Combining Non-Replicable Forecasts*

Chia-Lin Chang

Department of Applied Economics
National Chung Hsing University
Taichung, Taiwan

Philip Hans Franses

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
The Netherlands

Michael McAleer

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
The Netherlands
and
Tinbergen Institute
The Netherlands
and
Department of Economics and Finance
University of Canterbury
New Zealand

EI 2010-44

Revised: May 2010

* For financial support, the first author wishes to thank the National Science Council, Taiwan, and the third author wishes to thank the Australian Research Council, National Science Council, Taiwan, Center for International Research on the Japanese Economy (CIRJE), Faculty of Economics, University of Tokyo, and a Visiting Erskine Fellowship, College of Business and Economics, University of Canterbury.

Abstract

Macro-economic forecasts are often based on the interaction between econometric models and experts. A forecast that is based only on an econometric model is replicable and may be unbiased, whereas a forecast that is not based only on an econometric model, but also incorporates an expert's touch, is non-replicable and is typically biased. In this paper we propose a methodology to analyze the qualities of combined non-replicable forecasts. One part of the methodology seeks to retrieve a replicable component from the non-replicable forecasts, and compares this component against the actual data. A second part modifies the estimation routine due to the assumption that the difference between a replicable and a non-replicable forecast involves a measurement error. An empirical example to forecast economic fundamentals for Taiwan shows the relevance of the methodological approach.

Key words: Combined forecasts, efficient estimation, generated regressors, replicable forecasts, non-replicable forecasts, expert's intuition.

JEL Classifications: C53, C22, E27, E37.

1. Introduction

Econometric models are frequently used to provide base-level forecasts in macro-economics. Usually, these model-based forecasts are adjusted by experts who have domain knowledge. For example, Franses, Kranendonk and Lanser (2010) document that this holds for all forecasts (for GDP, inflation, and so on) generated from the large macro-economic model created at the CPB Netherlands Bureau for Economic Policy Analysis. The difference between the pure model-based forecast and the final forecast is often called expertise, intuition or judgment. It is a trade secret owned by a forecaster, as it is rarely written down, but it can have significant value in forecasting key economic fundamentals. A forecast that is based on an econometric model is replicable and may be unbiased, whereas a forecast that is not based on an econometric model is non-replicable and is typically biased. In practice, most macro-economic forecasts (from CPB, but also from the FED, the World Bank, OECD and IMF) are non-replicable forecasts.

In this paper we examine the evaluation of the quality of a range of available non-replicable forecasts with a specific focus on the combinations of these potentially biased forecasts. For this, we propose a methodology that approaches this issue from two different angles. The first aims to de-bias the non-replicable forecast by retrieving and comparing their replicable components. The second modifies the estimation method.

In order to illustrate, we use data from Taiwan for three reasons. First, a consistent data set is available for the government and two professional quarterly forecasts of economic fundamentals over an extended period. Second, no previous comparison seems to have been made of the competing combined forecasts. Third, there does not seem to have been any comparison of individual and combined forecasts based on an optimal subset of the multiple forecasts.

The plan of the remainder of the paper is a follows. Section 2 presents the econometric model specification, analyses replicable and non-replicable forecasts, considers optimal forecasts and efficient estimation methods, compares individual replicable forecasts with an optimal subset combined replicable forecast, and presents a direct test of forecasting expertise. The data analysis and a relevant empirical example of multiple forecasts of economic fundamentals for Taiwan are discussed in Section 3. Some concluding comments are given in Section 4.

2. Model Specification

In this section we present an econometric model for both government forecasts and multiple professional forecasts, where this setting is chosen as it matches the empirical data that are available. This will enable the generation of replicable forecasts, permit a comparison to be made with non-replicable forecasts, enable the government forecasts to be compared with multiple professional forecasts, lead to an optimal subset combined forecast, and provide a direct test of forecasting expertise.

2.1. Individual Forecasts

Let the econometric model for forecast *i* be given as

$$y = Z\gamma + \beta_i X_i^* + u_i, \quad u_i \sim (0, \sigma_i^2 I), \qquad (1)$$

where $i = 1,\ldots,m$, y is a $T \times 1$ vector of observations to be explained (typically, an economic fundamental, such as the inflation rate or the real GDP growth rate), Z is a $T \times g$ matrix of T observations on g variables that is public information, and hence is known to the government and (m-1) multiple professional forecasters and to the analyst. The X_i^* is the latent expertise of forecaster *i*, which is not observed by the analyst or by any of the forecasters. The assumptions on the error term in (1) can be relaxed easily, and it is also assumed that $E(Z' u_i) = 0$ and $E(X_i^{*\prime} u_i) = 0$.

If X_i^* were observable data, the OLS estimates of the parameters in (1) would be consistent and efficient. Under the assumption of correct specification and a mean squared error (MSE) loss function, the optimal forecast of *y*, given the information set, is its conditional expectation (see Patton and Timmermann (2007a, 2007b)).

Let the $T \times 1$ vector, X_i, represent the non-replicable forecast of forecaster *i*, which is observable for the analyst. A key notion (see also Franses et al. (2009)) is the assumption that the connection between X_i and the expertise of forecaster *i* , X_i^*, is assumed to be given by

$$X_i = X_i^* + \eta_i, \quad \eta_i \sim (0, \sigma_\eta^2 I) \qquad (2)$$

where i =1, …, m, X_i, X_i^* and η_i are T x 1 vectors, and η_i in (2) denotes the measurement error for forecast i. It is assumed that X_i^* and η_i are uncorrelated for all i =1,…,m.

Moreover, the non-replicable forecast can be modelled as

$$X_i = W_i\delta_i + \eta_i,\ \eta_i \sim (0, \sigma_i^2 I) \tag{3}$$

where the $T \times k_i$ matrix, W_i, denotes the measurable expertise of forecaster i at time t-1. It is assumed that $E(W_i\eta_i) = 0$ for all i, δ_i is a $k_i \times 1$ vector of unknown parameters, and

$$W_i \subset I_{-1}^i, \tag{4}$$

i =1,…,m, where I_{-1}^i is the information set of forecaster i at time t-1. As Z is public information, it follows from (4) that $\{Z, W_i\} \subset I_{-1}^i$, for all i =1,…,m. The information set I_{-1}^i is used to obtain optimal forecasts of y under a MSE loss function. It should be emphasized that an econometric model enables optimal forecasts to be generated, and hence the absence of an econometric model means that optimal forecasts under a MSE loss function cannot be obtained.

It follows from (3) and I_{-1}^i that

$$E(X_i | I_{-1}^i) = X_i^* = W_i\delta_i, \tag{5}$$

where the conditional expectation, $E(X_i | I_{-1}^i) = W_i\delta_i$ can be estimated by

$$\hat{X}_i^* = \hat{X}_i = W_i\hat{\delta}_i = W_i(W_i'W_i)^{-1}W_i'X_i = P_iX_i, \tag{6}$$

where $P_i = W_i(W_i'W_i)^{-1}W_i'$ is the standard 'hat' matrix. Equation (6) shows that the estimate of the latent expertise, X_i^*, is equivalent to the estimate of the non-replicable forecast, $\hat{X}_i$. In the context of equation (6), it is well known that the use of rational expectations reduces the number of unknowns in (5) from T to k_i, where $k_i \ll T$ for all i.

Replacing the latent X_i^* in (1) with the observable $\hat{X}_i$ gives

$$y = Z\gamma + \beta_i \hat{X}_i + \varepsilon_i \tag{7}$$

where

$$\begin{aligned} \varepsilon_i &= u_i + \beta_i\left(X_i^* - \hat{X}_i\right) \\ &= u_i + \beta_i(W_i\delta_i - P_iX_i) \\ &= u_i + \beta_i(W_i\delta_i - P_i(W_i\delta_i + \eta_i)) \\ \varepsilon_i &= u_i - \beta_i P_i \eta_i \end{aligned} \tag{8}$$

which is a composite error term, involving the measurement error, η_i, for forecast *i*.

If $\beta_i = 0$ for all *i*, such that forecaster *i* bases forecasts solely on public information rather than on intuition, then $\varepsilon_i = u_i$ for all *i*. However, if forecaster *i* does have intuition, and hence, X_i^* adds relevant information to *Z* when explaining *y*, for *i* =1,…,*m*, then there are *m* non-nested forecasting models in (7). These can be compared on the basis of standard forecasting criteria and/or can be tested using non-nested methods (for a detailed discussion see, for example, McAleer (1995)).

The correlation between $\hat{X}_i$ and ε_i is $-\beta_i\sigma_i^2(T - k_i)$, but OLS for the parameters in (7) is consistent as $\hat{X}_i$ is asymptotically uncorrelated with ε_i for all *i*.

If u_i and η_i are mutually uncorrelated, then

$$V = E(\varepsilon_i\varepsilon_i') = E(u_iu_i') + \beta_i^2 P_i E(\eta_i\eta_i')P_i$$

so that

$$V = \sigma_u^2 I + \beta_i^2 \sigma_i^2 P_i, \quad i = 1, \ldots, m . \tag{9}$$

It is obvious that serial correlation and heteroskedasticity are present in (9) through the measurement error, η_i, in X_i in (2). Thus, if OLS is used to estimate (9), the correct covariance matrix in (9), or a consistent estimator thereof, such as the Newey-West HAC covariance matrix estimator, should be used.

The necessary and sufficient conditions for OLS to be efficient in the presence of serial correlation and heteroskedasticity are given in Kruskal's Theorem, of which a special case is the Gauss-Markov Theorem (see, for example, McAleer (1992), Fiebig et al. (1992), McAleer and McKenzie (1991), Franses et al. (2009), Chang et al. (2009)), and for any *i* are given by

(i) $VZ = ZA_1$, for some A_1;

(ii) $V\hat{X}_i = \hat{X}_i A_2$, for some A_2.

Condition (i) is satisfied if $Z \perp W_i$ or if $Z \subset W_i$, while condition (ii) is satisfied automatically as $\hat{X}_i = P_i X_i$ in (6). In short, GLS is equivalent to OLS if the first step of the two step OLS estimator is satisfied as the transformation matrix will be proportional to the data matrix.

Defining $G_i = [Z_i \;\vdots\; \hat{X}_i]$ and $\phi'_i = (\gamma'_i \quad \beta'_i)$ for all *i*, (7) may be rewritten as

$$y = G_i \phi_i + \varepsilon_i. \tag{10}$$

If conditions (i) and (ii) are satisfied, OLS is efficient for ϕ_i and the correct OLS covariance matrix is given by

$$Var(\hat{\phi}_i) = (G'_i G_i)^{-1} G'_i V G_i (G'_i G_i)^{-1}, \tag{11}$$

where *V* is given in (9). Substitution for *V* in (11) gives

$$Var(\hat{\phi}_i) = \sigma_u^2 (G'_i G_i)^{-1} + \beta_i^2 \sigma_i^2 (G'_i G_i)^{-1} G'_i P_i G_i (G'_i G_i)^{-1}, \tag{12}$$

which shows that the standard OLS covariance matrix of $\hat{\phi}_i$, namely $\sigma_u^2 (G'_i G_i)^{-1}$, gives a downward bias in the covariance matrix and an upward bias in the corresponding t-ratios (see Pagan (1984) and Oxley and McAleer (1993) for examples in the case of generated regressors).

An alternative to estimating equation (7), which is the second part of our methodology, is to substitute from (2) into (1) to obtain

$$y = Z\gamma + \beta_i(X_i - \eta_i) + u_i$$

$$y = Z\gamma + \beta_i X_i + (u_i - \beta_i \eta_i). \qquad (13)$$

It is clear that OLS is inconsistent for (13) as X_i is correlated with η_i. Therefore, GMM should be used if the non-replicable forecast, X_i, is used to explain the variable of interest, *y*. Moreover, as X_i is not the conditional mean of *y*, it is not an optimal forecast under a MSE loss function. Indeed, X_i amounts to a biased forecast.

The effect of W_i, on the non-replicable forecast, X_i, for forecaster *i*, can be tested directly in (3):

$$X_i = W_i \delta_i + \eta_i, \ \eta_i \sim (0, \sigma_i^2 I),$$

in which OLS is efficient given the information set. Moreover, the conditional expectation of X_i is an optimal forecast under a MSE loss function.

2.2. Combined Forecasts

An alternative to evaluating the *m* forecasts individually is to combine the government and (*m*-1) professional forecasts into a combined forecast, namely:

$$y = Z\gamma + \beta_i \sum_{i=1}^{m} \lambda_i X_i^* + u \qquad (14)$$

where λ_i is a known constant, for *i* =1,…,*m*, and sum to unity. As X_i^* is not observed in (14), it can be replaced by $\widehat{X}_i$ from (6) to give

$$y = Z\gamma + \beta \sum_{i=1}^{m} \lambda_i \widehat{X}_i + \varepsilon \qquad (15)$$

where

$$\varepsilon = u + \beta \sum_{i=1}^{m} \lambda_i (X_i^? - \hat{X}_i) \quad (16)$$

If $\lambda_i = 1/m$ for all i, then $\sum_{i=1}^{m} \lambda_i \hat{X}_i$ in (15) is the mean of the m forecasts, which is a popular and frequently reported combined forecast. If $\lambda_i = 0$ for two values of $i = 1,\ldots,m$, that correspond with the minimum and maximum values of $\hat{X}_i$, with the remaining constants being set to $1/(m-2)$, then $\sum_{i=1}^{m} \lambda_i \hat{X}_i$ in (15) would correspond to a trimmed mean. If the $\hat{X}_i$ are ranked in increasing order, then setting $\lambda_m = 1$, $\lambda_1 = -1$ and $\lambda_i = 0$ for $i = 2,3,\ldots,m-1$ would give the range as the weighted sum.

It would be possible to replace the mean or trimmed mean by a median or modal forecast, but this would not be consistent with the purpose of the paper, as the median and mode are not based on replicable models. A similar comment applies to the use of the range as the weighted sum. . If the conditional mean of y is not given by the linear combination, $\sum_{i=1}^{m} \lambda_i \hat{X}_i$, with or without any of the weights being set to zero, then the linear combination is not an optimal forecast under a MSE loss function.

If the λ_i in (14) are unknown parameters for all i, they would have to be estimated. Although the $\hat{X}_i$ are likely to be highly correlated, for at least some i, the OLS estimates of λ_i would lead to an optimal combined forecast in the sense of minimizing MSE. If the statistically insignificant estimates of λ_i were set to zero, this would yield an optimal subset combined forecast.

The composite error in (16) can be rewritten as

$$\varepsilon = u + \beta \sum_{i=1}^{m} \lambda_i (W_i \delta_i - P_i X_i)$$

or equivalently

$$\varepsilon = u - \beta \sum_{i=1}^{m} \lambda_i P_i \eta_i \quad . \tag{17}$$

The covariance matrix of ε is given by

$$V = E(\varepsilon\varepsilon') = \sigma_u^2 I + \beta^2 \sum_{i=1}^{m} \lambda_i^2 \sigma_i^2 P_i \tag{18}$$

if u and η_i are uncorrelated for all i =1,…,m. .

Necessary and sufficient conditions for OLS in (15) to be efficient are given by:

(iii) $VZ = ZA_1$, for some A_1;

(iv) $V\left(\sum_{i=1}^{m} \lambda_i X_i\right) = \left(\sum_{i=1}^{m} \lambda_i X_i\right) A_2$, for some A_2.

Condition (iii) is satisfied if $Z \perp W_i$ for all for i =1, …, m,. or $Z \subset W_i$ for all for i =1, …, m, while condition (iv) is satisfied if $W_i \perp W_j$ for all for i, j =1, …, m, or if $W_i = W_j$ for all i, j =1, …, m. It is straightforward for condition (iii) to be satisfied by defining Z as a subset of W_i for all i =1,…,m.. However, it is unlikely that condition (iv) will be satisfied, especially for large m, as forecasters, by definition, differ in their expertise.

If OLS is used to estimate (15), the covariance matrix should be based on (18). Defining

$$H = \left[Z, \sum_{i=1}^{m} \lambda_i X_i \right]$$

and $\theta' = (\gamma', \beta)$, (15) may be rewritten as

$$y = H\theta + \varepsilon \tag{19}$$

so that the covariance matrix of $\hat{\theta}$ is given by

$$Var(\hat{\theta}) = (H'H)^{-1}H'VH(H'H)^{-1}. \tag{20}$$

Substitution of *V* from (18) into (20) gives

$$Var(\hat{\theta}) = \sigma_u^2(H'H)^{-1} + \beta^2(H'H)^{-1}H'\left(\sum_{i=1}^{m}\lambda_i\sigma_i^2 P_i\right)H(H'H)^{-1}, \tag{21}$$

which shows that, as in the case of (12), the standard OLS covariance matrix of $\hat{\theta}$, namely the first term on the right-hand side of (21), leads to a downward bias in the covariance matrix and an corresponding upward bias in the corresponding t-ratios. The covariance matrix in (21) can be consistently estimated by the Newey-West HAC covariance matrix. Smith and McAleer (1994) evaluate the finite sample properties of the HAC estimator for purposes of testing hypotheses and constructing confidence intervals in the case of generated regressors.

Again, an alternative combined forecast to (15) is to substitute from (2) into (14) to give

$$y = Z\gamma + \beta\sum_{i=1}^{m}\lambda_i(X_i - \eta_i) + u$$

$$y = Z\gamma + \beta\sum_{i=1}^{m}\lambda_i X_i + \left(u - \beta\sum_{i=1}^{m}\lambda_i\eta_i\right). \tag{22}$$

As in the previous discussion, λ_i may be known constants or unknown parameters. For estimation, as $\sum_{i=1}^{m}\lambda_i X_i$ is correlated with $\sum_{i=1}^{m}\lambda_i\eta_i$, GMM should be used rather than OLS to yield consistent estimators. Moreover, as the linear combination of the X_i in (22) is not the conditional mean of *y*, it is not an optimal forecast under a MSE loss function.

The individual and combined forecasting models given in (7) and (15), respectively, are non-nested, and hence may be tested against each other using a variety of non-nested tests. If the λ_i are known constants for all *i* =1, …, *m*, then the difference between (7) and (15) lies in the choice of whether an individual forecast, as given in (7), from *m* possible models is superior to the combined forecast, as given in (15). For purposes of statistical testing, the choice is one of whether $\hat{X}_i$ or $\sum_{i=1}^{m} \lambda_i \hat{X}_i$ is superior in forecasting *y* conditional on *Z*, such as comparing one forecast with the mean of the *m* forecasters. If the λ_i are unknown parameters, $\hat{X}_i$ would not be linearly independent of $\sum_{i=1}^{m} \lambda_i \hat{X}_i$, so that one of the *m* values of $\hat{X}_i$ would need to be omitted from $\sum_{i=1}^{m} \lambda_i \hat{X}_i$.

3. Data and Empirical Analysis

Since 1978, actual data and three sets of updated forecasts of the inflation rate and real GDP growth rate have been released by the Government of Taiwan (for further details, see Chang et al. (2009)). The unemployment rate is not regarded as a key economic fundamental in Taiwan. In this paper, we use the most recent revised government forecasts. The government forecasts (F1) and actual values of the inflation rate and real GDP growth rate are obtained from the Quarterly National Economic Trends, Directorate-General of Budget, Accounting and Statistics, Executive Yuan, Taiwan, 1980-2009. The forecasts from the two private forecasting institutions are obtained from the Chung-Hua Institution for Economic Research (F2) and Taiwan Institute of Economic Research (F3).

In addition to comparing actual data on both the inflation rate and real growth rate with three sets of forecasts, four combined forecasts are also considered, namely the mean of all three forecasts and three pairs of mean forecasts. In the Tables, M refers to the mean of all three forecasts, M12 refers to the mean of F1 and F2, M13 refers to the mean of F1 and F3, and M23 refers to the mean of F2 and F3.

As the actual values of the inflation rate and real GDP growth rate are available, the accuracy of the government and two private forecasts, as well as the effects of econometric model versus intuition, can be compared and tested. The sample period used for the actual values and the three sets of forecasts of seasonally unadjusted quarterly inflation rate and real growth rate of GDP is 1995Q3-2009Q2, for a total of 56 observations.

We have analyzed the data on unit roots and structural breaks. The diagnostics for unit roots (which are unreported) indicate that we can work with the growth rates data, as in Figures 1 and 2. Visual inspection from the same graphs does not suggest potential structural breaks, and there is also no evidence of structural breaks caused by any changes in measurement methods at the government agency and two private forecasting institutions in Taiwan.

The inflation rate and the three forecasts, F1, F2 and F3, are given in Figure 1, and the corresponding plots of the real GDP growth rate and the three forecasts are given in Figure 2. Figure 3 gives the inflation rate, the mean of the three forecasts, and the means of pairs of forecasts, while the corresponding plots of the real GDP growth rate, the mean of the three forecasts, and the means of pairs of forecasts are given in Figure 4.

Table 1 gives the correlations of the inflation rate, three forecasts, the mean of three forecasts, the means of pairs of forecasts (and their replicable counterparts, which are obtained from Tables 4 and 5 (to be discussed below) , with the corresponding plots of the real GDP growth rate given in Table 2. In these two tables, hats (circumflex) denote their replicable counterparts. In Tables 1 and 2, the highest correlations for both the actual inflation rate and the real GDP growth rate are with F1, followed by M13; for both variables, F1 is highly correlated with M12, M13 and M23, F2 is highly correlated with M12 and M23, F3 is highly correlated with M23, M is highly correlated with M12 and M13, M12 is highly correlated with M13, and M13 is highly correlated with M23. The correlations are generally higher between the original variables than between their fitted counterparts.

The goodness-of-fit measures, namely root mean square error (RMSE) and mean absolute deviation (MAD), of the replicable and non-replicable forecasts are given in Table 3 for both variables. For the non-replicable forecasts, in the upper panel of Table 3, the single forecast, F1, is best for both variables using RMSE and MAD, while the mean of two forecasts, M13, is second best for the inflation rate, and M12 is second best for the real GDP growth rate. A similar outcome holds for the replicable forecasts, with $\hat{F}1$ best for both variables using RMSE and MAD, while $\hat{M}13$ is second best for both variables using RMSE and MAD. These results suggest that, in general, the first single forecast is best in terms of both RMSE and MAD, followed by a mean combination of the first and third forecasts, for both the inflation rate and real GDP growth rate, regardless of whether a non-replicable or replicable forecast is used. Table 3 also shows that the biased non-replicable forecasts are apparently much more accurate than the replicable forecasts. Hence, the added intuition of the experts seems to lead to substantial improvement. This improvement is most evident for F1, where RMSE for the replicable forecast is abot twice as large as for the non-replicable forecast.

In Tables 4a-4b and 5a-5b, we report on the retrieval of a replicable part from the non-replicable forecasts based on public information for the inflation rate and real GDP growth rate, respectively. This public information is set at one-period lagged real growth, one-period lagged inflation, one period lagged forecast for forecaster 1, one period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3.

It is evident that the lagged values of the forecasts of all three forecasters are insignificant in all four tables, so the forecasters do not seem to include each other's predictions. The one-period lagged real GDP growth rate is significant for all seven forecasts for both the inflation rate and real GDP growth rate. Apart from the significant case of F1 in Table 4a, the one-period lagged inflation rate is not significant in capturing expertise for any of the seven forecasts for either variable. The F tests for the significance of the replicable part in Tables 4a-4b and 5a-5b indicate clearly that the expertise in equation (3) is captured by the one-period lagged variables, specifically the one-period lagged real GDP growth rate.

In order to examine if the replicable forecasts are unbiased, we consider equation (7) for three forecasts and four mean forecasts, which are given in Tables 6a-6b for the inflation rate and real GDP growth rate. As the replicable forecasts lead to generated regressors, the appropriate Newey-West HAC standard errors are calculated for valid inference. The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$ for i = 1,2,3. If the null hypothesis is not rejected, then the model via the replicable forecast can predict the actual value, whereas rejection of the null means that expert intuition could triumph over the model in case the non-replicable forecasts are not biased. Except for F1 and F2 for the real GDP growth rate in Table 6a, the null hypothesis is rejected in all cases, which makes it clear that intuition is significant in explaining actual values, and hence dominates the model. This supports the RMSE and MAD scores in Table 3.

Tables 7a-7b and 8a-8b focus on the accuracy of the non-replicable forecasts for three forecasts and four mean forecasts in equation (13) for the inflation rate and real GDP growth rate. As the non-replicable forecasts are correlated with the measurement errors, GMM is necessary for valid inference, where the instrument list for GMM for forecaster i includes one-period lagged real growth, one-period lagged inflation, one-period lagged forecast for forecaster 1, and one-period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$ for i = 1,2,3. Conditional on the information set, if the null hypothesis is not rejected, then the non-replicable forecast can accurately predict the actual value, whereas rejection of the null means means that the non-replicable forecast is biased.

Except for one case, namely GMM estimation of M for the inflation rate in Table 7b, the null hypothesis is rejected for all individual forecasts and mean forecasts. Thus, conditional on the information set, the non-replicable forecast cannot predict the actual inflation rate. Ignoring the OLS results in Tables 8a-8b, mirroring the results in Tables 7a-7b, except for one case, namely GMM estimation of F1 for the real GDP growth rate in Table 8a, the null hypothesis is rejected for all individual forecasts and mean forecasts. Thus, conditional on the information set, the non-replicable forecast cannot predict the actual real GDP growth rate. If we compare the F test values in Tables 7 and 8 with those in Table 6, we see that the non-replicable forecasts have greater bias than the replicable forecasts. Again, the non-replicable forecasts are much more accurate than the replicable forecasts, which means that the intuition of the forecasters greatly improves any model-based forecasts.

As in many other studies, combining forecasts can be beneficial. For inflation, we see that the GMM-based results in Table 7b indicate the M delivers unbiased forecasts. For GDP growth, matters are somewhat different. There we see that the non-replicable F1 is unbiased (Table 8a), and Table 3 also suggests it has the smallest forecast error. Table 8b clearly shows that combining forecasts is not sensible as all the combinations examined in Table 8b lead to biased forecasts.

4. Concluding Remarks

A forecast that is based on an econometric model is replicable and may be unbiased, whereas a forecast that is not based on an econometric model is non-replicable and is typically biased. Government and professional forecasters alike can, and do, provide both replicable and non-replicable forecasts. Both types of forecasts can be combined into a single combined forecast, such as a mean or trimmed mean forecast.

This paper developed a model to generate replicable forecasts by multiple professional forecasters, including the government, compared replicable and non-replicable forecasts using efficient estimation methods, and compared individual replicable forecasts with combined forecasts. An empirical example to forecast economic fundamentals for Taiwan showed the relevance of the methodological approach proposed in the paper. The empirical analysis showed that replicable and non-replicable forecasts could be distinctly different from each other, that efficient and inefficient estimation methods, as well as consistent and inconsistent covariance matrix estimates, could lead to significantly different outcomes, combined forecasts could yield different forecasts from their multiple individual components, and the relative importance of econometric model versus intuition could be evaluated in terms of forecasting performance.

It was shown that individual forecasts could perform quite differently from the mean forecasts of two or three individual forecasts, that intuition was significant in explaining actual values, and hence dominated the model, and that expert intuition that has been used to obtain the non-replicable forecasts of the inflation rate and real GDP growth rate was not sufficient to forecast accurately the actual values.

One of the major findings is that a proper analysis of combined forecasts could suggest a weaker dominance of other forecasts, as is typically documented in the literature. The GMM-based analysis shows that the combined forecasts could well be found to be biased, while the OLS-based analysis did not give any such warning signals.

References

Chang, C.-L., P.H. Franses and M. McAleer (2009), How accurate are government forecasts of economic fundamentals? The case of Taiwan, Available at SSRN: http://ssrn.com/abstract=1431007.

Fiebig, D.G., M. McAleer and R. Bartels (1992), Properties of ordinary least squares estimators in regression models with non-spherical disturbances, *Journal of Econometrics*, 54, 321-334.

Franses, P.H., H. Kranendonk, and D. Lanser (2010), One model and various experts: Evaluating Dutch macroeconomic forecasts, *International Journal of Forecasting*, to appear.

Franses, P.H., M. McAleer and R. Legerstee (2009), Expert opinion versus expertise in forecasting, *Statistica Neerlandica*, 63, 334-346.

McAleer, M. (1992), Efficient estimation: the Rao-Zyskind condition, Kruskal's theorem and ordinary least squares, *Economic Record*, 68, 65-72.

McAleer, M. (1995), The significance of testing empirical non-nested models, *Journal of Econometrics*, 67, 149-171.

McAleer, M. and C. McKenzie (1991), When are two step estimators efficient?, *Econometric Reviews*, 10, 235-252.

Oxley, L. and M. McAleer (1993), Econometric issues in macroeconomic models with generated regressors, *Journal of Economic Surveys*, 7, 1-40.

Pagan, A.R. (1984), Econometric issues in the analysis of regressions with generated regressors, *International Economic Review*, 25, 221-247.

Patton, A.J. and A. Timmermann (2007a), Properties of optimal forecasts under asymmetric loss and nonlinearity, *Journal of Econometrics*, 140, 884-918.

Patton, A.J. and A. Timmermann (2007b), Testing forecast optimality under unknown loss, *Journal of the American Statistical Association*, 102, 1172-1184.

Smith, J. and M. McAleer (1994), Newey-West covariance matrix estimates for models with generated regressors, *Applied Economics*, 26, 635-640.

Figure 1. Inflation Rate and Three Forecasts, 1995Q3-2009Q2

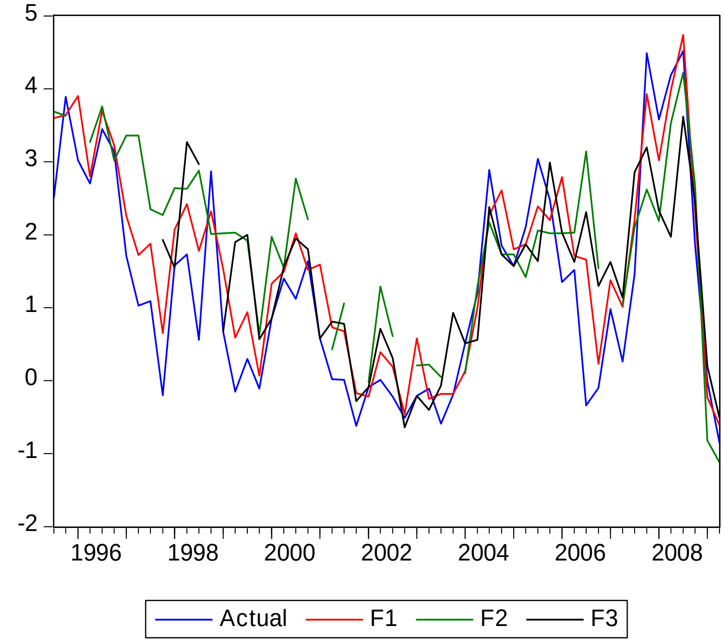

Figure 2. Real GDP Growth Rate and Three Forecasts, 1995Q3-2009Q2

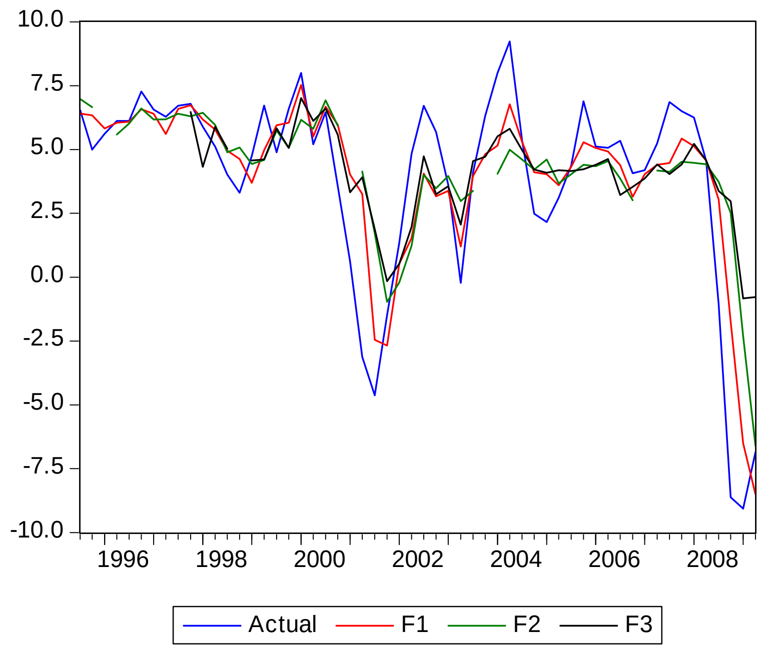

Figure 3. Inflation Rate, Mean of Three Forecasts, Means of Pairs of Forecasts, 1995Q3-2009Q2

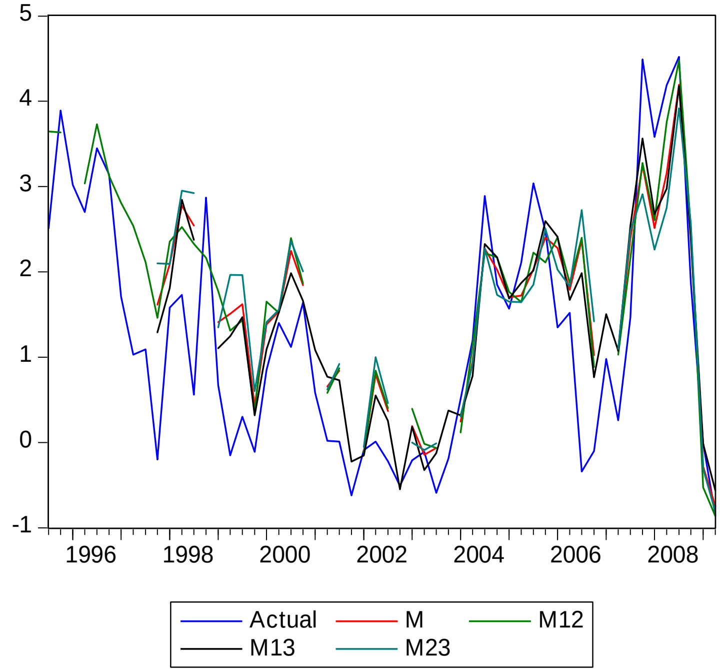

Figure 4. Real GDP Growth Rate, Mean of Three Forecasts, Means of Pairs of Forecasts, 1995Q3-2009Q2

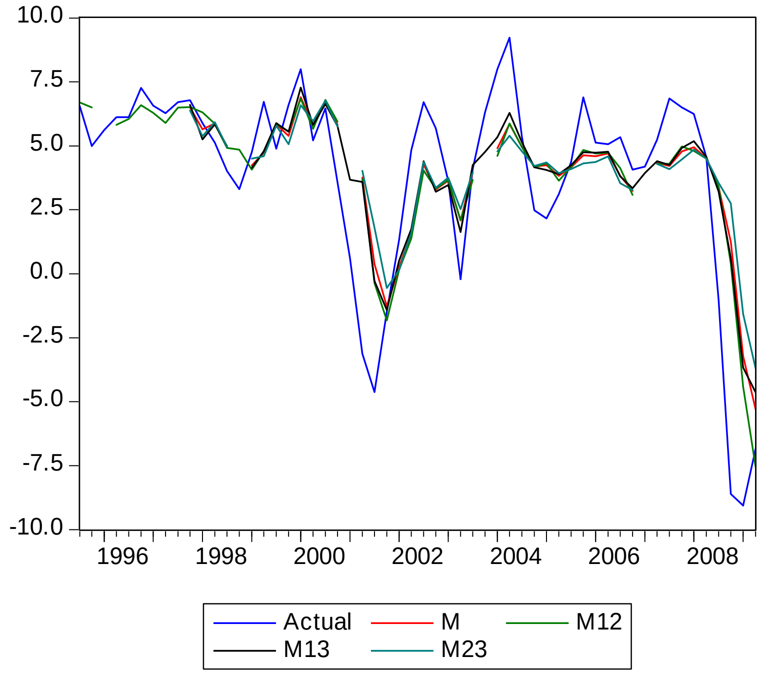

Table 1. Correlations of Inflation Rate, Three Forecasts, Mean of Three Forecasts, Means of Pairs of Forecasts, and their Replicable Counterparts

	Actual	F1	F2	F3	M	M12	M13	M23	$\hat{F}1$	$\hat{F}2$	$\hat{F}3$	$\hat{M}$	$\hat{M}12$	$\hat{M}13$	$\hat{M}23$
Actual	1.000														
F1	0.915	1.000													
F2	0.656	0.839	1.000												
F3	0.678	0.826	0.850	1.000											
M	0.803	0.947	0.947	0.939	1.000										
M12	0.828	0.964	0.953	0.873	0.987	1.000									
M13	0.845	0.964	0.883	0.946	0.987	0.966	1.000								
M23	0.693	0.865	0.964	0.960	0.981	0.950	0.950	1.000							
$\hat{F}1$	0.783	0.853	0.741	0.741	0.829	0.835	0.840	0.771	1.000						
$\hat{F}2$	0.699	0.778	0.822	0.769	0.836	0.833	0.810	0.828	0.901	1.000					
$\hat{F}3$	0.709	0.793	0.793	0.789	0.838	0.827	0.828	0.822	0.942	0.966	1.000				
$\hat{M}$	0.760	0.834	0.805	0.777	0.854	0.855	0.845	0.823	0.970	0.978	0.981	1.000			
$\hat{M}12$	0.766	0.840	0.802	0.770	0.853	0.857	0.845	0.817	0.974	0.974	0.971	0.999	1.000		
$\hat{M}13$	0.769	0.843	0.775	0.771	0.846	0.846	0.848	0.804	0.991	0.942	0.978	0.990	0.989	1.000	
$\hat{M}23$	0.710	0.791	0.817	0.784	0.844	0.838	0.824	0.833	0.925	0.994	0.987	0.988	0.981	0.965	1.000

Notes: F1: DGBAS: Directorate General of Budget, Accounting and Statistics (Government), F2: Chung-Hua: Chung-Hua Institution for Economic Research, F3: Taiwan: Taiwan Institute of Economic Research, M: Mean of three forecasts, M12: Mean of F1 and F2, M13: Mean of F1 and F3, M23: Mean of F2 and F3. Hats (circumflex) denote the replicable counterparts.

Table 2. Correlations of Real GDP Growth Rate, Three Forecasts, Mean of Three Forecasts, Means of Pairs of Forecasts, and their Replicable Counterparts

	Actual	F1	F2	F3	M	M12	M13	M23	$\hat{F}1$	$\hat{F}2$	$\hat{F}3$	$\hat{M}$	$\hat{M}12$	$\hat{M}13$	$\hat{M}23$
Actual	1.000														
F1	0.898	1.000													
F2	0.736	0.942	1.000												
F3	0.758	0.916	0.921	1.000											
M	0.832	0.984	0.978	0.960	1.000										
M12	0.842	0.990	0.980	0.931	0.996	1.000									
M13	0.866	0.990	0.953	0.964	0.995	0.988	1.000								
M23	0.760	0.950	0.986	0.973	0.990	0.979	0.976	1.000							
$\hat{F}1$	0.814	0.931	0.916	0.862	0.932	0.938	0.925	0.911	1.000						
$\hat{F}2$	0.702	0.898	0.950	0.874	0.931	0.933	0.907	0.936	0.963	1.000					
$\hat{F}3$	0.753	0.918	0.941	0.874	0.938	0.941	0.922	0.933	0.986	0.990	1.000				
$\hat{M}$	0.765	0.924	0.941	0.881	0.940	0.944	0.925	0.932	0.991	0.990	0.997	1.000			
$\hat{M}12$	0.771	0.925	0.939	0.875	0.940	0.944	0.925	0.930	0.993	0.988	0.997	0.999	1.000		
$\hat{M}13$	0.797	0.930	0.927	0.870	0.937	0.942	0.927	0.921	0.999	0.975	0.994	0.996	0.997	1.000	
$\hat{M}23$	0.718	0.906	0.949	0.878	0.935	0.937	0.913	0.937	0.972	0.999	0.995	0.995	0.993	0.983	1.000

Notes: F1: DGBAS: Directorate General of Budget, Accounting and Statistics (Government), F2: Chung-Hua: Chung-Hua Institution for Economic Research, F3: Taiwan: Taiwan Institute of Economic Research, M: Mean of three forecasts, M12: Mean of F1 and F2, M13: Mean of F1 and F3, M23: Mean of F2 and F3. Hats (circumflex) denote the replicable counterparts.

Table 3

Goodness-of-fit of Replicable and Non-Replicable Forecasts for Three Forecasts, Means of Three Forecasts, Means of Pairs of Forecasts, 1995Q3-2009Q2

Non-replicable Forecasts	Inflation Rate		Real GDP Growth Rate	
	RMSE	MAD	RMSE	MAD
F1	0.413	0.524	3.795	1.323
F2	1.409	0.943	8.079	1.888
F3	1.082	0.758	9.919	2.123
M	0.856	0.726	7.433	1.865
M12	0.790	0.715	5.568	1.584
M13	0.627	0.619	6.383	1.744
M23	1.201	0.836	9.690	2.130
Replicable Forecasts	**Inflation Rate**		**Real GDP Growth Rate**	
	RMSE	MAD	RMSE	MAD
$\hat{F}1$	0.895	0.754	6.209	1.946
$\hat{F}2$	1.325	0.964	9.678	2.262
$\hat{F}3$	1.108	0.851	10.51	2.217
$\hat{M}$	1.064	0.841	8.364	2.112
$\hat{M}12$	1.061	0.838	7.691	2.082
$\hat{M}13$	0.946	0.777	7.666	2.020
$\hat{M}23$	1.222	0.917	10.01	2.245

Note: RMSE and MAD denote root mean square error and mean absolute deviation, respectively.

Table 4a

Retrieving Replicable Components from the three Non-Replicable Forecasts

Included Variables	Inflation Rate		
	F1	F2	F3
Intercept	0.092	0.401	0.176
	(0.235)	(0.243)	(0.246)
Real GDP Growth(*t*-1)	0.127	0.156	0.103
	(0.030)***	(0.030)***	(0.031)***
Inflation(*t*-1)	0.544	0.133	0.119
	(0.228)**	(0.225)	(0.240)
F1(*t*-1)	0.040	0.266	0.255
	(0.368)	(0.373)	(0.383)
F 2(*t*-1)	-0.155	0.167	0.175
	(0.263)	(0.261)	(0.274)
F 3(*t*-1)	0.312	-0.079	0.072
	(0.224)	(0.213)	(0.240)
Adj. R^2	0.684	0.620	0.538
F test	17.89***	12.08***	9.840***

Notes: (i) The regression model (3) correlates the non-replicable forecasts, X_i , and W_i , in

$$X_i = W_i\theta_i + \eta_i\,,\ \eta_i \sim (0, \sigma_i^2 I)\,, \quad i = 1,2,3 \qquad (3)$$

where i = 1 for F1 forecast (government), i = 2 for F2 forecast (Chung-Hwa institution), and i = 3 for F3 forecast (Taiwan institution). W_i in (3) for the forecast for forecaster 1 is approximated by one-period lagged real growth, one-period lagged inflation, one period lagged forecast for forecaster 1, one period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. The F test is a test of expertise. Standard errors in parentheses.
** and *** denote significance at the 5% and 1% levels, respectively.

Table 4b

Retrieving Replicable Components from the Four Non-Replicable Mean Forecasts

Included Variables	Inflation Rate			
	M	M12	M13	M23
Intercept	0.304	0.291	0.153	0.347
	(0.221)	(0.229)	(0.218)	(0.226)
Real GDP Growth(*t*-1)	0.135	0.149	0.116	0.130
	(0.027)***	(0.029)***	(0.028)***	(0.028)***
Inflation(*t*-1)	0.274	0.312	0.353	0.146
	(0.204)	(0.211)	(0.212)	(0.209)
F1(*t*-1)	0.222	0.214	0.152	0.237
	(0.337)	(0.351)	(0.339)	(0.345)
F 2(*t*-1)	0.034	-0.040	0.002	0.190
	(0.236)	(0.246)	(0.242)	(0.242)
F 3(*t*-1)	0.035	0.090	0.157	-0.032
	(0.198)	(0.200)	(0.212)	(0.203)
Adj. R^2	0.682	0.682	0.665	0.639
F test	15.15***	15.55***	16.12***	12.68***

Notes: (i) The regression model (3) correlates the non-replicable forecasts, X_i , and W_i , in

$$X_i = W_i\theta_i + \eta_i ,\ \eta_i \sim (0, \sigma_i^2 I) ,\quad i = 1,2,3,4 \qquad (3)$$

where i = 1 for mean of 3 forecasters, i = 2 for mean of F1 and F2, i = 3 for mean of F1 and F3, and i = 4 for mean of F2 and F3. W_i in (3) for the forecast for forecaster 1 is approximated by one-period lagged real growth, one-period lagged inflation, one period lagged forecast for forecaster 1, one period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. The F test is a test of expertise. Standard errors in parentheses.
*** denotes significance at the 1% level.

Table 5a

Retrieving Replicable Components from the Three Non-Replicable Forecasts

Included Variables	Real GDP Growth Rate		
	F1	F2	F3
Intercept	0.495	0.765	2.077
	(0.761)	(0.502)	(0.546)***
Real GDP Growth(*t*-1)	0.664	0.246	0.222
	(0.141)***	(0.095)**	(0.102)**
Inflation(*t*-1)	-0.172	-0.093	-0.035
	(0.160)	(0.108)	(0.116)
F1(*t*-1)	0.131	0.383	0.220
	(0.382)	(0.256)	(0.275)
F2(*t*-1)	0.407	0.577	0.126
	(0.446)	(0.307)*	(0.321)
F3(*t*-1)	-0.344	-0.400	-0.069
	(0.386)	(0.259)	(0.277)
Adj. R^2	0.844	0.885	0.725
F test	45.52***	59.74***	22.05***

Notes: (i) The regression model (3) correlates the non-replicable forecasts, X_i , and W_i , in

$$X_t = W_t\delta_t + \eta_t\,,\ \eta_t \sim (0, \sigma_t^2 I)\,,\quad i = 1,2,3 \qquad (3)$$

where i = 1 for F1 forecast (government), i = 2 for F2 forecast (Chung-Hwa institution), and i = 3 for F3 forecast (Taiwan institution). W_i in (3) for the forecast for forecaster i is approximated by one-period lagged real growth, one-period lagged inflation, one period lagged forecast for forecaster 1, one period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. The F test is a test of expertise. Standard errors in parentheses.

* , ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 5b

Retrieving Replicable Components from the Four Non-Replicable Mean Forecasts

Included Variables	Real GDP Growth Rate			
	M3	M12	M13	M23
Intercept	1.053	0.577	1.283	1.391
	(0.554)*	(0.613)	(0.597)**	(0.477)***
Real GDP Growth(*t*-1)	0.392	0.471	0.447	0.235
	(0.106)***	(0.116)***	(0.111)***	(0.091)**
Inflation(*t*-1)	-0.072	-0.110	-0.099	-0.050
	(0.120)	(0.132)	(0.127)	(0.103)
F1(*t*-1)	0.200	0.212	0.168	0.291
	(0.284)	(0.313)	(0.301)	(0.244)
F2(*t*-1)	0.461	0.569	0.272	0.402
	(0.339)	(0.374)	(0.351)	(0.292)
F3(*t*-1)	-0.331	-0.418	-0.210	-0.271
	(0.286)	(0.315)	(0.303)	(0.246)
Adj. R^2	0.865	0.875	0.834	0.859
F test	48.55***	53.98***	41.21***	46.10***

Notes: (i) The regression model (3) correlates the non-replicable forecasts, X_i , and W_i , in

$$X_t = W_t\delta_t + \eta_t, \; \eta_t \sim (0, \sigma_t^2 I), \quad i = 1,2,3 \qquad (3)$$

where i = 1 for mean of 3 forecasters, i = 2 for mean of F1 and F2, i = 3 for mean of F1 and F3, and i = 4 for mean of F2 and F3. W_i in (3) for the forecast for forecaster i is approximated by one-period lagged real growth, one-period lagged inflation, one period lagged forecast for forecaster 1, one period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. The F test is a test of expertise. Standard errors in parentheses.

*, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 6a

Are Replicable Forecasts for Three Forecasts Accurate?

Estimation Method	Inflation Rate					
	Intercept	F1	F2	F3	Adj. R^2	F Test
OLS	-0.340	1.035			0.598	3.58**
	(0.248)	(0.135)***				
HAC	[0.156]***	[0.115]***				
OLS	-0.729		1.126		0.493	6.17***
	(0.358)**		(0.185)**			
HAC	[0.305]***		[0.180]***			
OLS	-0.673			1.249	0.517	5.03**
	(0.328)**			(0.191)***		
HAC	[0.237]***			[0.176]***		
Estimation Method	**Real GDP Growth Rate**					
	Intercept	F1	F2	F3	Adj. R^2	F Test
OLS	-0.374	1.081			0.637	0.20
	(0.591)	(0.127)				
HAC	[0.710]	[0.128]***				
OLS	-1.107		1.220		0.447	0.56
	(0.909)		(0.209)***			
HAC	[1.094]		[0.209]***			
OLS	-4.396			1.982	0.531	5.63***
	(1.216)***			(0.288)***		
HAC	[1.434]***			[0.296]***		

Notes: The regression model is

$$y_t = \gamma + \beta_i \hat{x}_t + \varepsilon_t , \quad i = 1,2,3 \qquad (7)$$

where i = 1 for F1 forecast (government), i = 2 for F2 forecast (Chung-Hwa institution), and i = 3 for F3 forecast (Taiwan institution). Standard errors in parentheses. Newey-West HAC standard errors are in brackets.

** and *** denote significance at the 5% and 1% levels, respectively.

The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$, i = 1,2,3.

Table 6b

Are Replicable Forecasts for Four Combined Forecasts Accurate?

Estimation Method	Intercept	M	M12	M13	M23	Adj. R^2	F Test
	Inflation Rate						
OLS	-0.693	1.195				0.562	4.55**
	(0.306)**	(0.167)***					
HAC	[0.264]**	[0.179]***					
OLS	-0.632		1.134			0.568	4.38**
	(0.295)**		(0.157)***				
HAC	[0.257]**		[0.167]***				
OLS	-0.534			1.171		0.583	4.39**
	(0.276)*			(0.157)***			
HAC	[0.190]***			[0.145]***			
OLS	-0.788				1.216	0.505	4.50**
	(0.351)**				(0.190)***		
HAC	[0.325]**				[0.225]***		
	Real GDP Growth Rate						
Estimation Method	Intercept	M	M12	M13	M23	Adj. R^2	F Test
OLS	-1.576	1.353				0.548	1.93
	(0.823)*	(0.190)***					
HAC	[1.215]	[0.208]***					
OLS	-0.784		1.172			0.559	0.65
	(0.719)		(0.161)***				
HAC	[1.074]		[0.176]***				
OLS	-1.830			1.412		0.605	2.30*
	(0.771)**			(0.177)***			
HAC	[1.100]			[0.186]***			
OLS	-2.314				1.500	0.472	2.47*
	(1.043)**				(0.244)***		
HAC	[1.572]				[0.286]***		

Notes: The regression model is

$$y = \gamma + \beta_t \hat{X}_t + \varepsilon_t, \quad i = 1,2,3,4 \qquad (7)$$

where i = 1 for mean of 3 forecasters, i = 2 for mean of F1 and F2, i = 3 for mean of F1 and F3, and i = 4 for mean of F2 and F3. Standard errors in parentheses. Newey-West HAC standard errors are in brackets.
*, **, and *** denote significance at the10%, 5% and 1% levels, respectively.
The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$, i = 1,2,3.

Table 7a

Examining Bias in Non-Replicable Forecasts for Three Forecasts

Estimation Method	Inflation Rate					
	Intercept	F1	F2	F3	Adj. R^2	F Test
OLS	-0.357 (0.118)***	1.009 (0.056)***			0.853	9.29***
GMM	-0.306 (0.092)***	0.993 (0.060)***			0.838	11.33***
OLS	-0.206 (0.280)		0.822 (0.124)***		0.467	7.77***
GMM	-0.394 (0.273)		0.747 (0.174)***		0.314	10.05***
OLS	-0.231 (0.235)			0.902 (0.135)***	0.492	3.41**
GMM	-0.323 (0.201)			0.738 (0.186)***	0.400	10.44***

Notes: The regression model is

$$y = Z\gamma + \beta_i X_i + (u_i - \beta_i \eta_i), \quad i = 1,2,3 \qquad (13)$$

where i = 1 for F1 forecast (government), i = 2 for F2 forecast (Chung-Hwa institution), and i = 3 for F3 forecast (Taiwan institution). The instrument list for GMM for forecaster i includes one-period lagged real growth, one-period lagged inflation, one-period lagged forecast for forecaster 1, and one-period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. Standard errors in parentheses.

*** denotes significance at the 1% level.

The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1, i = 1,2,3$.

Table 7b

Examining Bias in Non-Replicable Forecasts for Four Combined Forecasts

Estimation Method	Inflation Rate						
	Intercept	M	M12	M13	M23	Adj R2	F Test
OLS	-0.471 (0.231)**	1.044 (0.124)***				0.636	4.67**
GMM	-0.410 (0.249)	1.210 (0.128)***				0.577	1.44
OLS	-0.455 (0.203)**		1.010 (0.094)***			0.700	7.64***
GMM	-0.382 (0.191)*		0.893 (0.133)***			0.631	8.69***
OLS	-0.440 (0.168)**			1.065 (0.096)***		0.730	5.68***
GMM	-0.326 (0.152)**			0.828 (0.145)***		0.659	11.73***
OLS	-0.324 (0.286)				0.925 (0.152)***	0.472	3.90**
GMM	-0.262 (0.242)				0.666 (0.184)***	0.321	8.98***

Notes: The regression model is

$$y = Z\gamma + \beta_i X_i + (u_t - \beta_i \eta_t), \quad i = 1,2,3 \qquad (13)$$

where i = 1 for mean of 3 forecasters, i = 2 for mean of F1 and F2, i = 3 for mean of F1 and F3, and i = 4 for mean of F2 and F3. The instrument list for GMM for forecaster i includes one-period lagged real growth, one-period lagged inflation, one-period lagged forecast for forecaster 1, and one-period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. Standard errors in parentheses.

** and *** denote significance at the 5% and 1% levels, respectively.

The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1, i = 1,2,3$.

Table 8a

Examining Bias in Non-Replicable Forecasts for Three Forecasts

Estimation Method	Real GDP Growth Rate					
	Intercept	F1	F2	F3	Adj R^2	F Test
OLS	-0.565 (0.429)	1.118 (0.085)***			0.760	1.03
GMM	0.177 (0.324)	0.960 (0.050)***			0.768	0.35
OLS	-1.160 (0.788)		1.217 (0.164)***		0.516	1.09
GMM	-8.903 (2.396)***		2.845 (0.559)***		-0.586	7.47***
OLS	-3.720 (1.789)***			1.789 (0.239)***	0.550	6.26***
GMM	-11.72 (2.098)***			3.515 (0.497)***	-0.098	15.8***

Notes: The regression model is

$$y = \Sigma\gamma + \beta_i X_i + (u_i - \beta_i \eta_i), \quad i = 1,2,3,4 \qquad (13)$$

where i = 1 for F1 forecast (government), i = 2 for F2 forecast (Chung-Hwa institution) and i= 3 for F3 forecast (Taiwan institution). The instrument list for GMM for forecaster i includes one-period lagged real growth, one-period lagged inflation, one-period lagged forecast for forecaster 1, one-period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3.Standard errors in parentheses.
*** denotes significance at the 1% level.
The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$, i = 1,2,3.

Table 8b

Examining Bias in Non-Replicable Forecasts for Four Combined Forecasts

Estimation Method	Real GDP Growth Rate						
	Intercept	M	M12	M13	M23	Adj R2	F Test
OLS	-1.845 (0.720)**	1.411 (0.160)***				0.647	3.59**
GMM	-6.926 (1.469)***	2.439 (0.345)***				0.187	11.5***
OLS	-1.012 (0.577)*		1.209 (0.117)***			0.674	1.72
GMM	-5.328 (1.240)***		2.068 (0.293)***			0.241	10.1***
OLS	-2.019 (0.632)***			1.447 (0.140)***		0.703	5.56***
GMM	-5.978 (1.215)***			2.232 (0.287)***		0.426	12.5***
OLS	-2.473 (2.521)**				1.529 (0.586)***	0.534	3.38**
GMM	-11.26 (2.521)***				3.410 (0.586)***	-0.514	10.2***

Notes: The regression model is

$$y = \delta\gamma + \beta_i X_i + (u_t - \beta_i \eta_t), \quad i = 1,2,3 \qquad (13)$$

where i = 1 for mean of 3 forecasters, i = 2 for mean of F1 and F2, i = 3 for mean of F1 and F3, and i = 4 for mean of F2 and F3. The instrument list for GMM for forecaster i includes one-period lagged real growth, one-period lagged inflation, one-period lagged forecast for forecaster 1, and one-period lagged forecast for forecaster 2 and one period lagged forecast for forecaster 3. Standard errors in parentheses.

*, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

The F test is a test of the null hypothesis $H_0: \gamma = 0, \beta_i = 1$, i = 1,2,3.