Tourism Forecasting: to Combine or not to Combine?

Kevin K. F. Wong
Haiyan Song*
Stephen F. Witt
Doris C. Wu

School of Hotel and Tourism Management
The Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong SAR

* Corresponding author. Tel: +852 27666372; Fax: +852 2362 9362; Email: hmsong@polyu.edu.hk

Tourism Forecasting: to Combine or not to Combine?

Abstract

Existing non-tourism related literature shows that forecast combination can improve forecasting accuracy. This study tests this proposition in the tourism context by examining the efficiency of combining forecasts based on three different combination methods. The data used for this study relate to tourist arrivals in Hong Kong from the top ten tourism generating countries/regions. The forecasts are derived from four different forecasting models: integrated autoregressive moving average (ARIMA) model, autoregressive distributed lag model (ADLM), error correction model (ECM) and vector autoregressive (VAR) model. All forecasts are *ex post* and the empirical results show that the relative performance of combination versus single model forecasts various according to the origin-destination tourists flow under consideration, which parallels previous findings regarding the relative performance of individual forecasting methods. The results also vary with the combination techniques used. Furthermore, although the combined forecasts do not always outperform the best single model forecasts, almost all the combined forecasts are not outperformed by the worst single model forecasts. This suggests that forecast combination can considerably reduce the risk of forecasting failure. This conclusion also implies that combined forecasts are likely to be preferred to single model forecasts in many practical situations.

Key Words: Combination forecasting; econometric model; forecasting accuracy; tourism demand

1. Introduction

Along with the development of forecasting techniques, a large number of quantitative methods have been applied to the forecasting of tourism demand. Before the 1990s, traditional regression approaches dominated the tourism forecasting literature, but this trend changed from the mid-1990s as more researchers began to use modern econometric techniques, such as cointegration and error correction models, to model and forecast tourism demand; these studies include Song et al. (2003c), Kulendran and King (1997) and Morley (1998). However, each method has its own particular advantages/disadvantages. Empirical results demonstrate that no single forecasting method can generate the best forecasts in all situations and the relative accuracy of the different models varies with the origin/destination pairs and the lengths of the forecasting horizons (Witt and Song, 2002). No definitive criteria can be used to determine which forecasting method should be employed when a particular tourism demand forecasting task is performed.

This study aims to examine whether combining the tourism forecasts generated from different models can improve forecasting accuracy. The technique of combining forecasts was first introduced to the general forecasting literature by Bates and Granger (1969). Since then a large number of studies on forecast combination have been carried out. Seminal works include Dickinson (1973, 1975), Granger and Ramanathan (1984) and Min and Zellner (1993). The main objective of this approach is to obtain more accurate and stable forecasts through combining the advantages of different individual forecasting models. A number of forecast combination methods have been developed and empirical results from the general forecasting literature show that combining the forecasts generated from different models can considerably improve forecasting performance over the forecasts generated by the single forecasting models (see, for example, Diebold and Pauly, 1990 and Chan, Stock and Watson, 1999). However, rather surprisingly there has been virtually no work in the tourism context on forecast combination, with a study by Fritz *et al* (1984) being the exception. In the Fritz *et al* (1984) study only a traditional econometric model and an ARIMA model were considered and two forecast combination techniques were used. Both combination methods used simple weighting systems that took the historical performance of the individual forecasting methods into account. The study concluded that combining forecasts could improve the accuracy of forecasting airline visitors to the state of Florida. However, as traditional econometric models ignore data non-stationary, the empirical results obtained using these models are suspect.

The purposes of this study are to first provide a much more comprehensive examination than in the previous study (Fritz *et al* 1984) of whether or not it makes sense to combine tourism forecasts generated by different models in order to improve forecasting accuracy; and second to include modern econometric techniques for the first time in the comparison of combination versus single model forecasting accuracy. In this study, four modelling techniques - ARIMA, ADL, ECM and VAR - are used to generate the single model forecasts of tourist flows to Hong Kong; and three combination methods are applied to these four forecasting models in order to explore the relative efficiency of combining forecasts in forecasting tourism demand for Hong Kong. Two-, three- and four-model combinations are examined.

The rest of the paper is structured as follows. Section 2 reviews the existing studies on tourism modelling and forecasting and the development of forecast combination techniques. Section 3 explains the factors that affect the demand for tourism and describes the data sources. Section 4 explains the four forecasting models and three forecast combination methods. The empirical results are shown in section 5 and the last section concludes the study.

2. Literature Review

2.1. Tourism Forecasting

Tourism forecasting has become an important component in tourism research and different approaches have been used to generate forecasts of tourism demand. Witt and Witt (1995) provided a comprehensive review of the early tourism demand forecasting literature. Together with the rapid development of modern econometrics, many researchers have now applied these recent developments in forecasting tourism demand in various settings. Li, Song and Witt (2005) reviewed eighty-four post-1990 empirical studies of international tourism demand modelling and forecasting and gave an extensive and detailed view on issues such as data types and frequencies, independent and dependent variables, estimation methods and reported diagnostic test statistics. Their review suggested that the most frequently used forecasting methods in tourism are the static regression model, ADLM, ECM, VAR models, time varying parameter (TVP) model, almost ideal demand system (AIDS) and basic structural model (BSM).

Song, Witt and Li (2003c) used the general-to-specific modelling approach to obtain *ex ante* forecasts of the demand for Thai tourism. Song and Witt (2006) used the VAR modelling technique to forecast the demand for Macau tourism over the period 2003-2008. Kulendran and King (1997) considered four time series models and one econometric model when predicting quarterly tourist flows into Australia from four major tourist markets. Song *et al* (2003b) compared the forecasting performance of the ECM, ADLM, TVP and VAR models with those generated by the two univariate time series models in forecasting the demand for Denmark tourism and found that the TVP model generates the most accurate one-year-ahead forecasts. Li, Song and Witt (2006) reported the forecasts of tourist expenditure by UK residents in a number of Western European countries using the TVP and constant parameter linear AIDS models.

Although researchers have utilized the recent developments in econometrics to forecast tourism demand, the idea of combining the forecasts generated by different models, which has been widely used in forecasting macroeconomic and microeconomic activities, has attracted very little attention in the tourism literature and no attention since the adoption of recent developments in econometrics to forecast tourism demand. This study addresses this major deficiency in the tourism literature.

2.2. Forecasting Combination

Bates and Granger (1969) first introduced the idea of combining forecasts as a way of improving accuracy and since then the study of forecast combination techniques has mushroomed. Considerable efforts have been made to develop and improve the various forecast combination methods through empirical testing and/or simulations. Clemen (1989) reviewed a large number of published studies in this area and demonstrated that forecast combination generally leads to a considerable improvement in forecasting accuracy.

The simple average method is a straightforward combination technique, which assigns the same weight to each single forecast. Empirical results show that the simple average combination method can generate reliable forecasts in many situations. Makridakis and Winkler (1983) applied the simple average combination to a number of models and tested the effectiveness of this simple forecasting combination technique. Their study found that the average accuracy improves as the number of combined single methods increases. Palm and Zellner (1992) discussed the advantages and forecasting performance of the simple average combination technique also weighted combination techniques. They conclude that combining forecasts can reduce forecasting error and that a simple average combination may be more robust than weighted average combinations. The performance of the simple average combination method was found to be superior to the single forecasts by Fang (2003).

There are also many published studies on weighted average combination methods increased. These methods calculate the weights based on the past performance of each single forecast model. Among them the variance-covariance method was first introduced. In this method the weights are determined by a covariance matrix in which the accuracy of the single forecasts is embodied in the variances while the dependence between the single forecasts is interpreted by the covariance. Winkler and Makridakis (1983) tested a simple combination method and five variants of the variance-covariance combination method. They concluded that some variance/covariance procedures are more accurate than the simple combination technique and than individual forecasts, and the procedures in which covariance is ignored sometimes are more accurate than the ones in which variance is considered.

Extending this idea, Granger and Ramanathan (1984) showed that the optimal weights in the variance-covariance combination can be determined by a regression model and this regression-based combination technique has since attracted much attention among researchers. More sophisticated methods have also been developed in the literature. Through Monte Carlo experiments Chan *et al* (1999) demonstrated that principal component regression combinations are better than OLS combination methods in improving forecasting accuracy. Diebold and Pauly (1987) also used the principal component method to examine the accuracy of the combined forecasts in forecasting economic growth and they found that the best combined forecasts are much superior to the best single forecasts. Diebold and Pauly (1987) applied the TVP technique that utilizes the Kalman filter in the forecasting combination exercise.

In their study on forecasting combination, Diebold and Pauly (1990) developed a Bayesian shrinkage framework, which incorporates prior information in the estimation of the combination weights. The Bayesian combination method has been used in Anandalingam and Chen (1989a), Diebold and Pauly (1990), Min and Zellner (1993) and Walz and Walz (1989). These studies showed that Bayesian-based combination methods can improve the forecasting accuracy over other combination techniques.

Although the publications on the improvement of forecasting accuracy using various combination methods have been numerous, little effort has been made to explore why and when the forecasting combination techniques can improve forecasting accuracy. Flores and White (1989) suggested that combinations usually perform well when each forecast is based on different information/assumptions and they all cannot yield the needed accuracy. Hendry and Clements (2004) gave five potential explanations for the improvement in accuracy using forecast combination techniques: (i) if two models provide partial not completely overlapping explanations, the combination can better reflect all the information; (ii) when there is a structural break over the forecasting period, combining forecasts may help; (iii) when all models are mis-specified, combination can reduce variance; (iv) combination has an alternative interpretation of intercept correction which is well known to improve forecasting performance; and (v) combination can be viewed as “shrinkage” estimation.

Many of the above mentioned studies give support to the idea that forecast combination can significantly improve forecasting accuracy over the single forecasts. However, some researchers have suggested that forecast combinations do not always yield improvements in forecasting accuracy under all circumstances. For example, Winkler and Clemen (1992) found that combination forecasts performed poorly in their empirical studies due to the unstable combination weights being assigned to the different models and this was caused by the high correlations between the forecasts errors generated by the different models. More recently, Koning, Franses, Hibon and Stekler (2005) demonstrated that the combination of forecasting methods is not clearly more accurate than the single methods being combined using three univariate forecasting models and one combination technique. Hibon and Evgeniou (2005) reached a similar conclusion by testing more extensive combinations of many forecasting methods.

3. Data

This study focuses on the demand for Hong Kong tourism by residents from ten major origin countries/regions. The countries/regions that were ranked top ten in the period 2001-2004 according to *Visitor Arrival Statistics* published by the Hong Kong Tourism Board include: mainland China, Taiwan, Japan, USA, Macau, South Korea, Singapore, UK, Australia and Philippines.

The factors influencing tourism demand suggested by Song *et al.* (2003a) are followed in this study. These authors show that own price, substitute prices and consumer’s income are the primary factors influencing Hong Kong tourism demand. The own price P_{it} and substitute price P_{it}^{s} can be defined by equations (1) and (2) (Song, *et al.* (2003a):

$$P_{it} = \frac{CPI_{HK} / EX_{HK}}{CPI_{i} / EX_{i}} \tag{1}$$

$$P_{it}^{s} = \sum_{j=1}^{n} \frac{CPI_j}{EX_j} w_{ij} \tag{2}$$

where CPI and EX denote the consumer price index and the exchange rate respectively. i denotes the ith origin country/region, and j denotes the jth substitute destination. Here n=6 because six countries/regions are selected as substitute destinations of Hong Kong: China, Taiwan, Singapore, Thailand, Korea and Malaysia. w_j is the share of tourists visiting the jth substitute country/region among the summation of the tourists visiting these six countries. It can be denoted by $w_{ij} = \frac{TA_{ij}}{\sum_{j=1}^{6} TA_{ij}}$ in which TA_{ij} represents tourist arrivals from country/region i to the substitute destination j .

Consumer's income is measured by the real GDP index (2000=100) in constant prices of these ten origin countries/regions (*Song et al*, 2003a). Seasonal dummies are included in the forecasting models to capture the seasonal impacts, and one-off event dummies are used to capture the impacts of the hand-over of Hong Kong to China in 1997, SARS in 2003 and the "911" incident in the USA in 2001. Time trends are also considered in the models in order to improve the forecasting performance.

The sample covers the period from 1984q1 to 2004q2. The data period 1984q1 to 1999q2 is used to estimate the individual forecasting models and the subsequent period for forecasting evaluation. Most of the data are extracted from *Visitor Arrivals Statistics* published monthly by the Hong Kong Tourism Board, *Tourism Statistical Yearbook* published by the World Tourism Organization (WTO) and International Financial Statistics Online Service website of the International Monetary Fund (IMF). Some missing data are generated through extrapolation. All variables except the dummies are transformed to logarithms and the log-log linear models are used to explain the relationship between tourism demand and its determinants (Witt and Witt, 1995).

4. The Models

4.1. Individual forecasting methods

In this study, one time series method (seasonal ARIMA) and three econometric methods (ADLM, ECM and VAR) are used to generate the *ex post* forecasts. The choice of these models in this study was because these methods have been widely and successfully used in forecasting tourism demand (Li *et al.* 2005).

(1) Seasonal ARIMA

The seasonal ARIMA model is specified based on the standard Box-Jenkins method (Box and Jenkins, 1976). This method incorporates seasonal autoregressive and seasonal moving average structures and has been proved to be reliable in modelling and forecasting monthly or quarterly time series. The seasonal ARIMA models were specified based on the general-to-specific approach. That is, all potential terms - AR, MA, SAR and SMA, are included in the initial ARIMA model. Then the model was estimated and insignificant terms excluded stepwise until all terms in the model were significant and the model passed all the diagnostic statistics.

(2) ADLM

The dynamic econometric modelling technique advocated by (Hendry, 1986) is used to model the demand for Hong Kong tourism in this study. This methodology is known as the general-to-specific approach. This approach starts with a general ADLM and a stepwise reduction process is followed from the estimation of this general ADLM, which can be written as:

$$Q_{it} = \alpha + \sum_{n=1}^{p} \phi_{i,n} Q_{i,t-n} + \sum_{n=0}^{q_j} \beta_{i,n} X_{i,t-n} + \varepsilon_{it} \tag{3}$$

where Q_{it} denotes tourist arrivals from the origin country/region *i*. $X_{i,t-n}$ and β_{in} denote a vector of exogenous variables with a lag length of q_j and the coefficient vector, respectively. In our study the exogenous variables include the GDP index of country/region i, the own price P_{it}, the substitute price P_{it}^{s}, a time trend, the one-off event dummies and seasonal dummies.

Equation (3) is estimated first and the insignificant variables are removed from the equation. This process is repeated until there are no insignificant variables left in the equation. The final specific model should be simple in structure and posses the desirable statistical properties, that is, the estimated model should display lack of autocorrelation, heteroscedasticity, forecasting failure and non-normality.

(3) ECM

The Engle and Granger two-stage approach is used to specify the ECM (Engle and Granger, 1987). The long-run cointegrating regression model is estimated using OLS in the first step:

$$y_t = \beta_0 + \sum_{i=1}^{n} \beta_i x_{it} + u_t \tag{4}$$

Equation (4) indicates a long-term relationship among the independent and dependent variables. In the second step, the long-run cointegration relationship is transformed into an ECM process with the term $[\, y_{t-1} - \beta_0 - \sum_{i=1}^{n} \beta_i x_{it-1} \,]$ being added to equation (4), and the ECM is in the form of:

$$\Delta y_t = \sum_{i=1}^{p} \Delta y_{t-i} + \sum_{j=0}^{q} \Delta x_{t-j} + (y_{t-1} - \beta_0 - \sum_{i=1}^{n} \beta_i x_{it-1}) + \varepsilon_t \tag{5}$$

Equation (5) is termed the short-run dynamic ECM, which reflects the short-term relationship among the variables under consideration.

Since quarterly data are used in the model estimation, all the time series were subject to seasonal unit roots testing (the dummy variables were exempted from this test). Discussion of seasonal unit roots tests may be found in Dickey, Hasza and Fuller (1984), Engle, Granger and Hallman (1988), Hylleberg, Engle, Granger and Yoo (1990) and Engle, Granger, Hylleberg and Lee (1993). The HEGY test is used in this study to test for seasonal unit roots (Hylleberg, Engle, Granger and Yoo, 1990) and the test can be illustrated as follows.

Let y_t to represent a time series, which can be written in the form of

$$\Delta_4 y_t = (1-L^4)y_t = \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-2} + \pi_4 z_{3,t-1} + \varepsilon_t \tag{6}$$

where

$$z_{1,t} = (1+L+L^2+L^3)y_r \tag{7}$$

$$z_{2,t} = -(1-L+L^2-L^3)y_r \tag{8}$$

$$z_{3,t} = -(1-L^2)y_r \tag{9}$$

where L is a lag operator.

Equation (6) is then estimated and the one-tailed t test for π_1=0 and π_2 =0 and the F statistic test for $\pi_3 = \pi_4 = 0$ are carried out. The rejection of null π_1=0 means that the series y_t does not have non-seasonal unit roots. The rejection of null π_2=0 suggests that the series y_t does not have annual seasonal unit roots. And the rejection of the null $\pi_3 = \pi_4 = 0$ suggests that the series y_t do not posses semi-annual unit roots.

The test results show that all of the variables have non-seasonal unit roots. Tourist arrivals from Macau, USA, Singapore and Philippines have annual seasonal unit roots while the series of tourist arrivals from the UK has both annual and semi-annual seasonal unit roots. For the GDP variables, one semi-annual seasonal unit root is found in the Taiwan series. In addition, two seasonal unit roots were identified in the China, Macau, Korea, UK and Philippines series. To make sure that every available series has a unit root at the same frequency, the seasonal filter $(1+L)$ is used to remove the seasonal unit root at annual frequency in the variables in which π_2=0 is not rejected. While $(1+L^2)$ is used to eliminate the semi-annual seasonal roots of the variables in which $\pi_3 = \pi_4 = 0$ cannot be rejected. Furthermore, $(1+L+L^2+L^3)$ is employed to take out both the annual and semi-annual seasonal unit roots. As a result, each of the variables has only an annual seasonal unit root. That means they are all integrated at the same frequency. Following this, an Engle and Granger two-stage ECM approach can be specified.

(4) VAR

The VAR method is a system estimation technique which was first suggested by Sims (1980). This method treats all the variables as endogenous. The VAR method has been used widely in macroeconomic modelling and forecasting. Witt *et al.* (2003) and Song and Witt (2006) have successfully applied this technique to tourism demand forecasting. In this study, all explanatory variables are considered as endogenous except the constant, time trend and dummies. The lag lengths of the explanatory variables are determined by the Aikake Information Criterion (AIC) (Song and Witt, 2000, pp.93-94).

4.2. Forecasting Combination

As mentioned above, several forecast combination methods have been developed in the literature. In this study, three combination methods are used to test the performance of the different forecasting models. These are: simple, variance-covariance and discounted combination methods. The reason why these methods are chosen in this study is that the first two methods have been widely used in empirical studies in the general forecasting literature and most of the other combination approaches are developed from or modified versions of these methods. The third method involves ignoring the covariance among the single model forecasts when calculating the weights. This may be beneficial because it reduces potential instability.

(1) Simple Combination

In combining the forecasts generated by two or more models, it is important to decide the weights which will be assigned to each of the participating models. In the simple forecasting combination, the combination weight is assigned equally to each of the forecasts. The combination forecast is given by:

$$f_c = \sum_{i=1}^{n} w_i f_i \tag{10}$$

where f_i is the ith single forecast, f_c is the combined forecast generated by the n single forecasts f_i, and w_i is the combination weight assigned to f_i. In the simple average combination the weights can be specified as follows:

$$w_i = \frac{1}{n} \tag{11}$$

(2) Variance-Covariance Method

The variance-covariance method calculates the weights by taking the historical performance of the individual forecasts into consideration. The variance-covariance method determinates the weight vector according to:

$$w' = u' \Sigma^{-1} \big/ u' \Sigma^{-1} u \tag{12}$$

with the constraint $w_i \geq 0$. In equation (12) $w = (w_1, w_2, ..., w_n)'$; $\sum_{i=1}^{n} w_i = 1$; Σ denotes the sample covariance matrix; and u is a conformable column vector of ones: $(1,1,\ldots,1)'$.

A pair of single forecasts combination is used as an example. If $e_{it} = y_t - f_{it}$, $i = 1,2$, and y_t is the actual value of the corresponding forecast series, then

$$\varepsilon_t = y_t - f_{ct} = w_1 e_{1t} + w_2 e_{2t} \tag{13}$$

Considering $\sum_{i=1}^{n} w_i = 1$, equation (13) can be rewritten as:

$$\varepsilon_t = y_t - f_{ct} = w_1 e_{1t} + (1 - w_1) e_{2t} \tag{14}$$

which has mean zero. The variance is obtained as:

$$\sigma_c^2 = E(\varepsilon_t^2) = w_1^2 \sigma_1^2 + (1 - w_1)^2 \sigma_2^2 + 2 w_1 (1 - w_1) \sigma_{12} \tag{15}$$

where σ_c^2 denotes the variance of the combination forecast. σ_i^2 denotes the variance of the ith single forecast, and $\sigma_{i,j}$ denotes the covariance between the ith and the jth single forecasts.

Now by minimizing σ_c^2, the weight vector (w_i) can be expressed as:

$$w_1 = (\sigma_2^2 - \sigma_{12}) / (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) \tag{16}$$

$$w_2 = (\sigma_1^2 - \sigma_{12}) / (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) \tag{17}$$

with the constraint $w_i \geq 0$. In practice, σ_i^2 and σ_{ij} in the covariance matrix Σ can be replaced by $\sum e_{it}^2$ and $\sum e_{it} e_{jt}$, respectively. So the corresponding weights can be written as:

$$w_1 = (\sum e_{2t}^2 - \sum e_{1t} e_{2t}) / (\sum e_{1t}^2 + \sum e_{2t}^2 - 2\sum e_{1t} e_{2t}) \tag{18}$$

$$w_2 = (\sum e_{1t}^2 - \sum e_{1t} e_{2t}) / (\sum e_{1t}^2 + \sum e_{2t}^2 - 2\sum e_{1t} e_{2t}) \tag{19}$$

with the constraint $w_i \geq 0$ also is added in the weights determination. Because of the existence of correlations among the forecast errors, negative weights may appear in some cases, which might be considered unreasonable (see, for example, Newbold et al, 1987 and Clemen and Winkler, 1986).

(3) Discounted MSFE (Mean Square Forecast Error) method

The discounted MSFE method weights recent forecasts more heavily than distant ones. Winkler and Makridakis (1983) suggest that the weights can be written as:

$$w_i = \frac{1 / \sum_{t=1}^{T} \beta^{T-t+1} e_{it}^2}{1 / \sum_{i=1}^{n} \sum_{t=1}^{T} \beta^{T-t+1} e_{it}^2} \tag{20}$$

where β is the discounting factor with $0 < \beta \leq 1$, e_{it}^2 denotes the ith forecast error, and T and n denote the observation lengths used to obtain the weights and the number of combined single forecasts, respectively. The coefficient β assigns more weight to the most recent observations.

Apart from the β coefficient, the other difference between the discounted MSFE and the variance-covariance method is that equation (20) ignores the covariance among the errors. That is, in (12) Σ is diagonal and all off-diagonal elements are set to zero. The explanation can be found in Clemen and Winkler (1986). They suggested that when the correlations among the forecast errors are high, the combination weights are likely to be more sensitive to changes in the correlations. To avoid this instability caused by interdependence between the combination weights and correlations in the forecasting errors, the covariance matrix is ignored in equation (20).

4.3. Measures of Forecasting Accuracy

The existing published studies tend to use different error measures to compare the forecasting performance of different forecasting methods (e.g. Song *et al.,* 2003b, Martin and Witt, 1989). The most frequently used measure is the mean absolute percentage error (MAPE) (see Li *et al.*, 2005). Alternative error measures include the root mean square percentage error (RMSPE), mean absolute error (MAE) and Theil's U statistic. However, the use of these error measures is less frequent as compared with MAPE (Li, *et al.*, 2005). Following this tradition, this study uses MAPE to measure forecasting performance. A major advantage of this measure is that it does not depend on the magnitudes of the forecasting variables. Witt and Witt (1992) suggested that MAPE is the most appropriate error measure for evaluating the forecasting performance of tourism models. MAPE is calculated from:

$$MAPE = \frac{\sum_{t=1}^{n} \frac{|e_t|}{y_t}}{n} \tag{21}$$

where e_t is the forecast error, y_t is the actual value of the forecast variable, and *n* is the length of the forecasting horizon.

5. Empirical Results

In total, there are 11 forecast combination models for each combining method in this study. There are four single forecasting methods and of the combinations 6 involve two models, 4 involve three models and 1 combines all four models. The combination forecast series cover the period 1999q3 to 2004q2. The large gaps between the actual values and the forecasts in the second quarter of 2003 (caused by SARS) can potentially distort the forecasting evaluation as the outliers during the SARS period tend to have a much higher leverage effect on the calculation of the weights and the measure of MAPE.

To avoid this potential problem, we also recalculate the combination accuracy over the period 1999q3-2003q1. Therefore, our forecasting evaluation was carried out over two different sample periods.

The combination forecasts obtained by the simple average method for the period 1999q3-2003q1 are given in Table 1 (the summary results for the period 1999q3-2004q2 can be found in Tables 6-7).

Insert Tables 1-7 here

The combination forecasts that outperform the best single forecasts are labeled by an asterisk. The results suggest that the simple combination performs extremely well in the China and USA models in which almost all the combined forecasts are at least as good as the best single forecasts. In the Macau model six out of eleven combination forecasts perform better than the best single forecasts. As for the other models, the superiority of the combination forecasts is not evident.

The variance-covariance combination and discounted MSFE methods are assessed next. The combination weights are calculated from the previous performance of the single model forecasts. First, the individual forecasting models are re-estimated for the period 1984q1 to 1994q4 and then one-step-ahead forecasts are calculated from the four forecasting models (with iterative re-estimation). The post-sample combination forecasts were undertaken using two approaches: (i) the 38 forecasts (1995q1 to 2004q2) were divided into two parts and the first 18 observations were used to calculate the combination weights and these weights were then assigned to the latter 20 observations (over the period 1999q3-2004q2) for forecast comparison (Sankaran, 1989, Diebold and Pauly, 1990). (ii) The second approach is based on the one-step-ahead forecasts as explained by Clemen (1986). The optimal weights calculated from the previous 18 forecasts were assigned to the 19th forecast. This window was then continuously moved one-step ahead until the combination series included all 20 observations. The same assessment was repeated for the period 1999q3-2003q1 to avoid the influence of SARS.

The combination results based on the variance-covariance combination for the period 1999q3-2003q1 using the two approaches are shown in Tables 2-3. The results suggest that the variance-covariance method performs well for the China, USA, Macau and Australia models in the case of combination (a), and for the Macau, Singapore and UK models in the case of combination (b). There is no single model in which all the combinations outperform the single forecasts.

In terms of the discounted MSFE combination, the values of 0.6 and 0.9 were imposed on β in this study. The combination assessment results with $\beta = 0.9$ for the period 1999q3-2003q1 are shown in Tables 4-5. The discounted MSFE combination yields good results for China, USA, Macau and UK for both combinations (a) and (b).

The figures in Table 6 demonstrate the proportion of combination forecasts that outperform the best individual forecast among the 11 combined forecast models for each of the forecast combination methods. It is clear that combining forecasts does not always improve forecasting accuracy for all countries/regions concerned, but in many cases it is a worthwhile procedure.

Tables 1-5 demonstrate that whether or not combining the tourism forecasts generated by different models results in more accurate forecasts than those generated by the single models depends on both the origin-destination pair under consideration and the combination technique used. For example, the simple average and discounted MSFE combining methods work very well for China to Hong Kong and USA to Hong Kong tourist flows outperforming the most accurate single model forecasts in 91 per cent of the cases for China and 100 per cent of cases for USA. Furthermore, the variance-covariance combination (a) method gives reasonably good results for these origins, generating more accurate forecasts that the most accurate single model in 64 per cent of cases for both China and USA. By contrast, combining forecasts is less accurate than the most accurate single model forecasts in more than 50 per cent of cases for Taiwan to Hong Kong, Japan to Hong Kong, Korea to Hong Kong and Philippines to Hong Kong tourist flows for all five combination methods considered; and Singapore to Hong Kong and Australia to Hong Kong tourist flows for four of the five combination methods. Forecast combination generally works better than the most accurate single model forecasts for tourist flows from Macau (all combination methods) and UK (variance-covariance combination (b) and discounted MSFE methods). For Macau combining forecasts outperforms the best single forecasts in 73 per cent of cases for the variance-covariance and discounted MSFE methods, and for UK combining outperforms the most accurate single model forecasts in 73 per cent of the cases for the variance-covariance combination (b) and discounted MSFE combination (b) and discounted MSFE combination (b) methods. The empirical finings in this study show that whether or not tourism forecast combination outperforms the most accurate single model forecasts depends on the origin-destination pair under consideration parallels the previous empirical finding that the relative accuracy of single model forecasts varies with the origin-destination pair (Witt and Song, 2002).

The empirical results are now reconsidered from a different angle. Combination forecasts which are outperformed by the worst single forecast are focused on. These combination forecasts are labeled by the double asterisks in Table 1-5. The results show that for the sample period 1999q3-2003q1, there are only a few cases where the combination forecasts are outperformed by the worst single forecast, and these occurred only in the post-sample combination (b) using the variance-covariance method. (see Table 7).

The results in Table 7 illustrate that combining the forecasts generated by individual models almost always helps to avoid the worst forecasts. In addition, Table 3 shows that those rare cases in which combined forecasts are less accurate than each of the component individual forecasts only occur for two-model combinations. Where three or four models are included in the combination forecasts, the latter always outperform the least accurate single model included in the combination. The ability to avoid really bad forecasts would be particularly useful in a case in which one does not have any knowledge about the performance of single forecasting models. It would much safer to combine the forecasts in such situations. Witt and Song (2002) found that no single model consistently performs well in all situations (forecasting horizon, data frequency, and origin/destination country pairs). Thus combining forecasts can reduce the risk of forecasting failure in the tourism context. This conclusion is consistent with the study by Hibon and Evgeniou (2005), where they suggested that using a single method from a set of available methods is more risky than using a combination of methods.

6. Concluding Remarks

The research shows that the forecasting performance of each econometric and ARIMA model varies across countries/regions. Three combination methods have been used to compare the forecasting accuracy of combined forecasts against single model forecasts. The simple combination method is easy to understand and apply. The weights of this simple combination are known in advance and the past performance of each single forecast does not need to be considered. The variance-covariance and discounted MSFE methods require the calculation of the combination weights based on the past performance of the individual models.

The empirical results show that forecast combinations do not always outperform the best single forecasts. The research demonstrates that the relative performance of combination versus single model forecasts varies across origin countries/regions, which parallels the findings regarding the relative performance of individual forecasting methods. Whether or not combination forecasts outperform single model forecasts also depends on the combination technique used. The time series for different origin-destination pairs exhibit different properties and therefore it is not possible to generalize as to whether or not forecast combination is always likely to improve forecasting performance. The historical relative performance of combination/single model forecasts must be examined for each origin-destination tourist flow in order to ascertain whether forecast combination is likely to be beneficial.

A possible reason why combining forecasts may not result in greater accuracy is that the information included in each forecast tends to overlap. Hendry and Clements (2004) suggested that if two forecasts provide overlapping information, the combination of the two forecasts tends to not outperform the forecasts generated by each of the individual models. A second possible reason is that the forecasting performance of each individual model may be unstable over different time horizons. Batchelor and Dua (1995) suggested that forecast combination would only perform well under conditions in which the correlations of the forecast errors between individual models are low. Therefore, in practice, we should avoid combining forecasts if the correlations between the forecast errors are high. Winkler and Clemen (1992) found that the performance of combined forecasts is very sensitive to the weights assigned to each of the models and this is another reason why forecasting combination does not always improve the forecasting performance over the single model forecasts. However, this study shows that in many cases combining the forecasts generated by individual models results in greater accuracy than the forecasts produced by the most accurate individual model included in the combination.

This study also suggests that combination forecasts are almost certain to outperform the worst individual forecasts and avoid the risk of complete forecast failure. Therefore, in circumstances where a few forecasting models are available and the researcher has to generate forecasts but is uncertain as to which model is likely to generate the best forecasts, combining the forecasts from these alternative models would be the best and safest way forward.

References

Anandalingam, G., and Chen, L.(1989). Linear Combination of Forecasts: A General Bayesian Model. *Journal of Forecasting* **8**, 199.

Anandalingam, G., and Chen, L. (1989). Bayesian Forecast Combination and Kalman Filtering. *International Journal of Systems Science* **20**, 1499-1507.

Batchelor, R., and Dua, P. (1995). Forecaster Diversity and the Benefits of Combining Forecasts. M*anagement Science* **41**, 68-75.

Bates, J. M., and Granger, C. W. J. (1969). The Combination of Forecasts. *Operational Research Quarterly* **20**, 451-468.

Box, G. E. P., and Jenkins, G. M. (1976). Time Series Analysis. *Forecasting and Control (revised edition)*. San Francisco: Holden-Day.

C.Palm, F., and Zellner, A. (1992). To Combine or not to Combine? Issues of Combining Forecasts. *Journal of Forecasting* **11**, 687-701.

Chan, Y. L., Stock, J. H., and Watson, M. W. (1999). A Dynamic Factor Model Framework for Forecast Combination. *Spanish Economic Review* **1**, 91-121.

Clemen, R. T. (1986). Combining Economic Forecasts. *Journal of Business and Economic Statistics* **4**, 39-46.

Clemen, R. T. (1989). Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting* **5**, 559-583.

Clemen, R. T., and Winkler, R. L. (1986). Combining Economic Forecasts. *Journal of Business and Economic Statistics, **4**,* 39–46.

Dickey, D.A., Hasza, H. and Fuller, W.A. (1984) Testing for Unit Roots in Seasonal Time Series. *Journal of the American Statistical Association*, **79**, 355-367.

Dickinson, J.P. (1973) Some Statistical Results on the Combination of Forecasts. *Operational Research Quarterly*, **24**, 253-260.

Diebold, F. X., and Pauly, P. (1987). Structural Change and the Combination of Forecasts. *Journal of Forecasting* **6**, 21.

Diebold, F. X., and Pauly, P. (1990). The Use of Prior Information in Forecast Combination. *International Journal of Forecasting* **6**, 503-508.

Engle, R. F., Granger C. W. J. (1987) Co-integration and Error Correction: Representation, estimation, and testing. *Econometrica*, **55**, 251-276.

Engle, R.F., Granger, C.W.J. and Hallman, J. (1988) Merging Short- and Long-run Forecasts An Application of Seasonal Co-integration to Monthly Electricity Sales Forecasting. *Journal of Econometrics,* **40**, 45-62.

Engle, R.F., Granger, C.W.J., Hylleberg, S. and Lee H. S. (1993) Seasonal Cointegration, the Japanese Consumption Function. *Journal of Econometrics*, **55**,275-298.

Fang, Y. (2003). Forecasting Combination and Encompassing Tests. *International Journal of Forecasting* **19**, 87-94.

Flores, B. E., and White, E. M. (1989). Combining Forecasts:Why, When and How. *Journal of Business Forecasting Methods & Systems* **8**, 2-5.

Fritz, Brandon and Xander (1984) Combining Time Series and Econometric Forecast of Tourism Activity. *Annals of Tourism Research.* **11,** 219-229.

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

Hendry, D. F. (1986). Empirical Modeling in Dynamic Econometrics. *Applied Mathematics and Computation,* **20***, 201-236*

Hendry, D. F., and Clements, M. P. (2004). Pooling of Forecasts. *Econometrics Journal* **7**, 1-31.

Hibon, M., and Evgeniou, T. (2005) To Combine or not to Combine: Selecting among Forecasts and Their Combintions. *International Journal of Forecasting*, **21**, 15-24

Hylleberg, S., Engle R. F., Granger C. W.J., Yoo B.S. (1990) Seasonal Integration and Cointegration. *Journal of Econometrics*, **44**, 215-238.

Koning, J., Franses, H., Hibon, M. and Stekler, H.O. The M3 Competition:Statistical Tests of the Results. *International Journal of Forecasting,* **21**, 397-409

Kulendran, N., and King, m. L. (1997). Forecasting International Quarterly Tourist Flows Using Error Correction and Time-series Models. *International Journal of Forecasting* **13**, 319-327.

Li, G., Song, H., and Witt, S. F. (2005) Recent Developments in Econometric Modeling and Forecasting. *Journal of Travel Research.* **44,** *82-99.*

Li, G., Song, H., and Witt, S. F. (2006) Time Varying Parameter and Fixed Parameter Linear AIDS: An Application to Tourism Demand Forecasting. *International Journal of Forecasting.* 22.

Makridakis, S., and Winkler, R. L. (1983). Averages of Forecasts: Some Empirical Results. *Management Science* **29**, 987-996.

Martin, C. A., and Witt, S. F. (1989). Forecasting Tourism Demand: a Comparison of the Accuracy of Several Quantitative Methods. *International Journal of Forecasting* **5**, 7-19.

Min, C.K., and Zellner, A. (1993). Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates. *Journal of Econometrics* **56**, 89-118.

Morley, C.L.(1998) A Dynamic International Demand Model. *Annals of Tourism Research,* **25,** 70-84.

Newbold, P., Zumwalt, J. K., & Kannan, S. (1987). Combining Forecasts to Improve Earnings Per Share Prediction: An Examination of Electric Utilities. *International Journal of Forecasting, **3**,* 229-238.

Palm and Zellner (1992) To Combine or Not to Combine? Issues of Combining Forecasts. *Journal of forecasting*.**11**,687-701

Sankaran, S. (1989). A Comparative Evaluation of Methods for Combining Forecasts. *Akron Business and Economic Review* **20**, 33-39.

Sim,C.(1980) Macroeconomics and Reality. *Econometrica,***48**,1-48.

Song, H., Wong, K. F., and Chon, K.S. (2003a). Modelling and Forecasting the Demand for Hong Kong Tourism. *International Journal* of *Hospitality Management* **22**, 435-451.

Song, H., and Witt, S.F. (2000) Tourism Demand Modelling and Forecasting: Modern Econometric Approaches. Oxford: Pergamon.

Song, H., and Witt, S. F. (2006). Forecasting International Tourist Flows to Macau. *Tourism Management* **27,** 214-224**.**

Song, H., Witt, S. F., and Jensen, T. C. (2003b). Tourism Forecasting: Accuracy of Alternative Econometric Models. *International Journal of Forecasting* **19**, 123-141.

Song, H., Witt, S. F., and Li, G. (2003c). Modelling and Forecasting the Demand for Thai Tourism. *Tourism Economics* **9**, 363-387.

Tibiletti, L. (1994). A Non-linear Combination of Experts' Forecasts: A Bayesian Approach. *Journal of Forecasting* **13**, 21-31.

Walz, D. T., and Walz, D. B. (1989). Combining Forecasts: Multiple Regression Versus a Bayesian Approach. *Decision Sciences* **20**, 77-89.

Winkler, R. L., and Clemen, R. T. (1992). Sensitivity of Weights in Combining Forecasts. *Operations Research* **40**, 609-614.

Winkler, R. L., and Makridakis, S. (1983). The Combination of Forecasts. *Journal of the Royal Statistical Society. Series A* **146**, 150-157.

Witt, S. F., and Song, H. (2002). Forecasting Tourism Flows, in Lockwood, A. and Medlik, S. (eds) *Tourism and Hospitality in the 21st Century*, 106-18.

Witt, S. F., and Witt, C.A. (1992) Modeling and Forecasting Demand in Tourism. London: Academic Press: London.

Witt, S. F., and Witt, C.A. (1995) Forecasting tourism demand: a review of empirical research. *International Journal of Forecasting,* **11**, 447-475.

Table 1: MAPE for Single and Combined Forecasts of Tourist Arrivals in Hong Kong from by Origin Country/Region - Simple Average Combining Method, Sample Period 1999q3-2003q1

	China	Taiwan	Japan	USA	Macau	Korea	Singapore	UK	Australia	Philippines
A^1	6.55	4.37	6.10	7.07	13.71	10.92	10.63	5.31	7.16	8.24
A^2	8.33	4.49	8.60	7.99	7.86	7.41	9.00	5.23	6.52	6.68
E	8.28	6.96	6.64	6.64	13.44	9.80	8.99	5.69	7.56	9.59
V	6.90	11.51	5.37	6.39	8.85	9.66	11.68	4.55	7.71	8.79
A^1A^2	5.80*	3.77*	5.91*	6.95*	7.92	8.62	9.61	4.93*	6.84	7.21
A^1E	6.04*	5.38	5.60*	6.31*	11.38*	9.87	9.52	5.50	6.95*	7.84*
A^1V	6.61	6.99	5.43	6.32*	10.67	9.79	9.74*	4.62	7.09*	8.28
A^2E	8.22*	5.10	5.85*	6.58*	7.81*	7.06*	8.68*	5.11*	6.85	7.48
A^2V	6.10*	6.79	6.25	6.15*	6.03*	7.91	9.61	4.54*	7.09	7.31
EV	6.57*	8.37	5.73	6.20*	10.32	9.01*	8.98*	4.81	7.51*	7.95*
A^1A^2E	6.05*	4.61	5.25*	6.26*	8.08	8.23	9.21	5.00*	6.68	7.13
A^1A^2V	5.76*	5.43	5.53	6.15*	7.52*	8.55	9.35	4.44*	6.90	7.49
A^1EV	6.08*	6.62	5.47	6.19*	10.04	9.23*	9.30	4.97	6.96*	7.80*
A^2EV	6.61*	6.44	5.60	6.22*	6.96*	7.57	8.80*	4.62	7.03	7.23
A^1A^2EV	5.87*	5.59	5.16*	6.10*	7.59*	8.16	9.12	4.71	6.76	7.27

Notes: A^1, A^2, E and V denote ADL model, ARIMA model, ECM and VAR model respectively.
* Denotes that combination forecast is at least as good as the corresponding best combined single forecast involved in combination.

Table 2: MAPE for Single and Combined Forecasts of Tourist Arrivals in Hong Kong by Origin Country/Region - Variance-covariance Combining Methods, Sample Period 1999q3-2003q1, Combination (a)

	China	Taiwan	Japan	USA	Macau	Korea	Singapore	UK	Australia	Philippines
A^1	6.55	4.37	6.10	7.07	13.71	10.92	10.63	5.31	7.16	8.24
A^2	8.33	4.49	8.60	7.99	7.86	7.41	9.00	5.23	6.52	6.68
E	8.28	6.96	6.64	6.64	13.44	9.80	8.99	5.69	7.56	9.59
V	6.90	11.51	5.37	6.39	8.85	9.66	11.68	4.55	7.71	8.79
A^1A^2	6.55*	3.93*	8.60	7.11	6.95*	10.92	9.69	4.94*	6.61	7.74
A^1E	6.55*	5.34	6.27	6.38*	11.40*	10.92	9.28	5.31*	7.01*	7.76*
A^1V	6.55*	4.42	5.30*	6.35*	12.73	10.11	10.35*	5.02	7.14*	8.03*
A^2E	8.33	4.84	8.60	6.59*	7.27*	7.00*	8.99*	5.23*	6.51*	7.62
A^2V	8.33	4.74	8.60	6.38*	6.29*	8.20	8.75*	4.71	6.52*	6.78
EV	7.53	7.17	6.00	6.33*	10.97	9.39*	8.64*	4.51*	7.48*	8.42*
A^1A^2E	6.55*	4.29*	8.60	6.59*	6.67*	10.92	9.28	4.94*	6.61	7.76
A^1A^2V	6.55*	3.93*	8.60	6.40	6.29*	10.11	9.65	4.75	6.61	7.82
A^1EV	6.55*	5.34	6.00	6.31*	10.61	10.11	9.23	5.02	7.02*	7.76*
A^2EV	8.33	4.98	8.60	6.50	6.14*	8.20	8.64*	4.71	6.51*	7.63
A^1A^2EV	6.55*	4.29*	8.60	6.50	6.14*	10.11	9.23	4.75	6.61	7.76

Notes: Same as Table 1.

Table 3: MAPE for Single and Combined Forecasts of Tourist Arrivals to Hong Kong by Origin Country/Region - Variance-covariance Combining Methods, Sample Period 1999q3-2003q1, Combination (b)

	China	Taiwan	Japan	USA	Macau	Korea	Singapore	UK	Australia	Philippines
A^1	6.55	4.37	6.10	7.07	13.71	10.92	10.63	5.31	7.16	8.24
A^2	8.33	4.49	8.60	7.99	7.86	7.41	9.00	5.23	6.52	6.68
E	8.28	6.96	6.64	6.64	13.44	9.80	8.99	5.69	7.56	9.59
V	6.90	11.51	5.37	6.39	8.85	9.66	11.68	4.55	7.71	8.79
A^1A^2	7.52	3.92*	7.80	7.06*	7.15*	10.23	9.61	4.77*	7.20**	7.88
A^1E	7.80	5.22	6.31	6.43*	12.48*	11.04**	9.17	5.31*	6.98*	8.12*
A^1V	6.87	4.43	5.34*	6.52	9.67	9.97	10.38*	4.49*	7.32	8.02*
A^2E	8.54**	5.11	8.24	6.65	7.69*	7.94	8.81*	5.26	7.25	7.09
A^2V	7.90	4.71	8.00	6.14*	6.09*	8.51	8.81*	4.64	6.61	6.89
EV	7.19	7.30	5.91	6.83**	10.64	9.73	8.78*	4.48*	7.74**	8.29*
A^1A^2E	7.55	4.36*	8.26	6.65	7.27*	10.30	8.95*	4.77*	7.56	8.01
A^1A^2V	7.52	3.92*	8.04	6.38*	6.28*	9.54	9.77	4.49*	7.23	7.95
A^1EV	7.81	5.22	5.70	6.83	9.99	10.20	9.08	4.49*	7.10*	8.10*
A^2EV	7.94	5.30	8.43	6.58	6.98*	9.25	8.78*	4.67	7.30	7.18
A^1A^2EV	7.55	4.36*	8.39	6.60	7.13*	9.82	9.07	4.49*	7.59	8.01

Notes: Same as Table 1.

** Denotes that the combination forecast is inferior to the worst single forecast.

Table 4: MAPE for Single and Combined Forecasts of Tourist Arrivals in Hong Kong by Origin Country/Region - Discounted MSFE Method with β =0.9, Sample Period 1999q3-2003q1, Combination (a)

	China	Taiwan	Japan	USA	Macau	Korea	Singapore	UK	Australia	Philippines
A^1	6.55	4.37	6.10	7.07	13.71	10.92	10.63	5.31	7.16	8.24
A^2	8.33	4.49	8.60	7.99	7.86	7.41	9.00	5.23	6.52	6.68
E	8.28	6.96	6.64	6.64	13.44	9.80	8.99	5.69	7.56	9.59
V	6.90	11.51	5.37	6.39	8.85	9.66	11.68	4.55	7.71	8.79
A^1A^2	5.74*	3.83*	6.19	7.01*	6.87*	9.70	9.65	4.94*	6.80	7.29
A^1E	5.95*	5.45	5.60*	6.32*	11.40*	10.25	9.37	5.42	6.96*	7.80*
A^1V	6.59	5.70	5.43	6.31*	11.43	9.91	10.35*	4.67	7.10*	8.20*
A^2E	8.23*	5.04	6.09*	6.59*	7.03*	7.08*	8.68*	5.11*	6.77	7.44
A^2V	6.24*	5.42	6.53	6.19*	6.73*	8.19	8.78*	4.59	6.89	7.24
EV	6.58*	7.46	5.73	6.20*	10.96	9.19*	8.58*	4.51*	7.47*	7.97*
A^1A^2E	5.95*	4.52	5.44*	6.29*	6.68*	9.34	9.14	4.97*	6.65	7.14
A^1A^2V	5.69*	4.63	5.75	6.15*	6.41*	9.23	9.58	4.49*	6.85	7.48
A^1EV	6.02*	6.12	5.48	6.19*	10.46	9.52*	9.32	4.86	6.96*	7.77*
A^2EV	6.85*	5.42	5.69	6.23*	6.36*	7.91	8.60*	4.49*	6.87	7.21
A^1A^2EV	5.78*	4.90	5.31*	6.11*	6.63*	8.98	9.12	4.59	6.71	7.26

Notes: Same as Table 1.

Table 5: MAPE Single and Combined Forecasts of Tourist Arrivals in Hong Kong by Origin Country/Region - Discounted MSFE Method with β =0.9, Sample Period 1999q3-2003q1, Combination (b)

	China	Taiwan	Japan	USA	Macau	Korea	Singapore	UK	Australia	Philippines
A^1	6.55	4.37	6.10	7.07	13.71	10.92	10.63	5.31	7.16	8.24
A^2	8.33	4.49	8.60	7.99	7.86	7.41	9.00	5.23	6.52	6.68
E	8.28	6.96	6.64	6.64	13.44	9.80	8.99	5.69	7.56	9.59
V	6.90	11.51	5.37	6.39	8.85	9.66	11.68	4.55	7.71	8.79
A^1A^2	5.76*	3.87*	6.27	7.05*	6.83*	9.10	9.58	4.93*	6.94	7.35
A^1E	6.04*	5.25	5.65*	6.30*	12.09*	10.21	9.37	5.42	6.94*	8.04*
A^1V	6.68	4.95	5.42	6.26*	10.29	9.85	10.12*	4.51*	7.19	8.21*
A^2E	8.25*	5.03	6.25*	6.59*	7.35*	7.21*	8.75*	5.14*	6.89	7.32
A^2V	6.20*	5.13	6.53	6.13*	5.93*	8.29	8.74*	4.56	7.03	7.25
EV	6.60*	7.47	5.74	6.27*	10.23	9.26*	8.51*	4.45*	7.50*	8.16*
A^1A^2E	5.94*	4.46	5.70*	6.30*	7.32*	8.86	9.15	4.98*	6.77	7.14
A^1A^2V	5.76*	4.45	5.83	6.14*	6.58*	8.91	9.41	4.42*	7.00	7.40
A^1EV	6.16*	5.61	5.52	6.19*	10.01	9.50*	9.26	4.73	6.99*	7.97*
A^2EV	6.83*	5.41	5.94	6.22*	6.66*	7.96	8.61*	4.46*	7.01	7.18
A^1A^2EV	5.83*	4.78	5.50	6.10*	7.07*	8.63	9.09	4.54*	6.84	7.29

Notes: Same as Table 1.

Table 6: Percentage of combined forecasts which outperform the best individual forecasts

Method			Sample Period: 1999q3-2004q2	Sample Period: 1999q3-2003q1
Simple Average combination			53.64	47.27
Variance-covariance combination		**Post-sample combination (a)**	41.82	45.45
		Post-sample combination (b)	23.64	33.64
Discounted MSFE combination	β **=0.6**	**Post-sample combination (a)**	47.27	53.64
		Post-sample combination (b)	53.64	43.64
	β **=0.9**	**Post-sample combination (a)**	48.18	50.91
		Post-sample combination (b)	51.82	50.91

Table 7: Percentage of combined forecasts which are inferior to the worst individual forecasts

Method			Sample Period: 1999q3-2004q2	Sample Period: 1999q3-2003q1
Simple Average combination			0.00	0.00
Variance-covariance combination		**Post-sample combination (a)**	0.00	0.00
		Post-sample combination (b)	10.00	4.55
Discounted MSFE combination	β **=0.6**	**Post-sample combination (a)**	0.00	0.00
		Post-sample combination (b)	0.00	0.00
	β **=0.9**	**Post-sample combination (a)**	0.00	0.00
		Post-sample combination (b)	0.00	0.00