Department of
Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 15-16

Qiwei Chen, Mauro Costantini, and Bruno Deschamps

May 2015

http://www.brunel.ac.uk/economics

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

Qiwei Chen[1], Mauro Costantini [2], and Bruno Deschamps[3]

[1]Brunel University, Department of Economics and Finance, United Kingdom
[2]Brunel University, Department of Economics and Finance, United Kingdom
[3]Nottingham University Business School China

Abstract

This paper assesses the performance of professional GDP growth and in ation forecasts for ten Asian economies for the period 1995-2012. We evaluate the accuracy of the forecasts, and test for unbiasedness and e ciency. Our results show that (i) forecast errors are large for most of the countries, but large di erences exist between countries; (ii) forecasts improve slowly passing from long to short horizon, which contributes to explain the magnitude of forecast errors; (iii) GDP growth forecasts underreact to economic news but in ation forecasts are mostly e cient; (iv) the size or oiaseds-24p2fwiTJ/F17

1 Introduction

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

In this paper, we use the *Asian-Pacific Consensus Forecasts* to provide a first comprehensive evaluation of the macroeconomic forecasts 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 GDP growth and inflation forecasts, two key variables for macroeconomic analysis (see Golinelli and Parigi, 2008; Costantini and Kunst, 2011; Golinelli and Parigi, 2014).

Several studies have found differences in forecast performance between advanced and emerging economies, especially in terms of accuracy, information rigidities and efficient use of information (Loungani, 2001; Loungani et al., 2013; Dovern et al., 2015). After several decades of fast growth, some Asian economies have recently acquired the status of advanced economies, while some others are still emerging but growing rapidly. In this respect, it is worth investigating the performance of forecasts in these newly-advanced economies and compare them with those observed in previous studies for advanced and emerging countries. In addition, it is also impor-

[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 directional forecast accuracy.

tant to examine whether progress has been made in forecast performance over the years, since economies of many countries have transitioned from low/middle income to middle/high income.

Another aspect of Asian economies is that they have experienced economic fluctuations of large magnitude: while recessions tended to be more severe and longer-lasting than those in developed countries (Hong et al., 2010), sharp economic recoveries have also occurred. Furthermore, Asia has made remarkable progress in fighting against inflation (Filardo and Genberg, 2010), and it is interesting to examine how forecasters performed 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, it may be also important to correctly predict the direction of change of crucial variables. This is the case for GDP growth and inflation which are the most important macroeconomic goals for policy makers (a central banks can increase/decrease the interest rate if the inflation rises/decreases to stabilize the economy). To test for forecast unbiasedness and efficiency, we use the econometric approach developed by Davies and Lahiri, (1995) and later extended by Clements et al. (2007), Ager et al. (2009) and Dovern and Weisser (2011). We choose to analyze individual forecasts rather than consensus forecasts so as to shed light on individual heterogeneity across the forecasters and avoid any problem of aggregation bias.

It should be noticed that Loungani (2001), Loungani et al. (2013) and Dovern et al. (2015) use a larger data set which includes ours. However, our paper differs in several respects. First, they do not analyze inflation forecasts. Second, we focus on individual countries where those studies 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 address some other issues such as directional accuracy, long-term predictability, and acquisition

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

di erent versions of their name. For instance, the labels Citigroup and SSB Citibank refer to the same forecast institution. It is therefore essential to carefully clean the data and allocate the

2009 to 14.7% in 2010, and inflation in China fell from 17.1% to 8.4% between 1995 and 1996.

[Insert Figure 1]

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 forecast accuracy using the root mean squared error. We define $RMSE_{i;h} = \sqrt{T^{-1}\sum_{t=1} e^2_{i;t;h}}$ as the RMSE for forecaster i at horizon h and $RMSE_h = \frac{1}{N}\sum_{i=1} RMSE_{i;h}$ as the average of the individual RMSEs at horizon h. In Table 1, we report the $RMSE_h$ for selected forecast horizons. Similar to previous studies (see e.g. Lahiri and Sheng 2010), we find that 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 fully absorbed during the current year, with no potential impact on growth and inflation in the next year. After approximately the first 10 months (i.e. $h < 14$), forecasts become increasingly accurate as the horizon shortens, and information about the actual value accumulates.

Forecast errors vary considerably across countries, especially at long and middle horizons. For instance, when GDP growth forecasts are considered, the $RMSE_{12}$ (i.e. the RMSE for January of the year to be forecasted) is much higher in Singapore (3.55) and Malaysia (3.23) than in China (1.13) and India (1.70). Disparities are even wider for inflation, e.g. the $RMSE_{12}$ is equal to 8.63 for Indonesia and 0.50 for Japan. In most of the 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 and 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 forecasts of inflation in Japan. On average, forecasts for the advanced economies (Japan, Taiwan, Hong Kong, Singapore and Korea) are not more accurate than those of emerging economies (China, India, Indonesia, Malaysia, and Thailand). It should be noticed that these findings are not driven by outliers (i.e. forecasters with extremely high RMSE). For instance, using the median of individual RMSE rather than the mean would provide almost exactly the same results.

[Insert Table 1]

Table 1 also shows that the RMSE for inflation is lower than that for the GDP growth for most of the countries. This result, which has previously been reported for developed economies (e.g. Harvey et al. 2001), underscores the fact that actual inflation is easier to predict. One possible reason is that inflation is more stable than GDP growth. The reverse is however observed in China, India and Indonesia. Output in China has traditionally been relatively simple to forecast due to government control over the economic activity and its ability to meet growth targets. In India and Indonesia, inflation shocks have been rather large (it sometimes exceeds 10%), and inflation is difficult to predict compared to stable growth.

The comparison of absolute RMSE shows that GDP growth and inflation are more difficult to forecast in some countries than in others. However, it would be misleading to associate low RMSE with high forecast ability, 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 and 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 naive forecast of horizon 24. Predictability naturally increases moving from long to short horizons, and typically approaches 95%-100% at short horizons.

Figure 2 shows that 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

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

nd that for many countries predictability remains at zero until late in the forecasting cycle, in particular for GDP growth. For instance, for the GDP growth of Malaysia, $p_{h;24}$ only turns calculated as follows

$$RMSE^{adj}_{i,h} = \sqrt{T^{-1} \sum_{t=1}^{T} \tilde{e}^{2}_{i,t,h}} \tag{1}$$

with

$$\tilde{e}^{2}_{i,t,h} = e^{2}_{i,t,h} \frac{median_t(median_i(|e_{i,t,h}|))}{median_i(|e_{i,t,h}|)} \tag{2}$$

where $median_i$ is the cross-section median and $median_t$ is the median over t. Therefore, if the forecast errors are large at horizon h and year t compared with forecast errors for the same horizon but other t, then the weight $\frac{median_t(median_i(|e_{i,t,h}|))}{median_i(|e_{i,t,h}|)} < 1$ and the squared errors will be reduced. Note that the medianihr83-(usd)-2hethatn-345(the)-345(medn)-3hetho-345(tlessen-345(mhe)-345(r

3.3 Forecast errors over the horizons

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

We fit a non-parametric curve and find an inverted-L shape relationship for both GDP growth and inflation forecasts. Information gains are initially inexistent, but then gradually increase and peak at middle horizons as the economic news become increasingly informative. At short horizons, information gains remain remarkably high, especially for GDP growth and, to lesser extent, for inflation. These results contrast with those in Isiklar and Lahiri (2007), who find an inverted U-shape for advanced economies, and imply that forecasts in Asia improve relatively slowly. Large forecast errors in Asia may be also due to this. A possible explanation for this difference is that economic indicators in many Asian countries, including China and India (see Nilson and Brunet, 2006; Dovern et al., 2015) are often not as informative of growth as in countries such as the United States. Fewer quality indicators are available, which is expected to delay the acquisition of information. Consequently, it may take longer for forecasters to form accurate expectations about GDP growth. Thailand and Taiwan are two examples of countries where panelists keep making large forecast revisions for GDP growth even at the later stages the be qualitatively the same if other horizons were selected). It emerges that forecast errors are considerably higher during recessions years than during calm periods. For most of the countries, forecast errors increased sharply during the 1998 Asian crisis, before settling to low levels during the 2000-2007 calm period. Forecast errors increased again in 2008 and 2009, before starting to decline from 2010. China and India are two exceptions: forecast errors are less cyclical due to a stable economic growth and absence of recessions. Interestingly, there is no evidence that forecasts in Asia have become more accurate over time. For instance, the RMSE over period 2010-2012 is not lower than it was during the 1995-1997 and 2000-2007 periods for most of the countries.

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

[Insert Figure 4]

4 Testing forecast unbiasedness

In this section we test forecast unbiasedness. In order to do so, we use the error decomposition model initially proposed by Davies and Lahiri (1995) and later extended 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 forecast errors $e_{i;t;h}$, the difference between the actual value 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 ϕ_i captures a forecaster-specific bias, $\lambda_{t;h}$ represents the effects of unanticipated macroeconomic shocks occurring between the time 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 variance σ_u^2 and $\varepsilon_{i;t;h} = \sum_{k=1}^{h} \xi_{i;t;k}$, where $\xi_{i;t;k}$ has zero mean and variance σ_i^2 (see Deschamps and Ioannidis, 2013). We estimate the three components of the error model (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 model (3); $\phi_i > 0$ and $\phi_i < 0$ indicate forecast underestimation and overestimation, respectively. A simple OLS regression of forecast errors on a constant delivers a consistent estimate of the bias ϕ_i:

provide a 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.

[Insert Table 3]

It shows that the magnitude of the mean forecast errors is typically larger at long horizons than at short horizons. Intuitively, mean forecast errors are small at short horizons due to superior information. In spite of these differences, it is worthwhile to estimate the overall bias to assess the general tendency to over-/underpredict growth and inflation. Table 4 summarizes the results pooled over all the horizons (see equation 3). For growth forecasts, the hypothesis of unbiasedness can only be rejected for China (0.33 percentage point), 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 forecasters failed to predict. On the contrary, forecasts for China underpredict growth, indicating that China's strong growth over the past two decades has been unanticipated. 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 forecast errors leads to large standard errors. Overall, our analysis reveals differences in growth forecast biases between countries, both in terms of direction and magnitude. Nonetheless, forecast biases are shocks.[6] As a result, forecasts typically underpredict GDP during years of rapid growth and overpredict during recession years. For instance, forecasters have been 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 be observed for the 1998 Asian crisis. A similar pattern is observed for inflation: forecasters failed to predict unusual events such as 60% inflation in Indonesia in 1998, resulting in large forecast biases during those years.

[Insert Table 4]

5 Testing forecast efficiency

In this section we test for weak form efficiency (see Nordhaus, 1987). The forecasts are efficient when they incorporate all the past available information.[7] Nordhaus proposes a test 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 β_i=0 in the following regression of the forecast revisions on their lagged value:

$$r_{i;t;h} = \beta_i r_{i;t;h-1}$$

$$Cov(\ _{i;t_1;h_1};\ \ _{j;t_2;h_2}) = Cov(u_{t_1;h_1+1} + \ \ _{i;t_1;h_1+1}; u_{t_2;h_2+1} + \ \ _{j;t_2;h_2+1}) \quad (9)$$

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

[Insert Table 5]

Table 5 reports the efficiency test results. When considering the forecasts of GDP growth, the hypothesis of efficiency can be rejected for eight countries (at 1% significance level for six countries and at 10% significance level for two countries). The estimates of are positive for all the countries, indicating a general tendency to underreact to new information. However, these values are not larger than those reported in previous studies for developed economies (see for example Lahiri and Sheng, 2008). This indicates that the volatile macroeconomic environment in Asia does not seem to affect forecasters' ability, or willingness, to efficiently incorporate new information. However, at individual forecaster level, forecast efficiency can be rejected at the 5% level only for a small number of individual forecasters (35 out of 175). Among those 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 update of forecasters' information sets (i.e. "sticky information model"), as well as by the existence of noisy signals ("noisy information model"). However, the finding that individual forecast revisions are autocorrelated is not predicted by either of these two models. As long as forecasters place the optimal weight on new information (see e.g. Lahiri and Sheng, 2008), individual forecast revisions should be unpredictable. In other words, evidence that $_i > 0$ shows that there is more stickiness in the forecasts than what would be predicted by noisy information models.

The finding of forecast underreaction can be explained by behavioral aspects. Ehrbeck and Waldmann (1996) argue that forecasters may not care about accuracy per se, but rather seek to mimic the forecasting pattern of well-informed forecasters in order to enhance their own reputation. In this setting, they show that forecasters may be unwilling to make large forecast revisions because large revisions signal that previous forecasts were wrong. Therefore, forecasters are expected to insufficiently adjust forecasts upon the arrival of new information. This circumstance 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 forecasters who frequently change their forecasts may be perceived as erratic by their clients. As a result, forecasters may strategically choose to under-react to new information. Another possible explanation is 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 behavior, forecasts will be gradually rather than immediately adjusted to new information, 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 on GDP growth,

growth forecasts. Compared to previous analysis for developed countries (see for example Dovern and Weisser, 2011), no strong evidence against the efficiency of forecasts for inflation in 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 Asia forecasts. In this section, we discuss the role played by forecast under-/overreaction and systematic biases in explaining the high RMSE. In general, forecast under-/overreaction is expected to have an adverse effect on forecast accuracy. Forecast errors tend to be larger than those obtained when 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 performance of the forecasts in terms of RMSE is due to inefficient use of information. For GDP growth, the degree of underreaction is also low (maximum of 0.16 for Taiwan and cross-country average of 0.09) and it is 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 RMSE in Asia cannot be explained by the inefficient use of information. To further investigate this issue, we also compute the cross-country correlation between the RMSE and the estimated β. 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 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 $\tilde{f}_{i,t,h} = f_{i,t,h} + \hat{\alpha}_{i,h}$, where $\hat{\alpha}_{i,h} = \frac{1}{T}\sum_{t=1}^{T}(A_t - f_{i,t,h})$ is the forecaster- and horizon-specific bias. We denote by $\widetilde{RMSE}_h$ the mean of the individual RMSE for the bias-adjusted forecasts[8] and we expect that $\widetilde{RMSE}_h < RMSE_h$.

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

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

[Insert Table 6]

Overall, we argue that biases and forecast underreaction do not seem to explain much of the poor performance of forecasts in Asia. The performance of the forecasts would remain poor, and 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 accurately forecast the direction of the change is particularly important for investors and policymakers (Blaskowitz and Herwatz, 2009, 2011, 2014; Altavilla and De Grauwe, 2010; Bergmeir et al., 2014). For investors, an investment decision driven by a specific macroeconomic forecast with a small forecast error may not necessarily be as profitable as an investment decision guided by an accurate prediction of the direction of change. For policymakers, directional predictions are crucial to adjust policy instruments as to increase or decrease interest rates (Öller and Barot, 2000).

In this section, we analyse the directional accuracy of the professional forecasts in Asia. To

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

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

To further eeDA

turns out that these very low values of DA are observed during years of positive change, and this explains why DA is lower for accelerations. In all those cases, the low value of the DA for that year was preceded by another acceleration and forecasters usually failed to predict the second acceleration. For instance, in 2003 and 2004 GDP growth accelerates in China and panelists were surprised until 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 for several countries we find that positive changes are correctly predicted more often than negative changes. This finding also reflects the fact that Asia has made great progress in fighting against inflation (see Filardo and Genberg, 2010) and forecasters have regularly failed to anticipate inflation slowdowns, resulting in relatively low DA for negative changes. Interestingly, those 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 is 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 which performs well in terms of DA does not necessarily perform well in terms of RMSE, and vice-versa. For GDP growth, for example, China ranks first in terms of RMSE, but shows the worst result for DA, whereas Indonesia does the opposite for inflation. For some other countries, the forecast performance is 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 performance 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 RMSE and a directional forecast accuracy measure, and tested for unbiasedness and efficiency. The results are as follows. First, forecast errors are large for most of the countries, but the forecasts are nonetheless directionally accurate. Large disparities in the magnitude of forecast errors (and long-term predictability) are also observed across countries, for both GDP growth and inflation. For most of the countries, forecast accuracy is higher for inflation than for growth, which underscores that inflation is intrinsically easier to predict. Further, the accuracy of the forecasts in Asia improve relatively slowly from long to short horizons. This result may also contribute to explain the high RMSE. Second, the hypothesis of unbiasedness cannot be rejected for the majority of the countries. However, inflation forecasts show a tendency to overpredict, which may be caused by the decline of inflation in Asia. Finally, the hypothesis that forecasters incorporate new information efficiently is widely rejected for the forecasts of GDP growth, indicating a tendency to underreact, whereas for inflation we find little evidence of information stickiness.

This paper also contributes to the literature on the forecasting performance across advanced and emerging economies. Our results show that there is no correlation between forecast accuracy (and predictability) and the degree of economic development. Yet, unlike previous studies, we surprisingly find that underreaction for the 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 performance 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 target revised figures or the initial announcement, and it is important to verify that our main are robust to using revised figures. Starting with inflation, revised and initial IMF figures are actually extremely close. The mean absolute difference between initial and revised inflation estimates is less than 0.1%, with the exception of Indonesia (0.3%). None of the 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 upwards 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% in Korea to 1.2% in Singapore. Using revised figures as the benchmark, estimated RMSEs are mostly unaffected except for China, where RMSE would almost double. In general, RMSEs are smaller 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

[1] 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.

[2] Altavilla, C., De Grauwe, P. 2010. Forecasting and combining competing models of exchange rate determination. *Applied Economics*, 42, 3455-3480.

[3] Ashiya, M. 2005. Twenty-two years of Japanese institutional forecasts. *Applied Financial Economics Letters*, 12, 79-84.

[4] Batchelor, R., Dua, P. 1992. Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, 11, 169-181.

[5] 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.

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

[7] Blaskowitz, O., Herwartz, H. 2011. On economic evaluation of directional forecasts. *International Journal of Forecasting*, 27, 1058-1065.

[8] Blaskowitz, O., Herwartz, H. 2014. Testing the value of directional forecasts in the presence of serial correlation. *International Journal of Forecasting* 30, 30-42.

[9] Capistran, C., Lopez-Moctezuma, G. 2014. Forecast revisions of Mexican in ation and GDP growth. *International Journal of Forecasting*, 30, 177-191.

[10] 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.

[11] Clements, M.P., Taylor, N. 2001. Robust evaluation of xed-event forecast rationality. *Journal of Forecasting*, 20, 285-295.

[12] 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.

[13] Clements, M.P. 2014. Forecast Uncertainty | Ex Ante and Ex Post: U.S. In ation and Output Growth. *Journal of Business and Economics Statistics*, 32(2), 206-216.

[14] Coibion, O. and Gorodnichenko, Y., 2012. What Can Survey Forecasts Tell Us about Information Rigidities?, *Journal of Political Economy*, 120(1), 116 { 159.

[15] Costantini M., Kunst, R. M. 2011. Combining forecasts based on multiple encompassing tests in a macroeconomic core system. *Journal of Forecasting*, 30, 579-596.

[20] 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.

[21] 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.

[22] Ehrbeck, T., Waldmann, R. 1996. Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics*, 111, 21-41.

[23] Filardo, A., Genberg H. 2010. Targeting inflation in Asia and the Pacific: Lessons from the recent past. *The international financial crisis and policy challenges in Asia and the Pacific*. BIS Papers, 52.

[24] 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.

[25] Golinelli, R., Parigi, G. 2014. Tracking world trade and GDP in real time. *International Journal of Forecasting*, 30, 847-862.

[26] Harvey, D. Leybourne, S., Newbold, P. 2001. Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal*, 4, 37-55.

[27] Hong, K., Lee, J., Tang, H. 2010. Crises in Asia: Historical perspectives and implications. *Journal of Asian Economies*, 21, 265-279.

[28] 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.

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

[30] Krkoska, L., Teksoz, U. 2009. How reliable are forecasts of GDP growth and inflation for countries with limited coverage? *Economic Systems*, 33, 376-388.

[31] Lahiri, K., and Isiklar, G. 2009. Estimating International Transmission of Shocks Using GDP Forecasts: India and Its Trading Partners. In *Development Macroeconomics, Essays in Memory of Anita Ghatak* (Eds. S. Ghatak and P. Levine), Routledge.

[32] Lahiri, K., Sheng, X. 2008. Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, 144, 325-340.

[33] Lahiri, K., Sheng, X. 2010. Learning and Heterogeneity in GDP and Inflation Forecasts. *International Journal of Forecasting*, 26, 265-292.

[34] 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.

[35] Loungani, P., Steckler, H. and Tamirisa, N. 2013. Information rigidity in growth forecasts: Some cross-country evidence. *International Journal of Forecasting*, 29, 605-621.

[36] Nilsson, R. and O. Brunet, 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.

[37] Nordhaus W. 1987. Forecasting efficiency: concepts and applications. *Review of Economics and Statistics*, 69, 667-674.

[38] Öller, L, Barot, B. 2000. The accuracy of European growth and inflation forecasts. *International Journal of Forecasting*, 16, 293-315.

[39] Ottaviani M., Sorensen, P. 2006. The Strategy of Professional Forecasting. *Journal of Financial Economics*, 81, 441-466.

Table 1: Root mean squared error averaged across forecasters.

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
In ation										
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			In ation			
		$_i > 0$	$_i < 0$		$_i > 0$	$_i < 0$	No. forecasters
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: indicates the bias parameter (see Section 4). $_i > 0$ and $_i < 0$ (see equation (3)) refers the number of forecasters with a positive (negative) bias at the 5% level. No. Forecasters denotes the number of forecasters. Standard errors are in parenthesis. ; and indicate the level of signi cance at 10%, 5% and 1%, respectively.

Table 5: Efficiency test results

	GDP			Inflation			
		$_i > 0$	$_i < 0$		$_i > 0$	$_i < 0$	No.forecasters
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: denotes the pooled estimates of equation (8). For the interpretation of $_i > 0$ and $_i > 0$, see Section 5. Standard errors are in parenthesis. No. Forecasters denotes the number of forecasters. ; and indicate the level of significance at 10%, 5% and 1%, respectively.

Table 6: Bias-adjusted RMSE averaged across forecasters

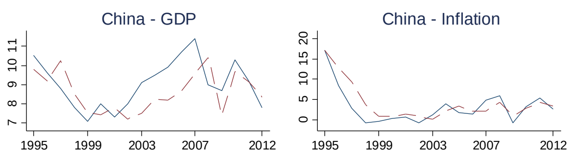
	China	Japan	Taiwan	Hong Kong	Korea	Singapore	India	Indonesia	Malaysia	Thailand
GDP										
h=1	0.33	0.42	0.61	0.52	0.55	0.59	0.65	0.51	0.38	0.82
h=12	0.98	1.80	2.32	2.92	2.30	3.34	1.61	2.35	3.05	3.12
h=24	1.47	2.30	2.85	3.85	3.54	4.08	1.71	3.28	3.50	4.02
Inflation										
h=1	0.34	0.11	0.27	0.28	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: This 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
$A_t > 0$										
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
$A_t < 0$										
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
$A_t > 0$										
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
$A_t < 0$										
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
$A_t > 0$	0.25	0.50	0.50	0.25	1.00	0.71	0.00	0.11	0.11	-0.43
$A_t < 0$	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
$A_t > 0$	0.11	0.11	0.00	0.11	0.20	0.00	0.11	0.00	0.20	0.27
$A_t < 0$	0.00	0.00	0.33	0.00	0.71	-0.43	0.25	-0.33	0.71	-0.33

Notes: Figures indicate the directional accuracy loss given by Equation (10).



.

Figure 1: Actual values (solid line) and consensus forecast at h=12 (dash line) for GDP growth and in ation.

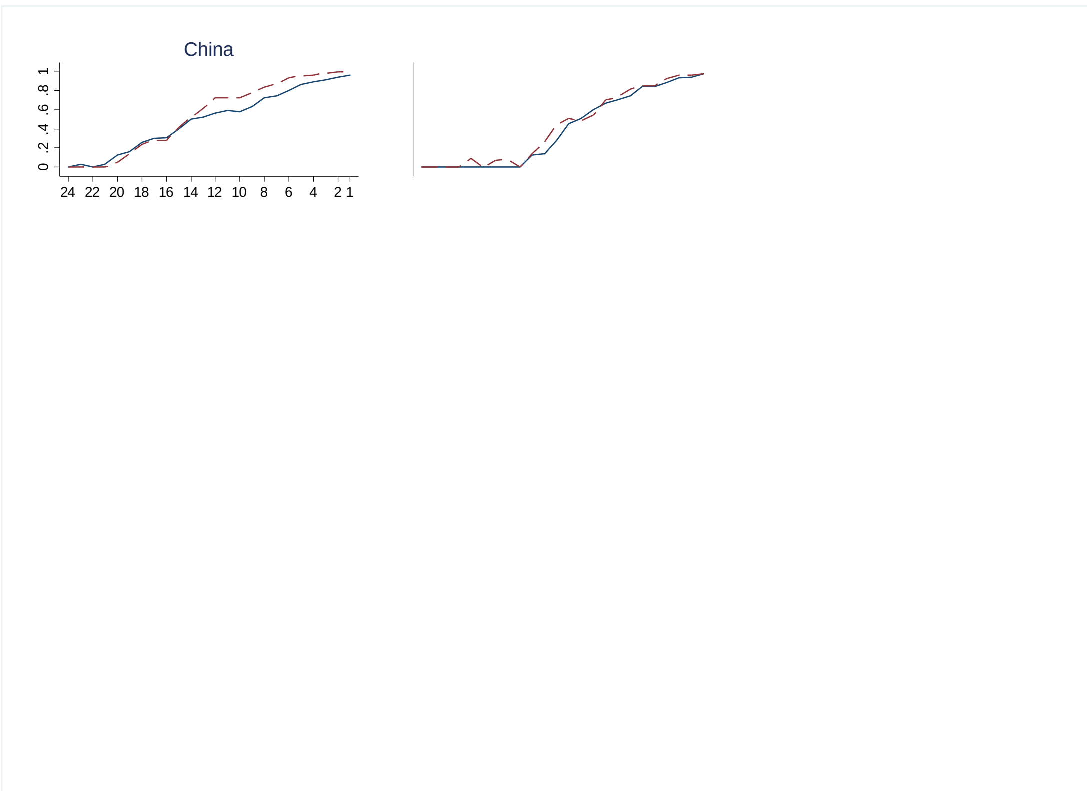


Figure 2: Predictability of GDP growth (solid line) and in ation (dash line). Diebold-Kilian statistics.

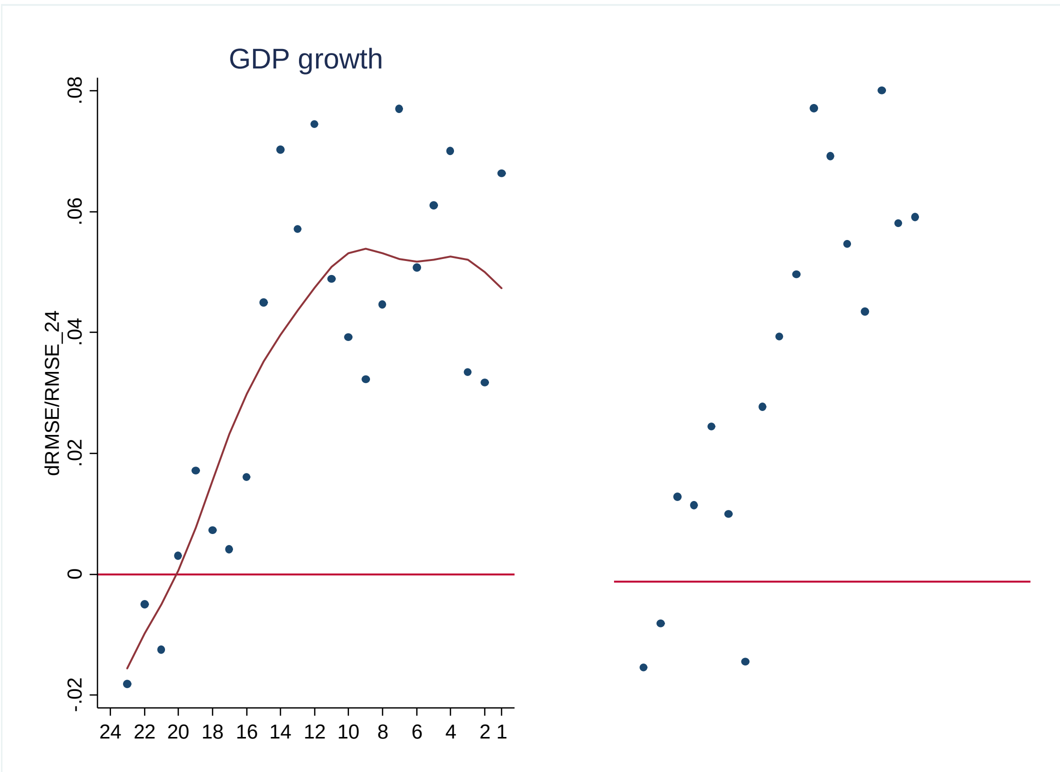


Figure 3: Changes in RMSE between two consecutive horizons, averaged across forecasters.

Figure 4: RMSE of GDP growth forecasts (solid line, left scale) and in ation forecasts for horizon 12.