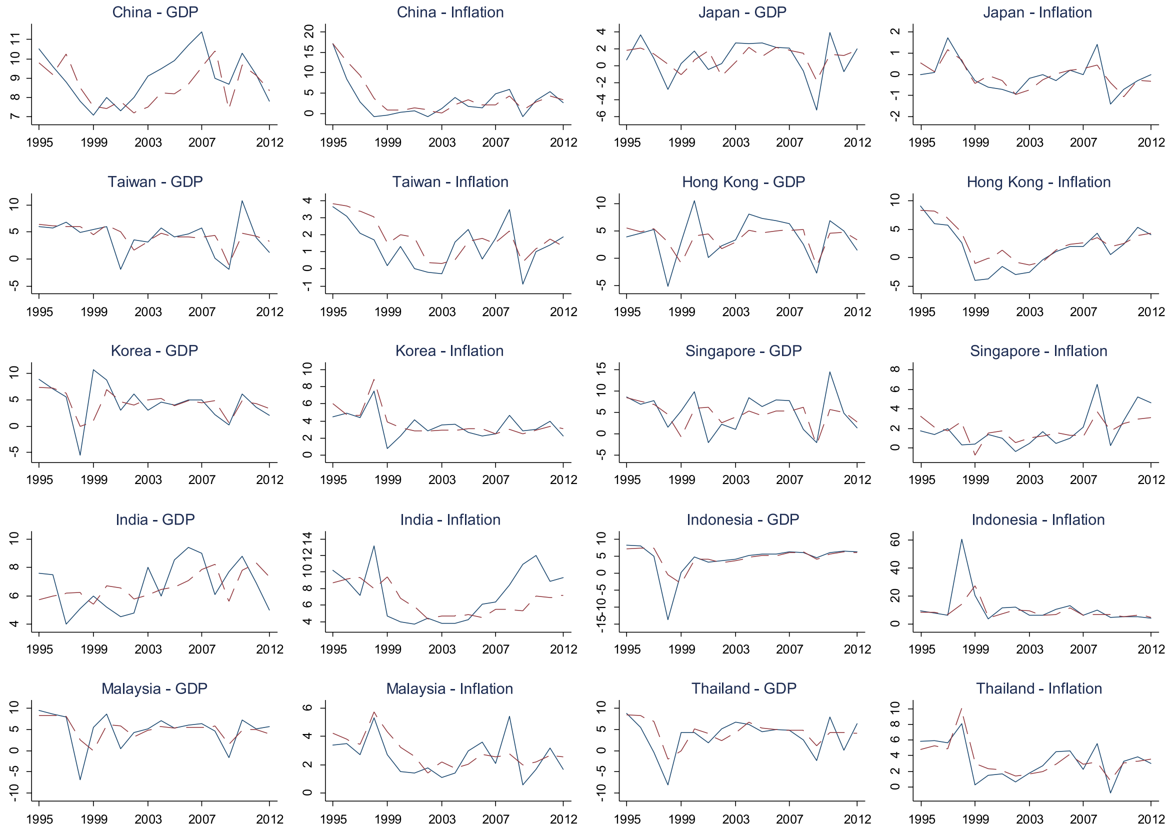


**Fig. 1.** Actual values (solid line) and consensus forecast at $h = 12$ (dashed line) for GDP growth and inflation.

forecasts made for the target years 1995–2012.[3] It follows that, for each forecaster, the maximum number of releases for each series is 420 (that is, $17 \times 24 + 12$).

One important aspect of this data set is that it is heavily unbalanced, as the set of forecasters participating in the survey changes over time. In addition, there are also gaps among the participating panelists, because they submit new forecasts at irregular dates, and do not always report current forecasts when they have no new forecasts. We discard forecasters with fewer than 100 observations, leaving a total of more than 51,000 observations and 175 forecasters for each variable. In our final sample, the number of forecasters ranges from a minimum of 13 for India and Indonesia to a maximum of 23 for Japan, and the number of observations ranges from a minimum of 3,354 for India to a maximum of 6,248 for Japan. Considering the selected forecasters, the number of observations corresponds to 62% of the fully balanced panel for Thailand (minimum) and 79% for Indonesia (maximum). It should be noted that more observations are available for short horizons than for long horizons, with the number of observations at $h = 1$ being approximately double that for $h = 24$.

Another aspect of the data set is that a given forecaster may be represented by several different versions of their name. For instance, the labels Citigroup and SSB Citibank refer to the same forecast institution. It is therefore essential to clean the data carefully and allocate the same unique forecaster ID to different labels when it is clear that they correspond to the same forecaster. For the realized values of GDP growth and inflation, we use the first IMF estimates, which are typically included in the April release of the World Economic Outlook of the following year.[4] In the Appendix, we discuss the robustness of our results using the latest available estimates of actual figures rather than the first estimates. Following the conventional notation, we denote the forecast made by panelist $i$ for the target year $t$ at forecast horizon $h$ by $f_{i,t,h}$. The actual value of the variable of interest for year $t$ is denoted by $A_t$, and $e_{i,t,h} = A_t - f_{i,t,h}$ represents the forecast error.

Fig. 1 shows the actual values of GDP growth and inflation, as well as the 12-month-ahead consensus forecasts (the mean of the individual forecasts). Consensus forecasts are noticeably more stable than the actual values, as large fluctuations in growth and inflation are usually recognized late in the year, passing from the January survey ($h = 12$) to that of December ($h = 1$) of the year to be forecast. For instance, the 12-month-ahead growth forecasts for Hong Kong failed to predict the recessions of 1998 and 2009, and

[3] Note that only horizons $h = 1$ to $h = 12$ are available for the target year 1995, but all 24 horizons are available for the remaining years, 1996–2012.

[4] It should be noted that the forecasts for India are made for fiscal years rather than calendar years, and we use the World Bank estimates for the actual values.

**Table 1**
Root mean squared errors averaged across forecasters.

| | China | Japan | Taiwan | Hong Kong | Korea | Singapore | India | Indonesia | Malaysia | Thailand |
|---|---|---|---|---|---|---|---|---|---|---|
| *GDP* | | | | | | | | | | |
| $h = 1$ | 0.36 | 0.44 | 0.63 | 0.61 | 0.57 | 0.65 | 0.80 | 0.59 | 0.42 | 0.89 |
| $h = 4$ | 0.57 | 0.88 | 1.07 | 1.12 | 0.85 | 1.39 | 0.97 | 1.14 | 0.82 | 1.61 |
| $h = 8$ | 0.90 | 1.38 | 2.09 | 2.48 | 1.97 | 2.83 | 1.41 | 1.54 | 2.13 | 3.08 |
| $h = 12$ | 1.13 | 1.85 | 2.41 | 3.09 | 2.39 | 3.55 | 1.75 | 2.68 | 3.23 | 3.60 |
| $h = 16$ | 1.42 | 2.56 | 3.16 | 4.52 | 4.37 | 4.45 | 1.70 | 4.68 | 4.75 | 4.91 |
| $h = 20$ | 1.59 | 2.66 | 3.21 | 4.31 | 4.14 | 4.63 | 1.75 | 5.16 | 5.15 | 5.52 |
| $h = 24$ | 1.71 | 2.50 | 3.03 | 3.90 | 3.93 | 4.22 | 1.93 | 3.84 | 3.69 | 4.71 |
| *Inflation* | | | | | | | | | | |
| $h = 1$ | 0.36 | 0.11 | 0.30 | 0.32 | 0.17 | 0.19 | 0.88 | 0.70 | 0.24 | 0.33 |
| $h = 4$ | 0.97 | 0.21 | 0.39 | 0.63 | 0.33 | 0.36 | 1.01 | 3.25 | 0.50 | 0.77 |
| $h = 8$ | 2.06 | 0.37 | 0.78 | 1.24 | 0.92 | 0.84 | 1.84 | 2.61 | 1.14 | 1.10 |
| $h = 12$ | 2.59 | 0.50 | 1.10 | 1.82 | 1.15 | 1.52 | 2.58 | 8.63 | 1.18 | 1.25 |
| $h = 16$ | 4.19 | 0.72 | 1.51 | 2.70 | 1.54 | 1.77 | 2.57 | 12.69 | 1.80 | 2.27 |
| $h = 20$ | 4.82 | 0.68 | 1.64 | 3.25 | 1.76 | 1.74 | 2.60 | 13.12 | 1.67 | 2.51 |
| $h = 24$ | 4.93 | 0.71 | 1.64 | 3.44 | 1.62 | 1.80 | 2.79 | 10.26 | 1.69 | 2.26 |

Notes: $h$ is the forecast horizon.

also failed to predict the strong growth of 2000 and 2004. Forecasters also have a limited ability to predict extreme events, such as the Indonesia hyperinflation of 1998 (60.7% inflation), at $h = 12$.

With regard to the actual values, there are large differences in the unconditional variability between countries. Inflation is considerably more stable in Japan (standard deviation of 0.75) than in Indonesia (12.98). Likewise, GDP growth is much more stable in China (1.24) and India (1.70) than in small open economies such as Hong Kong (3.84) and Singapore (4.42), or in South East Asia (Indonesia, Malaysia, Thailand). Apart from a few exceptions, such as inflation in Japan and GDP growth in China and India, GDP growth and inflation are noticeably more volatile in Asia than in the United States and other large non-Asian advanced economies. In Asia, recessions tend to be deeper and recoveries sharper, resulting in large fluctuations in economic activity and inflation. For instance, the GDP growth in Singapore jumped from −2% in 2009 to 14.7% in 2010, and the inflation in China fell from 17.1% to 8.4% between 1995 and 1996.

## 3. Forecast errors

In this section, we first report the root mean squared forecast error (RMSE) and the long-term predictability of each series. We then examine the evolution of the RMSE over forecast horizons and target years, and highlight some important facts.

### 3.1. RMSE and predictability

We assess the forecast accuracy using the root mean squared error. We define the RMSE for forecaster $i$ at horizon $h$ as $\text{RMSE}_{i,h} = \sqrt{T^{-1}\sum_{t=1}^{T} e_{i,t,h}^2}$, and the average of the individual RMSEs at horizon $h$ as $\text{RMSE}_h = \frac{1}{N}\sum_{i=1}^{N}\text{RMSE}_{i,h}$. In Table 1, we report the $\text{RMSE}_h$ for selected forecast horizons. Similarly to previous studies (see e.g. Lahiri & Sheng, 2010), we find that the forecast errors are mostly flat for approximately the first 10 months (i.e., $h > 14$). At long horizons, there are virtually no information gains, as the economic shocks tend to be absorbed fully during the current year, with no potential impact on the growth and inflation in the next year. After approximately the first 10 months (i.e., $h < 14$), the forecasts become increasingly accurate as the horizon shortens, and information about the actual value accumulates.

The forecast errors vary considerably across countries, especially at long and middle horizons. For instance, when GDP growth forecasts are considered, the $\text{RMSE}_{12}$ (i.e., the RMSE for January of the year to be forecast) is much higher in Singapore (3.55) and Malaysia (3.23) than in China (1.13) and India (1.70). The disparities are even wider for inflation; for example, the $\text{RMSE}_{12}$ is equal to 8.63 for Indonesia and 0.50 for Japan. In most cases, these figures are much higher than those reported in previous studies for developed non-Asian economies using the same data set (see e.g. Dovern & Weisser, 2011), indicating that growth and inflation are inherently difficult to forecast for most Asian countries. A few exceptions are the forecasts of the output growth in China and India, and of inflation in Japan. On average, the forecasts for the advanced economies (Japan, Taiwan, Hong Kong, Singapore and Korea) are no more accurate than those for the emerging economies (China, India, Indonesia, Malaysia, and Thailand). It should be noted that these findings are not driven by outliers (i.e., forecasters with extremely high RMSEs); for instance, using the median of individual RMSEs rather than the mean would provide almost exactly the same results.

Table 1 also shows that the RMSEs for inflation are lower than those for the GDP growth for most of the countries. This result, which has been reported previously for developed economies (e.g., Harvey, Leybourne, & Newbold, 2001), underscores the fact that the actual inflation is easier to predict. One possible reason for this may be that inflation is more stable than GDP growth. However, the reverse is observed in China, India and Indonesia. The output has traditionally been relatively simple to forecast in China, due to government control over the economic activity and its ability to meet growth targets. In India and Indonesia, there have been rather large inflation shocks (with inflation sometimes exceeding 10%), and inflation is more difficult to predict than stable growth.

A comparison of absolute RMSEs shows that GDP growth and inflation are more difficult to forecast in

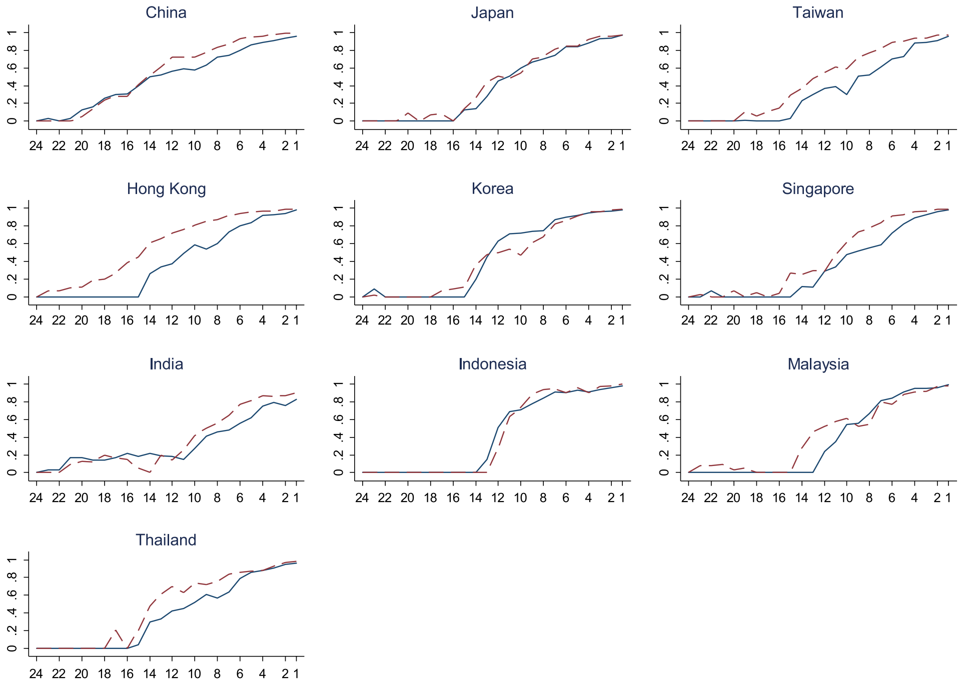


**Fig. 2.** Predictability of GDP growth (solid line) and inflation (dashed line): Diebold–Kilian statistics.

some countries than in others. However, it would be misleading to associate low RMSEs with high forecasting abilities, and some series can be intrinsically easier to predict than others for many different reasons. Therefore, we use the statistics by Diebold and Kilian (2001) to compare predictability performances (see also Lahiri & Sheng, 2010). More specifically, we define $p_{h,24}$ as the proportionate gain in mean squared error (MSE) between the horizon 24 forecasts and the horizon $h$ forecasts, such that $p_{h,24} = 1 - (MSE_h/MSE_{24})$.[5] The $p_{h,24}$ statistics shows the improvement in the forecast accuracy at horizon $h$ compared to the naïve forecast of horizon 24. The predictability naturally increases as we move from long to short horizons, and typically approaches 95%–100% at short horizons.

Fig. 2 shows that the predictability is higher for inflation than for growth for most of the countries and horizons, which confirms the impression that inflation is easier to predict. We find that the predictability remains at zero until late in the forecasting cycle for many countries, and for GDP growth in particular. For instance, for the GDP growth of Malaysia, $p_{h,24}$ only becomes positive after horizon 13 (i.e., December of the previous year), indicating that the first 12 months bring no useful information over and above the "naïve" forecast of horizon 24.

In general, we find considerable differences between countries. For instance, the predictability of GDP growth ranges from 0.18 for India (minimum) to 0.63 for Korea (maximum) when $h = 12$. For inflation, the predictability ranges from 0.29 for Singapore to 0.72 for Hong Kong when $h = 12$. On average, countries with a good predictability for GDP growth also tend to have a good predictability for inflation (the cross-country correlation is between 0.3 and 0.5 for most horizons). In general, China shows the best predictability of the various countries for both growth and inflation, whereas India has the lowest predictability for both series.

Finally, it should be noticed that the RMSE is correlated strongly with the unconditional variance of the actual values. Pooling across the forecasters, the correlation ranges from 0.9 to 1 for inflation, depending on the horizon, and from 0.7 to 1 for GDP growth. In contrast, the ranking of countries based on the Diebold–Kilian statistics is correlated only weakly with the ranking of countries based on the RMSE. Therefore, the high variability of growth and inflation in Asia may help explain the poor RMSE performance. The RMSEs of volatile series are particularly large at long horizons, when forecasters possess little economic information.

### 3.2. *Distribution of forecast errors*

Due to the unbalanced nature of our panel, it may not be particularly meaningful to compare RMSEs across individual forecasters directly. Indeed, panelists that have been active during time periods that are easy to forecast will obviously perform better. In order to take this issue into account, we follow Clements (2014) and use an adjusted RMSE measure, where the actual squared forecast errors of year $t$ are weighted by the cross-sectional average for year $t$ relative to the average over all years (see Clements, 2014).

[5] Note that we report the maximum between 0 and $p_{h,24}$. In practice, negative values for $p_{h,24}$ can occur when forecasters receive no meaningful information at the very long horizons and $MSE_h > MSE_{24}$.

**Table 2**
Distribution of adjusted-RMSE values across forecasters.

| | | China | Japan | Taiwan | Hong Kong | Korea | Singapore | India | Indonesia | Malaysia | Thailand |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $h = 6$ | | | | | | | | | | | |
| GDP | mean | 0.76 | 0.92 | 1.55 | 1.72 | 1.16 | 2.08 | 1.10 | 0.83 | 1.23 | 1.73 |
| | std | 0.10 | 0.08 | 0.21 | 0.19 | 0.17 | 0.20 | 0.08 | 0.19 | 0.16 | 0.17 |
| | min | 1.00 | 1.05 | 2.09 | 2.14 | 1.55 | 2.46 | 1.20 | 1.12 | 1.43 | 2.07 |
| | max | 0.60 | 0.76 | 1.21 | 1.39 | 0.96 | 1.69 | 0.99 | 0.48 | 0.99 | 1.49 |
| Inflation | mean | 1.06 | 0.29 | 0.59 | 0.68 | 0.48 | 0.52 | 1.30 | 1.66 | 0.62 | 0.81 |
| | std | 0.17 | 0.08 | 0.11 | 0.19 | 0.16 | 0.10 | 0.31 | 0.75 | 0.16 | 0.17 |
| | min | 1.30 | 0.45 | 0.74 | 1.23 | 0.78 | 0.74 | 1.73 | 3.54 | 0.87 | 1.08 |
| | max | 0.79 | 0.16 | 0.34 | 0.46 | 0.27 | 0.38 | 0.88 | 0.99 | 0.41 | 0.59 |
| corr. | | 0.46* | 0.39 | 0.10 | −0.28 | 0.47* | −0.59** | −0.39 | 0.10 | −0.12 | −0.08 |
| $h = 12$ | | | | | | | | | | | |
| GDP | mean | 1.12 | 1.62 | 2.00 | 2.51 | 1.79 | 3.08 | 1.65 | 1.59 | 2.41 | 2.69 |
| | std | 0.10 | 0.12 | 0.25 | 0.29 | 0.28 | 0.37 | 0.11 | 0.29 | 0.27 | 0.29 |
| | min | 1.34 | 1.81 | 2.35 | 3.00 | 2.28 | 3.58 | 1.84 | 1.96 | 2.76 | 3.00 |
| | max | 0.98 | 1.35 | 1.54 | 1.86 | 1.39 | 2.30 | 1.49 | 1.20 | 2.04 | 2.21 |
| Inflation | mean | 2.03 | 0.46 | 0.98 | 1.48 | 1.01 | 1.16 | 2.31 | 4.54 | 1.34 | 1.24 |
| | std | 0.24 | 0.08 | 0.10 | 0.15 | 0.21 | 0.12 | 0.34 | 0.96 | 0.52 | 0.16 |
| | min | 2.56 | 0.65 | 1.17 | 1.78 | 1.30 | 1.41 | 2.76 | 6.39 | 2.73 | 1.48 |
| | max | 1.63 | 0.28 | 0.86 | 1.25 | 0.71 | 0.97 | 1.53 | 2.90 | 0.87 | 1.03 |
| corr. | | −0.06 | 0.33 | 0.31 | −0.22 | 0.58** | 0.10 | 0.12 | 0.00 | 0.01 | 0.37 |

Notes: The table reports the distribution of adjusted-RMSE values across individual forecasters for horizons of 6 and 12, calculated as in Eqs. (1) and (2). Corr. indicates the cross-sectional correlation between the adjusted-RMSEs.
* Significant at the 10% level.
** Significant at the 5% level.

More specifically, the adjusted RMSE is calculated as

$$\mathrm{RMSE}_{i,h}^{adj} = \sqrt{T^{-1}\sum_{t=1}^{T} e_{i,t,h}^{*2}}, \qquad (1)$$

with

$$e_{i,t,h}^{*2} = e_{i,t,h}^{2} \times \frac{\mathrm{median}_t(\mathrm{median}_i(|e_{i,t,h}|))}{\mathrm{median}_i(|e_{i,t,h}|)}, \qquad (2)$$

where $\mathrm{median}_i$ is the cross-sectional median and $\mathrm{median}_t$ is the median over all $t$. Therefore, if the forecast errors are large at horizon $h$ and year $t$ compared with the forecast errors for the same horizon but other $t$, then the weight $\frac{\mathrm{median}_t(\mathrm{median}_i(|e_{i,t,h}|))}{\mathrm{median}_i(|e_{i,t,h}|)} < 1$ and the squared errors will be reduced. Note that we use the median rather than the mean in order to lessen the influence of outliers.

In Table 2, we show the cross-sectional distribution of the adjusted-RMSE across forecasters for two selected horizons, $h = 6$ and $h = 12$. We consider forecasters with at least 10 observations. In some cases, we find a large dispersion in accuracy. For instance, when considering the Korea GDP growth forecasts at $h = 12$, we find that the maximum and minimum adjusted-RMSEs are 2.28 and 1.39 (ratio of 1.64), respectively. In most cases, the adjusted-RMSE of the least accurate forecaster is about twice as large as the most accurate one, which may indicate differences in forecasting ability. For most of the countries, the dispersion is slightly larger for inflation than for GDP growth, and the largest cases of dispersion are all related to inflation forecasts.

Table 2 also shows the correlation between the adjusted-RMSEs of GDP growth and inflation. Some of the correlation coefficients are positive and some are negative, which leads us to conclude that there is no strong evidence that panelists with superior GDP growth forecasts also tend to produce more accurate inflation forecasts.

### 3.3. *Forecast errors over the horizons*

We indicate above that forecasts fail to improve substantially over approximately the first 10 months. Fig. 3 shows the evolution of information arrival across horizons. We calculate the change in RMSEs between two consecutive horizons as $\Delta\mathrm{RMSE}_h = \mathrm{RMSE}_{h+1} - \mathrm{RMSE}_h$, and scale it by $\mathrm{RMSE}_{24}$. A positive value for $\frac{\Delta\mathrm{RMSE}_h}{\mathrm{RMSE}_{24}}$ implies information gains between $h+1$ and $h$, whereas a negative value indicates that the forecasts have become less accurate. Rather than reporting the results for individual countries, we report the cross-country average in order get an indication of the timing of economic news in Asia.

We fit a non-parametric curve and find an inverted-L shaped relationship for both the GDP growth and inflation forecasts. The information gains are initially nonexistent but gradually increase, peaking at middle horizons as the economic news becomes increasingly informative. At short horizons, the information gains remain remarkably high, especially for GDP growth but also for inflation to lesser extent. These results contrast with those of Isiklar and Lahiri (2007), who find an inverted U-shape for advanced economies, and imply that the forecasts in Asia improve relatively slowly. The large forecast errors found in Asia may also be due to this. One possible explanation for this difference is the fact that the economic indicators in many Asian countries, including China and India (see Dovern et al., 2015; Nilsson & Brunet, 2006), are often not as informative of growth as is the case for countries such as the United States. Fewer quality indicators are available, which is expected to delay the acquisition of information. As a consequence, it may take longer for forecasters to form accurate expectations about GDP growth. Thailand and Taiwan are two examples of countries in which panelists keep on making large forecast revisions for GDP growth, even at the later stages of the

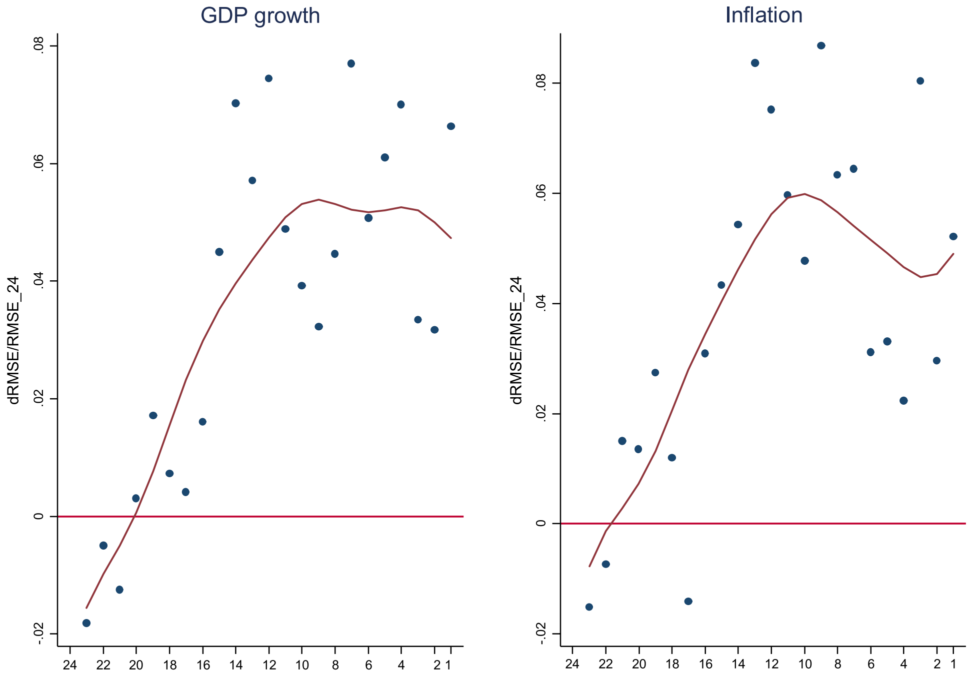


**Fig. 3.** Changes in RMSE between two consecutive horizons, averaged across forecasters.

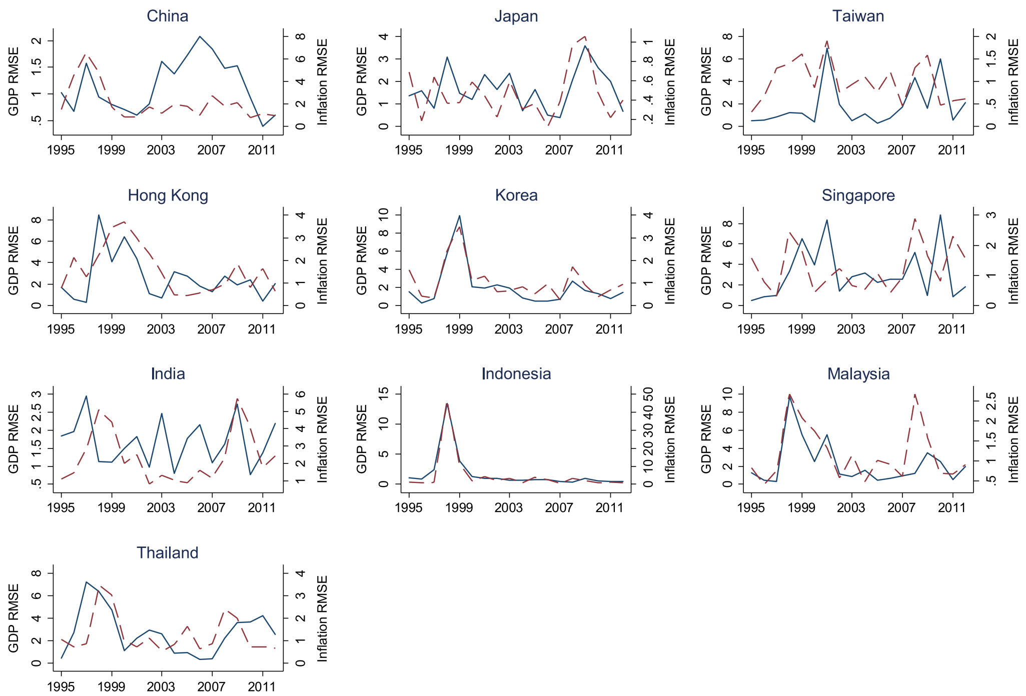


**Fig. 4.** RMSEs of GDP growth (solid line, left scale) and inflation forecasts for $h = 12$.

forecasting cycle, which leads to substantial accuracy improvements at short horizons.

### 3.4. Forecast errors over the years

We find that the forecast accuracy varies not only across forecasters, but also over time. In Fig. 4, we report the forecast errors when the horizon $h = 12$ is considered (note that the findings would be qualitatively the same if other horizons were selected). It turns out that the forecast errors are considerably higher during recession years than in calm periods. For most of the countries, the forecast errors increased sharply during the 1998 Asian crisis, before returning to low levels during the calm period of 2000–2007. The forecast errors increased again in 2008 and 2009, before starting to decline from 2010. China and India are two exceptions: their forecast errors are less cyclical, due to a stable economic growth and an absence of recessions. Interestingly, there is no evidence that forecasts in Asia have become more accurate over time. For instance, for most of the countries, the RMSEs over the period 2010–2012 are no lower than they were during the 1995–1997 and 2000–2007 periods.

Overall, our analysis indicates that the growing maturity of Asian economies has not been accompanied by an improved forecast accuracy. However, there are some notable exceptions. For instance, Indonesia's GDP growth and inflation forecasts have become more accurate over time, which reflects the country's long period of economic stability and lower inflation, beginning in the aftermath of the 1998 recession.

## 4. Testing forecast unbiasedness

In this section, we test for forecast unbiasedness. To do this, we use the error decomposition model that was proposed initially by Davies and Lahiri (1995) and extended later by Clements et al. (2007) and Dovern and Weisser (2011). The objective of this model is to have an estimator that accommodates the three-dimensional nature of the data set and provides standard errors that are consistent with the data structure. The model postulates that the forecast errors $e_{i,t,h}$, that is, the differences between the actual values and the forecasts, $e_{i,t,h} = A_t - f_{i,t,h}$, can be decomposed into three parts:

$$e_{i,t,h} = \phi_i + \lambda_{t,h} + \varepsilon_{i,t,h}, \tag{3}$$

where $\phi_i$ captures a forecaster-specific bias, $\lambda_{t,h}$ represents the effects of unanticipated macroeconomic shocks that occur between the time when the forecast is made and the end of year $t$, and $\varepsilon_{i,t,h}$ is the error term. For the analysis, it is assumed that $\lambda_{t,h} = \sum_{k=1}^{h} u_{t,k}$ (the sum of the shocks affecting the rational expectation value of the target variable), where $u_{t,k}$ has a mean of zero and a variance of $\sigma_u^2$, and $\varepsilon_{i,t,h} = \sum_{k=1}^{h} \eta_{i,t,k}$, where $\eta_{i,t,k}$ has a zero mean and a variance of $\sigma_i^2$ (see Deschamps & Ioannidis, 2013). We estimate the three components of the error model in Eq. (3) as follows:

$$\hat{\phi}_i = \frac{1}{TH} \sum_{t=1}^{T} \sum_{h=1}^{H} (A_t - f_{i,t,h}) \tag{4}$$

$$\hat{\lambda}_{t,h} = \frac{1}{N} \sum_{i=1}^{N} (A_t - f_{i,t,h} - \hat{\phi}_i) \tag{5}$$

$$\hat{\varepsilon}_{i,t,h} = A_t - f_{i,t,h} - \hat{\phi}_i - \hat{\lambda}_{t,h}. \tag{6}$$

In order to test unbiasedness for forecaster $i$, we test the hypothesis that $\phi_i = 0$ in Eq. (3); $\phi_i > 0$ and $\phi_i < 0$ indicate forecast underestimation and overestimation, respectively. A simple OLS regression of the forecast errors on a constant delivers a consistent estimate of the bias $\phi_i$. However, due to the error structure assumed in Eq. (3), we cannot use the OLS standard errors. In order to estimate standard errors, we therefore use a GMM-type estimator (see also Dovern & Weisser, 2011). This estimator accounts for the fact that the error terms are correlated across target years, forecast horizons and forecasters. The standard errors of the forecaster-specific bias $\hat{\phi}_i$ are estimated using the covariance matrix $(X'X)^{-1}X'\Sigma X(X'X)^{-1}$, where $\Sigma$ is the $NTH \times NTH$ error covariance matrix that is consistent with the error decomposition model. To estimate $\Sigma$, we need to compute the non-zero covariances between the composite error terms, which are given by:

$$\begin{aligned}&\mathrm{Cov}(A_{t_1} - f_{i,t_1,h_1}, A_{t_2} - f_{j,t_2,h_2}) \\ &= \mathrm{Cov}\left(\sum_{k=1}^{h_1} u_{t_1,k} + \sum_{k=1}^{h_1} \eta_{i,t_1,k}, \sum_{k=1}^{h_2} u_{t_2,k} + \sum_{k=1}^{h_2} \eta_{j,t_2,k}\right).\end{aligned} \tag{7}$$

To establish whether the forecasts are biased on average, we also perform a test of unbiasedness by imposing a common bias $\phi$ across forecasters. Due to sample size limitations, we do not provide any formal test of horizon-specific biases. Nonetheless, we report the mean forecast errors for selected horizons in Table 3 to show that they may vary across horizons.

The table shows that the magnitude of the mean forecast errors is typically larger at long horizons than at short horizons. Intuitively, the mean forecast errors are smaller at short horizons, due to the availability of superior information. In spite of these differences, it is worth estimating the overall bias in order to assess the general tendency toward over-/underpredicting growth and inflation. Table 4 summarizes the results pooled over all of the horizons (see Eq. (3)). For growth forecasts, the hypothesis of unbiasedness can only be rejected for China (0.33 percentage points), Thailand (−0.83) and Taiwan (−0.42). In the case of Thailand, the overprediction bias is explained by the fact that the country was hit by two deep recessions that the forecasters failed to predict. On the other hand, the forecasts for China underpredict growth, indicating that China's strong growth over the past two decades was not anticipated. For the remaining countries, the estimates are not significant.

Turning to individual forecasters, Table 4 shows that forecast unbiasedness cannot be rejected for most of the forecasters, in part because the correlation structure of the forecast errors leads to large standard errors. Overall, our analysis reveals differences in growth forecast biases between countries, in terms of both the direction and magnitude. Nonetheless, the forecast biases are only statistically significant for a minority of countries and forecasters.

As for inflation, the forecasts are significantly biased for five countries, namely China, Taiwan, India, Malaysia and Hong Kong. The estimates are negative for all countries except for Indonesia and India, indicating a broad tendency toward overestimation. Following the 1997–1998 crisis, Asia experienced a structural decline in inflation, and the forecasts have been adjusting too slowly, producing an overprediction bias. China is one example of this phenomenon. After experiencing an inflation of 17.1% in 1995, China saw a rapid reduction in inflation, which was largely unanticipated, causing an overprediction bias. India is an outlier. Its inflation has increased over the past

**Table 3**
Mean forecast errors.

| | China | Japan | Taiwan | Honk Kong | Korea | Singapore | India | Indonesia | Malaysia | Thailand |
|---|---|---|---|---|---|---|---|---|---|---|
| GDP | | | | | | | | | | |
| $h = 1$ | 0.11 | −0.06 | 0.12 | 0.14 | 0.02 | 0.16 | 0.16 | 0.19 | 0.14 | −0.11 |
| $h = 4$ | 0.19 | −0.20 | 0.02 | 0.33 | 0.06 | 0.20 | 0.15 | 0.28 | 0.20 | −0.31 |
| $h = 8$ | 0.32 | 0.09 | −0.04 | 0.19 | 0.08 | 0.48 | −0.01 | 0.01 | −0.06 | −0.66 |
| $h = 12$ | 0.44 | −0.10 | −0.31 | −0.01 | 0.03 | 0.35 | 0.04 | −0.37 | −0.19 | −0.99 |
| $h = 16$ | 0.30 | −0.62 | −0.93 | −0.35 | −0.76 | −0.44 | −0.21 | −1.40 | −0.93 | −1.75 |
| $h = 20$ | 0.36 | −0.87 | −1.03 | −0.73 | −0.76 | −0.69 | −0.38 | −1.59 | −1.33 | −1.97 |
| $h = 24$ | 0.41 | −0.63 | −0.94 | −0.46 | −0.84 | −0.42 | −0.31 | −1.24 | −0.81 | −1.86 |
| Inflation | | | | | | | | | | |
| $h = 1$ | −0.05 | 0.01 | −0.02 | −0.10 | −0.08 | −0.00 | 0.13 | 0.05 | −0.08 | −0.10 |
| $h = 4$ | −0.41 | −0.00 | −0.16 | −0.34 | −0.14 | −0.04 | 0.08 | −0.95 | −0.30 | −0.29 |
| $h = 8$ | −0.84 | −0.00 | −0.32 | −0.71 | −0.26 | −0.06 | 0.27 | 0.36 | −0.41 | −0.19 |
| $h = 12$ | −0.90 | 0.01 | −0.44 | −1.03 | −0.17 | −0.02 | 0.60 | 2.64 | −0.38 | 0.01 |
| $h = 16$ | −1.68 | −0.18 | −0.76 | −1.58 | −0.00 | −0.04 | 0.47 | 3.73 | −0.57 | −0.56 |
| $h = 20$ | −1.98 | −0.24 | −0.86 | −1.88 | −0.09 | −0.10 | 0.52 | 4.45 | −0.47 | −0.25 |
| $h = 24$ | −2.02 | −0.26 | −0.90 | −1.80 | −0.06 | −0.03 | 0.71 | 3.92 | −0.64 | −0.12 |

**Table 4**
Unbiasedness test results.

| | GDP | | | Inflation | | | No. of forecasters |
|---|---|---|---|---|---|---|---|
| | $\phi$ | $\phi_i > 0$ | $\phi_i < 0$ | $\phi$ | $\phi_i > 0$ | $\phi_i < 0$ | |
| Japan | −0.29 (0.23) | 1 | 5 | −0.07 (0.08) | 0 | 2 | 23 |
| China | 0.33*** (0.13) | 12 | 0 | −1.02*** (0.31) | 0 | 10 | 21 |
| Hong Kong | 0.02 (0.31) | 1 | 0 | −0.93*** (0.24) | 0 | 12 | 19 |
| Taiwan | −0.42* (0.29) | 0 | 2 | −0.45*** (0.14) | 0 | 10 | 18 |
| Korea | −0.30 (0.31) | 0 | 0 | −0.06 (0.21) | 0 | 0 | 17 |
| Singapore | −0.08 (0.68) | 1 | 0 | 0.01 (0.15) | 3 | 0 | 18 |
| Thailand | −0.83** (0.36) | 0 | 4 | −0.20 (0.25) | 0 | 2 | 16 |
| Malaysia | −0.28 (0.28) | 0 | 2 | −0.37** (0.17) | 0 | 7 | 16 |
| India | 0.01 (0.17) | 1 | 0 | 0.59** (0.29) | 2 | 1 | 13 |
| Indonesia | −0.49 (0.41) | 0 | 0 | 1.84 (1.26) | 1 | 0 | 13 |

Notes: $\phi$ indicates the bias parameter (see Section 4). $\phi_i > 0$ and $\phi_i < 0$ (see Eq. (3)) refer to the numbers of forecasters with positive (negative) biases at the 5% level. No. of forecasters denotes the number of forecasters. Standard errors are in parentheses.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

decade and the forecasters have failed to adjust, causing an underprediction bias.

We also compute the mean forecast errors for each month separately for every year, and find that the forecasts tend to be more heavily biased in months that precede large macroeconomic shocks.[6] As a result, the forecasts typically underpredict GDP during years of rapid growth and overpredict during recession years. For instance, forecasters were overly optimistic (by about 2–3 percentage points) for the 2009 GDP forecasts for most of the countries, as they failed to recognize the severity of the recession. Likewise, an overprediction bias can also be observed for the 1998 Asian crisis. A similar pattern is observed for inflation: the forecasters failed to predict unusual events such as an inflation of 60% in Indonesia in 1998, resulting in large forecast biases during those years.

## 5. Testing forecast efficiency

In this section, we test for weak form efficiency (see Nordhaus, 1987). The forecasts are efficient when they incorporate all of the available past information.[7] Nordhaus proposes a test that is based on restricting the set of information to the lagged forecast revisions. If the forecasts are efficient, future forecast revisions should be unpredictable. The hypothesis of efficiency implies $\beta_i = 0$ in the following regression of the forecast revisions on their lagged value:

$$r_{i,t,h} = \beta_i r_{i,t,h+1} + \xi_{i,t,h}, \tag{8}$$

where $r_{i,t,h} = f_{i,t,h} - f_{i,t,h+1}$ denotes the forecast revisions between horizons $h + 1$ and $h$. When $\hat{\beta}_i > 0$ in Eq. (8), the

[6] The results for this are not reported here, but are available from the authors upon request.

[7] The hypothesis of strong form efficiency is of limited practical use, as it requires a knowledge of the entire information set, which is not available to most econometricians.

**Table 5**
Efficiency test results.

| | GDP | | | Inflation | | | No. of forecasters |
|---|---|---|---|---|---|---|---|
| | $\beta$ | $\beta_i > 0$ | $\beta_i < 0$ | $\beta$ | $\beta_i > 0$ | $\beta_i < 0$ | |
| Japan | 0.12*** (0.03) | 9 | 0 | 0.04** (0.02) | 1 | 1 | 23 |
| China | 0.00 (0.03) | 0 | 0 | 0.00 (0.04) | 0 | 1 | 21 |
| Hong Kong | 0.08*** (0.03) | 2 | 1 | 0.03 (0.03) | 1 | 0 | 19 |
| Taiwan | 0.16*** (0.04) | 6 | 0 | 0.03 (0.03) | 0 | 0 | 18 |
| Korea | 0.10*** (0.03) | 6 | 0 | −0.06** (0.03) | 0 | 2 | 17 |
| Singapore | 0.14*** (0.03) | 7 | 0 | 0.02 (0.03) | 0 | 0 | 18 |
| Thailand | 0.07* (0.04) | 2 | 0 | 0.06** (0.03) | 2 | 0 | 16 |
| Malaysia | 0.08*** (0.03) | 1 | 0 | −0.02 (0.04) | 0 | 1 | 16 |
| India | 0.05 (0.04) | 0 | 0 | −0.01 (0.04) | 0 | 1 | 13 |
| Indonesia | 0.08* (0.04) | 3 | 0 | 0.03 (0.03) | 2 | 0 | 13 |

Notes: $\beta$ denotes the pooled estimates of Eq. (8). For the interpretation of $\beta_i > 0$ and $\beta_i > 0$, see Section 5. Standard errors are in parentheses. No. of forecasters denotes the number of forecasters.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

forecasts tend to be overly smooth. In other words, when new information arrives, forecasters may prefer to make several small revisions rather than a single large revision (underreaction), which results in positively autocorrelated revisions. On the other hand, when $\hat{\beta}_i < 0$, there is evidence of overreaction. The OLS estimator provides a consistent estimate for Eq. (8). However, statistical inference requires the correlation structure of forecast errors to be taken into account. Therefore, we use a GMM-type estimator to estimate $Var(\hat{\beta})$, and compute the elements of the covariance matrix as follows:

$$Cov(\xi_{i,t_1,h_1}, \xi_{j,t_2,h_2}) = Cov(u_{t_1,h_1+1} + \eta_{i,t_1,h_1+1}, u_{t_2,h_2+1} + \eta_{j,t_2,h_2+1}). \quad (9)$$

In our analysis, we also consider a pooled approach by imposing a common $\beta$ on all forecasters, in order to determine whether forecasters overreact or underreact to new information on average. We do not investigate horizon-specific $\beta$s, due to sample size limitations.

Table 5 reports the efficiency test results. When considering the GDP growth forecasts, the hypothesis of efficiency can be rejected for eight countries (at a 1% significance level for six countries and a 10% significance level for two countries). The estimates of $\beta$ are positive for all of the countries, indicating a general tendency to underreact to new information. However, these values are no larger than those reported for developed economies in previous studies (see for example Lahiri & Sheng, 2008). This indicates that Asia's volatile macroeconomic environment does not seem to affect the ability, or willingness, of forecasters to incorporate new information efficiently. However, at the individual forecaster level, forecast efficiency can be rejected at the 5% level for only a small number of individual forecasters (35 out of 175). Of these 35 forecasters, 34 show underreaction and just one shows overreaction.

As for the consensus forecast, Coibion and Gorodnichenko (2012) have shown that the correlation of the revisions can be explained by the infrequent updating of forecasters' information sets (i.e., "sticky information model"), as well as by the existence of noisy signals ("noisy information model"). However, neither of these models predicts the finding that individual forecast revisions are autocorrelated. As long as forecasters place the optimal weight on new information (see e.g. Lahiri & Sheng, 2008), individual forecast revisions should be unpredictable. In other words, evidence that $\beta_i > 0$ shows that there is more stickiness in the forecasts than noisy information models would predict.

The finding of forecast underreaction can be explained by behavioral factors. Ehrbeck and Waldmann (1996) argue that forecasters may not care about accuracy per se, but instead may seek to mimic the forecasting patterns of well-informed forecasters, in order to enhance their own reputations. In this setting, they show that forecasters may be reluctant to make large forecast revisions because large revisions signal that the previous forecasts were wrong. Therefore, forecasters would be expected to adjust forecasts insufficiently upon the arrival of new information. This is termed "rational stubbornness". Deschamps and Ioannidis (2013) find evidence of rational stubbornness among professional forecasters for the G-7 countries. In the same vein, Batchelor and Dua (1992) argue that clients may perceive forecasters who change their forecasts frequently to be erratic. As a result, forecasters may strategically choose to underreact to new information. Another possible explanation may be that forecasts are overly sticky due to herding behavior. For instance, Ottaviani and Sorensen (2006) show that it is optimal to bias forecasts towards the consensus, so as to appear better informed. Because of herding behaviors, forecasts will be adjusted to new information gradually rather than immediately, causing positive autocorrelation of revisions.

Dovern et al. (2015) also study forecast efficiency for a larger set of countries, including the Asian countries. However, they use a different methodology and focus only on GDP growth, whereas we study both GDP growth and inflation. They find forecast smoothing to be more pronounced for emerging economies than for advanced economies, which they explain as being due to the lower quality of economic statistics in emerging countries. Interestingly, we find the opposite result. In our sample, five countries may be classified as advanced economies (Japan, Singapore, Korea, Taiwan, and Hong Kong) and five as developing economies (China, Thailand, Malaysia, India, and Indonesia). We find that forecast underreaction for GDP growth is always larger for the five advanced countries (from a minimum of 0.08 in Hong Kong to a maximum of 0.16 in Taiwan) than for the five developing countries (from a minimum of 0.00 for China to a maximum of 0.08 in Malaysia).

Turning to inflation, forecast efficiency can be rejected for three countries, namely Japan (underreaction), Korea (overreaction) and Thailand (underreaction). It is noticeable that the estimates of $\beta$ for inflation are smaller than those for GDP growth. In addition, individual forecast efficiency can be rejected for only a very small number of forecasters (12 out of 175), further indicating that new information is incorporated into inflation forecasts more promptly than GDP growth forecasts. Compared to previous analyses for developed countries (see for example Dovern & Weisser, 2011), no strong evidence against the efficiency of inflation forecasts for Asia is found.

## 6. Assessment of forecast errors

We have argued in Section 3 that the low predictability and high unconditional variance of growth and inflation may have contributed to the overall high RMSE of Asian forecasts. In this section, we discuss the roles of forecast under-/overreaction and systematic biases in explaining the high RMSE. In general, forecast under-/overreaction is expected to have an adverse effect on the forecast accuracy. The forecast errors tend to be larger than those obtained when the individual forecasts are not optimal, e.g., when new information is incorporated overly slowly.

Our results for inflation show that the degree of forecast over-/underreaction is almost zero, indicating that there is no evidence that the poor performances of the forecasts in terms of RMSEs are due to an inefficient use of information. The degree of underreaction is also low for GDP growth (maximum of 0.16 for Taiwan and cross-country average of 0.09), being comparable to that found in previous studies for the G-7 economies. In other words, the intensity of forecast underreaction is not particularly high, and the high RMSEs in Asia cannot be explained by an inefficient use of information. To investigate this issue further, we also compute the cross-country correlation between the RMSE and the estimated $\beta$. The correlations are low and insignificant (0.20 for the GDP growth and −0.11 for inflation), confirming there is no evidence of a link between underreaction and forecast accuracy in our sample.

Systematic biases are also expected to have an adverse effect on the forecast accuracy. In order to assess the role played by biases, we filter the estimated biases from the actual forecasts and calculate bias-adjusted forecasts, which we denote by $f^*_{i,t,h} = f_{i,t,h} + \hat{\phi}_{i,h}$, where $\hat{\phi}_{i,h} = \frac{1}{T}\sum_{t=1}^{T}(A_t - f_{i,t,h})$ is the forecaster- and horizon-specific bias. We denote the mean of the individual RMSE for the bias-adjusted forecasts by $\text{RMSE}^*_h$,[8] and we expect that $\text{RMSE}^*_h < \text{RMSE}_h$.

Table 6 reports $\text{RMSE}^*_h$ values for the selected horizons $h = 1, 12, 24$. When comparing the results in Table 6 with those in Table 1, we find that $\text{RMSE}^*_h < \text{RMSE}_h$. In particular, for the GDP growth forecasts, RMSE would be 3%–19% lower if there was no bias (see Tables 1 and 6). For inflation, the range is from 3% to 25%. We find that the RMSE disparities for the bias-adjusted forecasts are as large as those for the unadjusted forecasts, which shows that biases do not seem to play a major role in explaining why some countries have such large RMSEs. For instance, the GDP growth forecasts for China are much more accurate than those for Thailand, and this would still be the case even after adjusting for biases. Furthermore, the RMSEs of the bias-adjusted forecasts are still well above the unadjusted RMSEs found in other studies for non-Asian advanced economies (see e.g. Dovern & Weisser, 2011), further indicating that biases cannot explain much of the poor RMSE performances of Asian forecasts.

Overall, we argue that biases and forecast underreaction do not seem to explain much of the poor forecast performance in Asia. The performances of the forecasts would remain poor and the RMSE disparities would persist even in the absence of systematic biases and underreaction.

## 7. Directional accuracy

Some studies have pointed out that being able to forecast the direction of the change accurately is particularly important for investors and policymakers (Altavilla & De Grauwe, 2010; Bergmeir, Costantini, & Benítez, 2014; Blaskowitz & Herwartz, 2009, 2011, 2014). For investors, an investment decision that is driven by a specific macroeconomic forecast with a small forecast error may not necessarily be as profitable as an investment decision that is guided by an accurate prediction of the direction of change. For policymakers, directional predictions are crucial for adjusting policy instruments, such as increasing or decreasing interest rates (Öller & Barot, 2000).

In this section, we analyse the directional accuracy of the professional forecasts in Asia. To this end, we use the following measure (see Blaskowitz & Herwartz, 2009):

$$L^{DA}_{i,t,h} = I((f_{i,t,h} - A_{t-1})(A_t - A_{t-1}) > 0) - I((f_{i,t,h} - A_{t-1})(A_t - A_{t-1}) < 0), \quad (10)$$

where $I(\cdot)$ is an indicator function and $L^{DA}_{i,t,h}$ takes the value 1 (−1) if the direction of change is predicted correctly (incorrectly). We calculate the average of $L^{DA}_{i,t,h}$ among the forecasters as $L^{DA}_h = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} L^{DA}_{i,t,h}$, for selected horizons $h = 1, 4, 8, 12$. Table 7 reports the results. A positive value of $L^{DA}_h$ indicates that the forecasts outperform a random toss of a coin. The values are largely positive for both

[8] More specifically, $\text{RMSE}^*_{i,h} = \sqrt{T^{-1}\sum_{t=1}^{T}(e_{i,t,h} - \hat{\phi}_{i,h})^2}$ and $\text{RMSE}^*_h = \frac{1}{N}\sum_{i=1}^{N}\text{RMSE}^*_{i,h}$.

**Table 6**
Bias-adjusted RMSE, averaged across forecasters.

| | China | Japan | Taiwan | Hong Kong | Korea | Singapore | India | Indonesia | Malaysia | Thailand |
|---|---|---|---|---|---|---|---|---|---|---|
| GDP | | | | | | | | | | |
| $h = 1$ | 0.33 | 0.42 | 0.61 | 0.52 | 0.55 | 0.59 | 0.65 | 0.51 | 0.38 | 0.82 |
| $h = 12$ | 0.98 | 1.80 | 2.32 | 2.92 | 2.30 | 3.34 | 1.61 | 2.35 | 3.05 | 3.12 |
| $h = 24$ | 1.47 | 2.30 | 2.85 | 3.85 | 3.54 | 4.08 | 1.71 | 3.28 | 3.50 | 4.02 |
| Inflation | | | | | | | | | | |
| $h = 1$ | 0.34 | 0.11 | 0.27 | 0.28 | 0.15 | 0.16 | 0.80 | 0.65 | 0.19 | 0.23 |
| $h = 12$ | 2.20 | 0.49 | 0.91 | 1.37 | 1.08 | 1.37 | 2.24 | 7.39 | 1.08 | 0.99 |
| $h = 24$ | 3.71 | 0.62 | 1.23 | 2.47 | 1.47 | 1.65 | 2.21 | 8.73 | 1.51 | 1.58 |

Notes: The table reports $RMSE^*_h$ (see Section 6) for selected horizons.

**Table 7**
Directional accuracy.

| | China | Japan | Taiwan | Hong Kong | Korea | Singapore | India | Indonesia | Malaysia | Thailand |
|---|---|---|---|---|---|---|---|---|---|---|
| *GDP* | | | | | | | | | | |
| All obs. | | | | | | | | | | |
| $h = 1$ | 0.88 | 0.74 | 0.94 | 0.98 | 0.94 | 0.90 | 0.76 | 0.92 | 0.92 | 0.92 |
| $h = 4$ | 0.64 | 0.48 | 0.94 | 0.92 | 0.94 | 0.92 | 0.64 | 0.82 | 0.70 | 0.78 |
| $h = 8$ | 0.34 | 0.44 | 0.60 | 0.86 | 0.84 | 0.78 | 0.50 | 0.68 | 0.56 | 0.60 |
| $h = 12$ | 0.16 | 0.46 | 0.56 | 0.88 | 0.72 | 0.64 | 0.30 | 0.52 | 0.58 | 0.44 |
| $\Delta A_t > 0$ | | | | | | | | | | |
| $h = 1$ | 0.72 | 0.72 | 0.94 | 0.98 | 0.86 | 0.92 | 0.66 | 0.86 | 0.86 | 0.80 |
| $h = 4$ | 0.32 | 0.48 | 0.60 | 0.84 | 0.82 | 0.80 | 0.52 | 0.72 | 0.44 | 0.64 |
| $h = 8$ | −0.18 | 0.48 | 0.46 | 0.82 | 0.66 | 0.54 | 0.42 | 0.54 | 0.20 | 0.62 |
| $h = 12$ | −0.42 | 0.32 | 0.50 | 0.92 | 0.50 | 0.30 | 0.22 | 0.32 | 0.22 | 0.56 |
| $\Delta A_t < 0$ | | | | | | | | | | |
| $h = 1$ | 0.98 | 0.74 | 0.94 | 0.98 | 1.00 | 0.88 | 0.86 | 0.98 | 1.00 | 1.00 |
| $h = 4$ | 0.88 | 0.50 | 0.84 | 0.98 | 1.00 | 0.98 | 0.78 | 0.96 | 1.00 | 0.88 |
| $h = 8$ | 0.76 | 0.40 | 0.74 | 0.88 | 0.94 | 0.96 | 0.58 | 0.84 | 0.98 | 0.56 |
| $h = 12$ | 0.62 | 0.58 | 0.62 | 0.84 | 0.86 | 0.84 | 0.36 | 0.78 | 1.00 | 0.34 |
| *Inflation* | | | | | | | | | | |
| All obs. | | | | | | | | | | |
| $h = 1$ | 0.90 | 0.98 | 0.94 | 0.90 | 0.98 | 0.98 | 0.40 | 0.94 | 0.92 | 0.84 |
| $h = 4$ | 0.88 | 0.90 | 0.80 | 0.82 | 0.86 | 0.82 | 0.44 | 0.86 | 0.78 | 0.72 |
| $h = 8$ | 0.78 | 0.64 | 0.42 | 0.76 | 0.70 | 0.58 | 0.06 | 0.74 | 0.56 | 0.72 |
| $h = 12$ | 0.54 | 0.42 | 0.22 | 0.62 | 0.62 | 0.44 | −0.14 | 0.62 | 0.38 | 0.50 |
| $\Delta A_t > 0$ | | | | | | | | | | |
| $h = 1$ | 0.82 | 0.96 | 1.00 | 0.92 | 0.96 | 0.98 | 0.38 | 0.92 | 0.98 | 0.88 |
| $h = 4$ | 0.86 | 0.96 | 0.94 | 0.88 | 0.84 | 0.92 | 0.50 | 0.88 | 1.00 | 0.80 |
| $h = 8$ | 0.96 | 0.76 | 0.80 | 0.84 | 0.68 | 0.80 | 0.10 | 0.66 | 0.92 | 0.60 |
| $h = 12$ | 0.76 | 0.58 | 0.72 | 0.74 | 0.62 | 0.74 | −0.22 | 0.46 | 0.72 | 0.34 |
| $\Delta A_t < 0$ | | | | | | | | | | |
| $h = 1$ | 0.98 | 0.98 | 0.90 | 0.88 | 1.00 | 1.00 | 0.44 | 0.94 | 0.86 | 0.76 |
| $h = 4$ | 0.90 | 0.86 | 0.66 | 0.70 | 0.88 | 0.88 | 0.34 | 0.86 | 0.56 | 0.54 |
| $h = 8$ | 0.62 | 0.52 | 0.10 | 0.64 | 0.74 | 0.74 | 0.00 | 0.80 | 0.26 | 0.94 |
| $h = 12$ | 0.34 | 0.24 | −0.16 | 0.42 | 0.64 | 0.64 | −0.04 | 0.80 | 0.02 | 0.86 |
| AR(1) | | | | | | | | | | |
| GDP | | | | | | | | | | |
| all obs. | 0.29 | 0.29 | 0.53 | 0.06 | 0.29 | 0.53 | 0.06 | 0.06 | 0.18 | −0.06 |
| $\Delta A_t > 0$ | 0.25 | 0.50 | 0.50 | 0.25 | 1.00 | 0.71 | 0.00 | 0.11 | 0.11 | −0.43 |
| $\Delta A_t < 0$ | 0.33 | 0.11 | 0.56 | −0.11 | −0.17 | 0.40 | 0.11 | 0.00 | 0.25 | 0.20 |
| Inflation | | | | | | | | | | |
| all obs. | 0.06 | 0.06 | 0.18 | 0.18 | 0.41 | −0.18 | 0.18 | −0.18 | 0.41 | 0.06 |
| $\Delta A_t > 0$ | 0.11 | 0.11 | 0.00 | 0.11 | 0.20 | 0.00 | 0.11 | 0.00 | 0.20 | 0.27 |
| $\Delta A_t < 0$ | 0.00 | 0.00 | 0.33 | 0.00 | 0.71 | −0.43 | 0.25 | −0.33 | 0.71 | 0.33 |

Notes: The values indicate the directional accuracy loss, given by Eq. (10).

growth and inflation, indicating that the professional forecasts have positive value for predicting directions.

To evaluate the performances of the panelists in terms of directional accuracy further, we also consider an AR(1) model as a benchmark model. The AR(1) model is estimated using actual values, and its out-of-sample forecasting performance is given in Table 7.[9] The professional forecasts beat the AR(1) in predicting the direction of change for almost all of the countries, regardless of the horizon. The directional accuracy is remarkably high in some cases (see the forecasts of GDP growth for Hong Kong), although we do observe large disparities across countries. Unsurprisingly, the directional accuracy improves vastly when moving to shorter horizons, as the information underlying the forecasts becomes increasingly accurate.

[9] The AR forecast is denoted $\hat{A}^{AR}_t = \hat{\theta}_0 + \hat{\theta}_1 A_{t-1}$.

The results in Table 7 also show that forecasters are not equally good at predicting positive and negative changes. For the forecasts of GDP growth, DA is typically higher when the change is negative ($\Delta A_t = A_t - A_{t-1} < 0$), which implies that panelists predict slowdowns better than accelerations. A close look at the data reveals that almost all forecasters predict the sign of change correctly for most of the years. However, for each country there are one or two years where DA is very low, even at short horizons.

When averaging across horizons, for instance, the loss function is −0.64 for Malaysia in 2007, −0.06 for Singapore in 1997, −0.68 for Taiwan in 2007, and −0.90 for China in 2005. It turns out that these very low values of DA are observed during years of positive change, which explains why DA is lower for accelerations. In each of these cases, the low DA value for that year was preceded by another acceleration, with forecasters usually failing to predict the second acceleration. For instance, GDP growth accelerated in China in 2003 and 2004, and panelists were surprised to the very end by the further acceleration in 2005. The same phenomenon occurred in Taiwan, Indonesia, Malaysia and Korea. In other words, forecasters seem to be relatively poor at forecasting changes when the economy accelerates for two consecutive years.

Turning to inflation, the results are more mixed, and we find that positive changes are predicted correctly more often than negative changes for several countries. This finding also reflects the fact that Asia has made great progress in fighting against inflation (see Filardo & Genberg, 2010), and forecasters have regularly failed to anticipate inflation slowdowns, resulting in a relatively low DA for negative changes. Interestingly, the countries that have adopted explicit inflation targeting (Indonesia in 2000, Korea in 1999, and Thailand in 2000) have been more successful at predicting negative changes. A possible explanation for this may be that the downward trend in inflation was predictable due to the government commitment to stick to low inflation for these three countries.

It is worth noting that a country that performs well in terms of DA will not necessarily perform well in terms of RMSE, and vice versa. For example, China ranks first in terms of RMSE for GDP growth, but shows the worst result for DA, whereas Indonesia does the opposite for inflation. For some other countries, the forecast performances are equally good/bad in terms of the two accuracy measures. This suggests that the two accuracy measures are distinct, and both should be considered when assessing the overall forecast performance.

## 8. Conclusion

In this paper, we have provided a comprehensive assessment of the performances of GDP growth and inflation forecasts for a set of ten Asian economies over the period 1995–2012. We have evaluated the accuracy of the forecasts using both RMSE and a directional forecast accuracy measure, and tested for unbiasedness and efficiency. The results are as follows. First, the forecast errors are large for most of the countries; nevertheless, the forecasts are still directionally accurate. Large disparities in the magnitudes of the forecast errors (and long-term predictability) are also observed across countries, for both GDP growth and inflation. For most of the countries, the forecast accuracy is higher for inflation than for growth, which underscores the fact that inflation is intrinsically easier to predict. Further, the forecast accuracy in Asia improves relatively slowly as we move from long to short horizons. This result may also help to explain the high RMSEs. Second, the unbiasedness hypothesis cannot be rejected for the majority of countries. However, the inflation forecasts show a tendency to overpredict, which may be due to the decline of inflation in Asia. Finally, the hypothesis that forecasters incorporate new information efficiently is rejected widely for the forecasts of GDP growth, indicating a tendency to underreact, whereas we find little evidence of information stickiness for inflation.

This paper also contributes to the literature on the relative forecasting performances of advanced and emerging economies. Our results show that there is no correlation between forecast accuracy (and predictability) and the degree of economic development. However, surprisingly and unlike previous studies, we find that the underreaction for forecasts of GDP growth is more pronounced for advanced economies. Overall, we find little evidence that forecasters perform better in advanced economies (Singapore or Korea) than in emerging countries (China or India). Future research exploring the channels through which economic development affects forecast performances would be very beneficial.

## Appendix. Initial versus revised figures

Throughout the paper, we have evaluated forecasts using the initial estimates of GDP growth and inflation rather than the revised figures. It is possible that some forecasters may target revised figures rather than the initial announcement, and it is important to verify that our main results are robust to the use of revised figures. Starting with inflation, the revised and initial IMF figures are actually extremely similar. The mean absolute difference between initial and revised inflation estimates is less than 0.1%, with the exception of Indonesia (0.3%). None of our main results would be affected if we used revised figures. For GDP, however, the situation is slightly different. In China and Singapore, we observe average upward GDP estimate revisions of 0.7% and 0.5% respectively. The mean absolute difference between initial and revised figures is considerably larger than for inflation, ranging from 0.2% for Korea to 1.2% for Singapore. Using revised figures as the benchmark, the estimated RMSEs are generally unaffected, except for China, where the RMSEs would almost double. In general, the RMSEs are smaller when using the initial figures, which is consistent with the view that panelists target initial estimates. In terms of GDP unbiasedness and efficiency tests, the statistical significance of the estimates would not be affected.

## References


Ager, P., Kappler, M., & Osterloh, S. (2009). The accuracy and efficiency of the consensus forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting*, *25*, 167–181.

Altavilla, C., & De Grauwe, P. (2010). Forecasting and combining competing models of exchange rate determination. *Applied Economics*, *42*, 3455–3480.

Ashiya, M. (2005). Twenty-two years of Japanese institutional forecasts. *Applied Financial Economics Letters*, *12*, 79–84.

Batchelor, R., & Dua, P. (1992). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, *11*, 169–181.

Bergmeir, C., Costantini, M., & Benítez, J. M. (2014). On the usefulness of cross-validation for directional forecast evaluation. *Computational Statistics and Data Analysis*, *76*, 132–143.

Blaskowitz, O., & Herwartz, H. (2009). Adaptive forecasting of the EURIBOR swap term structure. *Journal of Forecasting*, *28*(7), 575–594.

Blaskowitz, O., & Herwartz, H. (2011). On economic evaluation of directional forecasts. *International Journal of Forecasting*, *27*, 1058–1065.

Blaskowitz, O., & Herwartz, H. (2014). Testing the value of directional forecasts in the presence of serial correlation. *International Journal of Forecasting*, *30*, 30–42.

Capistrán, C., & López-Moctezuma, G. (2014). Forecast revisions of Mexican inflation and GDP growth. *International Journal of Forecasting*, *30*, 177–191.

Carvalho, A., & Minella, A. (2012). Survey forecasts in Brazil: A prismatic assessment of epidemiology, performance, and determinants. *Journal of International Money and Finance*, *31*, 1371–1391.

Clements, M. P. (2014). Forecast uncertainty—ex ante and ex post: U.S. inflation and output growth. *Journal of Business and Economic Statistics*, *32*(2), 206–216.

Clements, M. P., Joutz, F., & Stekler, H. O. (2007). An evaluation of the forecasts of the federal reserve: a pooled approach. *Journal of Applied Econometrics*, *22*, 121–136.

Clements, M. P., & Taylor, N. (2001). Robust evaluation of fixed-event forecast rationality. *Journal of Forecasting*, *20*, 285–295.

Coibion, O., & Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, *120*(1), 116–159.

Costantini, M., & Kunst, R. M. (2011). Combining forecasts based on multiple encompassing tests in a macroeconomic core system. *Journal of Forecasting*, *30*, 579–596.

Davies, A., & Lahiri, K. (1995). A new framework for analysing survey forecasts using three-dimensional panel data. *Journal of Econometrics*, *68*, 205–227.

Deschamps, B., & Bianchi, P. (2012). An evaluation of Chinese macroeconomic forecasts. *Journal of Chinese Economic and Business Studies*, *10*, 229–246.

Deschamps, B., & Ioannidis, C. (2013). Can rational stubbornness explain forecast biases? *Journal of Economic Behavior and Organization*, *92*, 141–151.

Diebold, F. X., & Kilian, L. (2001). Measuring predictability: theory and macroeconomic applications. *Journal of Applied Econometrics*, *16*, 657–669.

Dovern, J., Fritsche, U., Loungani, P., & Tamirisa, N. (2015). Information rigidities: Comparing average and individual forecasts for a large international panel. *International Journal of Forecasting*, *31*(1), 144–154.

Dovern, J., & Weisser, J. (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. *International Journal of Forecasting*, *27*, 452–465.

Ehrbeck, T., & Waldmann, R. (1996). Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics*, *111*, 21–41.

Filardo, A., & Genberg, H. (2010). Targeting inflation in Asia and the Pacific: Lessons from the recent past. In J. Caruana, A. Filardo, J. George, A. Munro, M. Loretan, G. Ma, et al. (Eds.), *BIS papers: Vol. 52. The international financial crisis and policy challenges in Asia and the Pacific*.

Golinelli, R., & Parigi, G. (2008). Real time squared: A real-time data set for real-time GDP forecasting. *International Journal of Forecasting*, *24*, 368–385.

Golinelli, R., & Parigi, G. (2014). Tracking world trade and GDP in real time. *International Journal of Forecasting*, *30*, 847–862.

Harvey, D., Leybourne, S., & Newbold, P. (2001). Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal*, *4*, 37–55.

Hong, K., Lee, J., & Tang, H. (2010). Crises in Asia: historical perspectives and implications. *Journal of Asian Economies*, *21*, 265–279.

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

Isiklar, G., Lahiri, K., & Loungani, P. (2006). How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics*, *21*, 703–725.

Krkoska, L., & Teksoz, U. (2009). How reliable are forecasts of GDP growth and inflation for countries with limited coverage? *Economic Systems*, *33*, 376–388.

Lahiri, K., & Isiklar, G. (2009). Estimating international transmission of shocks using GDP forecasts: India and its trading partners. In S. Ghatak, & P. Levine (Eds.), *Development macroeconomics, essays in memory of Anita Ghatak*. Routledge.

Lahiri, K., & Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, *144*, 325–340.

Lahiri, K., & Sheng, X. (2010). Learning and heterogeneity in GDP and inflation forecasts. *International Journal of Forecasting*, *26*, 265–292.

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

Loungani, P., Stekler, H., & Tamirisa, N. (2013). Information rigidity in growth forecasts: some cross-country evidence. *International Journal of Forecasting*, *29*, 605–621.

Nilsson, R., & Brunet, O. (2006). *Composite leading indicators for major OECD non-member economies: Brazil, China, India, Indonesia, Russian Federation, South Africa*. OECD Statistics Working Paper 2006:1, Paris.

Nordhaus, W. (1987). Forecasting efficiency: concepts and applications. *Review of Economics and Statistics*, *69*, 667–674.

Öller, L., & Barot, B. (2000). The accuracy of European growth and inflation forecasts. *International Journal of Forecasting*, *16*, 293–315.

Ottaviani, M., & Sorensen, P. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, *81*, 441–466.



**Qiwei Chen** is lecturer in finance and accounting at Brunel University. She earned her Ph.D. in Finance at the University of Essex. Her main interests are behavioural finance and behavioural accounting.

**Mauro Costantini** is senior lecturer in economics at Brunel University. He earned his Ph.D. in Mathematics, Statistics and Economics at the University of Rome 'La Sapienza'. He has published papers in journals such as the *Journal of Forecasting, International Journal of Forecasting, Computational Statistics and Data Analysis*, and *Oxford Bulletin of Economics and Statistics*, among others. His main research interests are time series analysis, forecasting and panel data econometrics.

**Bruno Deschamps** is assistant professor in economics at the Nottingham University Business School China. He earned his Ph.D in Economics at ECARES, University of Brussels. His main interests are macroeconomic forecast evaluation, theory of strategic forecasting, panel data econometrics, macroeconomy and financial markets betting markets. He has published papers in journals such as the *Journal of Economic Behavior and Organization, Manchester School*, and *Review of Quantitative Finance and Accounting*.