Forecast Combination through Factor Models: Assessing consensus and disagreement

Pilar Poncela

Dept. Análisis Económico: Economía Cuantitativa

Universidad Autónoma de Madrid

Avenida Tomás y Valiente, 5. 28049 Madrid. SPAIN

Telephone: +34-91-4975521.Fax:+34-91-4974091

e-mail: pilar.poncela@uam.es

Eva Senra

Dept. Estadística, Estructura Eca. y O.E.I.

Universidad de Alcalá

Plaza de la Victoria 2. 28802 Alcalá de Henares, Madrid. SPAIN.

Telephone: +34-91-8855232.Fax: +34-91-8854201

e-mail: eva.senra@uah.es

July 20, 2007

Abstract:

The combination of individual forecasts is often a useful tool to improve forecast accuracy. This paper considers factor models to produce a single forecast from several individual forecasts provided by the Survey of Professional Forecasters for the main US macroeconomic aggregates. We find that one factor models provide good adjustments and improvements in forecasting accuracy with respect to the average of the forecasts, but a second factor seems appropriate for some variables. There is also empirical evidence on high correlation between the first factor and the average of the forecasts and the second squared factor and the variance of them, which allows for an interpretation of the factors as measures related to consensus and disagreement among the forecasters. We develop necessary conditions based on the structure of the correlation matrix among the panelists, in order to obtain a related interpretation of the factors. We also check through a small simulation study how robust the result is.

Keywords:

Combining forecasts, dynamic factors, Survey of Professional Forecasters

1 Introduction

Dynamic factor models have spread a large amount of attention in the recent literature. Due to the possibility of larger databases nowadays, we need dimension reduction techniques. One of the most frequently tools used to achieve dimensionality reduction is factor analysis. In the context of time series analysis, some references related to this topic are, for instance, Anderson (1963), Priestley *et al.* (1974), Box and Tiao (1977), Geweke and Singleton (1981), Brillinger (1981), Velu *et al.* (1986), Peña and Box (1987), Stock and Watson (2002a), Forni *et al.* (2000, 2005) and Peña and Poncela (2004, 2006), among others.

Recently, factor models are commonly used to provide a single forecast of an endogenous variable from a large database of exogenous variables (see, for instance, Stock and Watson, 2002b and Favero et al., 2004, among others). In this paper we will use factor models to provide a single forecast of a certain variable from a larger set of individual forecasts, that is, we will use factor models as a forecast combination technique. The combination of such forecasts has been a useful tool to improving forecast accuracy since the early work of Bates and Granger (1969) and Newbold and Granger (1974), among others. Recent surveys on this topic can be found, for instance, in Clemen (1989), Diebold and Lopez (1996), Newbold and Harvey (2002) and Timmermann (2006). There are a large number of alternatives to combine forecasts. The most commonly used is the equal weights for all the individual forecasts (mean), that has proven to be hard to beat by most sophisticated alternatives, like basing the weighting mechanism in the variance-covariance matrix of the forecast errors or ridge regressions, among others, (see, for instance, Stock and Watson, 2004). As the combination of forecasts can be seen as a dimension reduction problem (from N panelists forecasts to just a single forecast), we consider the use of factor analysis as a tool to extract the common information to produce a consensus forecast and to reveal the level of disagreement amongst the different forecasters. The one factor model has been used previously in forecast combination by Figlewski (1983) and Figlewski and Urich (1983) who focused on forecast errors, and by Chan et al. (1999), who used an approximated factor model for the forecasts from different models.

Poncela and Senra (2006) used two factor models to combine US inflation forecast from the Survey of Professional Forecasters (SPF, from now on) and found that two factors were necessary to beat the mean of the forecasts. Moreover, they found that the first factor was very highly correlated with the average of the forecasts of the individual panelists and that the squared second factor was also correlated with the variance of these forecasts.

In this paper we extend the previous conclusions to several US macroeconomic aggregates from the Survey of Professional Forecasters and analyze the empirical characteristics that the point forecasts of different variables might share. We develop necessary conditions based on the structure of the correlation matrix among the panelists, in order to obtain a related interpretation of the factors. We also check through a small simulation study how robust the result is. Finally, we apply these results to the SPF.

The paper is organized as follows. In section 2 we present the factor model and the forecasting combination rule. In section 3 we describe the data and present the forecast combination results. In section 4 we seek for an interpretation of the common factors and provide necessary conditions for this interpretation to hold. We also illustrate the previous results through simulations. In section 5 we check if the previous results hold in the SPF. Finally, in section 6 we conclude.

2 The model

Let $y_{t+1|t,i}$ be the 1 period ahead forecast, with information up to time t, given by forecaster i, $i = 1, 2, ..., N$. Let $\mathbf{y}_{t+1|t} = (y_{t+1|t,1}, ..., y_{t+1|t,I})'$ be a N-dimensional vector of a certain macroeconomic forecasts. We use a factor model to separate the common information contained across the N forecasts, from the specific one to each forecaster. Consider the $r-$dimensional vector of common factors $\mathbf{f}_t = (f_{t1}, ..., f_{tr})', r < N$. Then, it can be assumed that the forecasts can be generated by a linear combination of the common information plus the specific components or error terms, such that

$$\mathbf{y}_{t+1|t} = \mathbf{\Lambda f}_t + \mathbf{e}_t, \tag{1}$$

where $\mathbf{\Lambda}$ is a $N \times r$ factor loading matrix, and $\mathbf{e}_t = (e_{t1}, ..., e_{tN})'$ is the vector of **specific errors**. Therefore, all the common correlated information comes through the common factors, $\mathbf{f}_t$, and the vector $\mathbf{e}_t$ contains information specific to each time series forecast. The error terms are assumed to have a diagonal variance-covariance matrix, since if there was information correlated among two terms in $\mathbf{e}_t$, it should be captured through the term $\mathbf{\Lambda f}_t$. We assume, furthermore, that cov($\mathbf{f}_t, \mathbf{e}_t$)=0.

We will follow Forni et al. (2000) and Stock and Watson (2002a) that use principal components to consistently estimate the dynamic factors. The procedure consists in estimating the r common factors through the first r principal components of the variance-covariance matrix of $\mathbf{y}_{t+1|t}$. It has been proven that estimated factor models in this way are useful in forecasting macroeconomic and financial variables (see, for instance, Stock and Watson, 2002b, Favero et al., 2004, and Zaher, 2007)

Let $\mathbf{v}_i$ be the eigenvector associated to the i- th largest eigenvalue of the $N \times N$ correlation matrix of $\mathbf{y}_{t+1|t}$. Then, the i-th principal component (which is a consistent estimate of the $i-$th common factor) is estimated as $\widehat{f}_{it} = \mathbf{v}_i'\mathbf{y}_{t+1|t}$. Let $\widehat{\mathbf{f}}_t = (\widehat{f}_{1t}, ..., \widehat{f}_{rt})'$ be the first estimated r principal components.

The factor combining rule to produce a unique one step ahead forecast $\widehat{y}^*_{t+1|t}$ from all the sources of information is given by

$$\widehat{y}^*_{t+1|t} = \widehat{\beta}_0 + \widehat{\beta}_1 \widehat{f}_{1t} + \cdots + \widehat{\beta}_r \widehat{f}_{rt}, \tag{2}$$

where $\widehat{\beta} = (\widehat{\beta}_0, ..., \widehat{\beta}_r)'$ is the ordinary least squares estimate from the regression

$$y_t = \beta_0 + \beta_1 \widehat{f}_{1t-1} + \cdots + \beta_r \widehat{f}_{rt-1} + u_t,$$

with y_t the observed variable in time t. Notice that $\widehat{y}^*_{t+1|t}$ are true ex-ante forecasts since the coefficients β_i are estimated only with information up to time t and do not include any information of the forecasting sample. Notice also that $\widehat{\beta}_0$ serves as a bias correction term.

3 The data

This paper considers the individual forecasts for several US macroeconomic aggregates from the Survey of Professional Forecasters, provided since 1990 by the Federal Reserve Bank of Philadelphia. This Survey gives regular macroeconomic forecasts from private sector economists (Wall Street financial firms, banks, economic consulting and other private firms). It is distributed free of charge and its forecasts are widely watched as they are reported in major newspapers like the Wall Street Journal and on financial newswires. The forecasts are anonymous and do not reflect the ideas of the institutions they belong to. The survey is conducted quarterly, and at each moment in time, it provides the last observed value and forecasts for one to five quarters ahead of the variable of interest, that allows us to have real one period ahead forecasts. Other surveys like the Livingston Survey, the Blue Chip Economic Indicators, the National Association of Business Economists (NABE) or the Consensus Forecast do not fill our requirements because of their periodicity or because they provide the forecasts for the mean of the year and in that case the forecasts are updated for the same time period but with different information sets.

The Survey of Professional Forecasters was thoroughly analyzed by Zarnowitz and Braun (1993) who found that the forecast combination of many of these individuals provided a consensus forecast with lower average errors than most individual forecasts. More recently, Harvey and Newbold (2003) have checked the non-normality of some of the forecast errors.

We consider one period ahead forecasts from the first quarter of 1991 until the last quarter of 2003 of the main U.S. Business Indicators at quarterly periodicity used in this paper that are shown in table 1. Information on the variables and their transformations in order to achieve stationarity are also given in table 1, where Δ denotes de difference operator, log is the natural logarithm, *sa* means seasonally adjusted, and *ar* means annual rate. The initial number of panelists provided by the Survey has to be reduced to those individuals that systematically collaborate, that is, they have been in the panel for a minimum of seven years and did not miss more than 4 consecutive forecasts.

Table 1: Information on the variables and their transformations.

Variable	Acronysm	Transformation
Gross Domestic Product (saar, $billions)	NGDP	$\Delta \log NGDP$
Civilian Unemployment Rate (sa, percent)	UNEMP	$\Delta \log UNEMP$
Housing Starts (saar, millions, monthly)	HOUSING	$\Delta \log HOUSING$
Consumer Price Index, CPI-U (saar, percent)	CPI	CPI
Treasury Bill Rate (three month, secondary market rate, discount basis percent)	TBILL	$\Delta \log TBILL$
AAA Corporate Bond Yield, Moody's (percent)	BOND	$\Delta BOND$

Regarding the treatment of missing data, we have estimated the unavailable one step ahead forecast by the two steps ahead forecasts made by the same individual in the previous quarter. With these considerations in mind, we consider 14 forecasters for each variable for the sample from 1991-III to 2003-IV, although they were not the same for each variable. Nevertheless we are going to use only the first part of the sample, from 1991-III to 1999-IV to estimate the models, and leave the remaining data to produce true ex-ante forecasts.

The benchmark forecast for comparison will be the average forecast of all the panelists. The procedure to analyze the different combinations of forecasts is as follows:

1. Estimate 1 and 2 factor models for the time span 1991-III to 1999-IV.

2. Generate one-step-ahead forecasts for the years in our forecasting sample, 2000-I to 2003-IV. In order to obtain one step-ahead forecasts, *models will be reestimated* adding one data point at the time, using all previous data prior to each forecast period.

3. Compute forecast errors for each forecast period. The root mean squared errors (RMSEs) by model will be computed to verify the forecasting performance of alternative combinations of

forecasts.

4. Compare the RMSE for the factor model to that obtained with the benchmark models by calculating the ratios between them.

$\frac{RMSE(\text{ factor model})}{RMSE(\text{average})}$

If this ratio is less than one, then the factor model improves the mean.

Table 2 compares the forecasting results in terms of the ratio of the RMSE of the different factor models over the RMSE of the average forecast (benchmark forecast) for different estimation and forecasting samples, attempting to verify the robustness of our forecasting exercise. So, a ratio less than one means that the factor model improves the benchmark forecast. The columns show the RMSE ratios for the different estimation and forecasting samples considered with 1 and 2 factor models.

Table 2: Ratio of the RMSE of the 1 and 2 factor models over the RMSE of the average forecast.

Forecasting sample	2000-2003		2001-2003		2002-2003		2003	
Number of factors	1	2	1	2	1	2	1	2
$\Delta \log NGDP$	0.980	0.965	0.967	0.951	0.958	0.943	0.975	0.963
$\Delta \log UNEMP$	0.833	0.964	0.796	0.873	0.929	0.965	0.731	0.882
$\Delta \log HOUSING$	0.815	0.807	0.772	0.763	0.775	0.762	0.807	0.779
CPI	1.009	0.982	0.975	0.939	1.045	0.921	0.980	0.892
$\Delta \log TBILL$	0.634	0.641	0.625	0.632	0.783	0.769	0.735	0.754
$\Delta BOND$	0.968	0.965	0.948	0.934	0.937	0.930	0.960	0.956

These results point out that factor models outperform the average of the forecasters in all the variables and periods considered but two, and that the gain in forecasting accuracy can be of an important magnitude, up to over 37%. The two losses are less important (0.9 and 4.5%). Notice also, that not always more factors mean better forecasting results, since it happens that one factor models may outperform 2 factor models. An interpretation of the factors will help in understanding the forecasting results.

4 Interpreting the factors: Consensus and disagreement in the forecasts

For the case of one factor models, it is usually found that the first common factor is a weighted average of all the individual forecasts. That is, all the components in the eigenvector $\mathbf{v}_1$, which is used to form the linear combination of the individual forecasts that estimates the common factor, are positive and about the same magnitude. If this is the case, one factor models will provide a first combination rule that captures the level of the variable to be forecast given by the consensus of all the individual forecasts. That is, it can be interpreted as the consensus forecast.

For all variables that we have analyzed the estimated first common factor resembles this behavior (see Table A.1 in Appendix 1). We have also checked its correlation with the equal weights forecast combination rule, that is the average of the individual forecasts. In Table 3 we show the correlation of the first common factor and the average forecast of all the panelists for each variable considered. Not surprisingly, it is *0.99* or higher.

Table 3: Correlations between the first factor ($\widehat{f}_{1t}$) and the average of the forecasts at each time t ($\overline{x_t}$) and correlations between the squared of the second factor ($\widehat{f}_{2t}$) and the variance (var) at each time t.

	corr($\widehat{f}_{1t}$,$\overline{x}_t$)	corr($\widehat{f}_{2t}^2$,var_t)
$\Delta \log NGDP$	0.996	0.69
$\Delta \log UNEMP$	0.995	0.52
$\Delta \log HOUSING$	0.998	0.88
CPI	0.999	0.85
$\Delta \log TBILL$	0.994	0.91
$\Delta BOND$	0.993	0.53

These high correlations between the first factor and the average of the forecasts at each moment in time could be due to the characteristics of the correlation matrix between the forecasters. (See table A.3 in appendix 1 for a summary of the bivariate correlations between forecasters for each variable). Notice that almost all of the correlations between the forecasts given by the different individuals happen to be positive and not too high.

Due to the estimation method, the second factor is orthogonal to the first common factor. The factor loadings are of different sign (see Table A.2 in Appendix 1). So, this second factor reveals different expectations in forecasting of the panelists. As our observations are forecasts, a measure of the magnitude of this second factor might be related to the disagreement in forecasting. We are going to see that this might be partly due to the structure of the correlation matrix. Nevertheless, the correlation about the measure of disagreement we have used and the second factor is not so high as the correlation between the first common factor and the mean of the forecasts of the individual panelists. The next proposition aims to explain the previous findings.

Proposition 1 *Let* $\mathbf{X} = \{x_{it}\}$ *the matrix of point forecasts provided by panelist* $i = 1, ..., N$ *at time* $t = 1, ..., T$ *whose correlation matrix is given by* $\mathbf{R}^T = \{r_{ij}^T\}$, $i, j = 1, ..., N$, *where the superindex* T *indicates the sample used to compute them. Let* $r_{ij}^T \xrightarrow{p} a$, $\forall i \neq j$; *where* a *is the population bivariate correlation among each pair of forecasters. Let the estimated common factors by principal components be given by* $\widehat{f}_{it}^T = \left(\mathbf{v}_i^T\right)' \mathbf{x}_{.t}$, *where* $\mathbf{v}_i^T$ *is the* i*-th eigenvector of matrix* $\mathbf{R}^T$ *and* $\mathbf{x}_{.t}$ *is the* t*-th row of matrix* $\mathbf{X}$. *Then, if* $a > 0$

(i) $\widehat{f}_{1t}^T \xrightarrow{p} \sqrt{N}\overline{x}_t = \frac{\sum_{i=1}^N x_{it}}{\sqrt{N}}$;

(ii) $\left(\widehat{f}_{2t}^T\right)^2 + \left(\widehat{f}_{3t}^T\right)^2 + \cdots + \left(\widehat{f}_{Nt}^T\right)^2 \xrightarrow{p} \sum_{i=1}^N (x_{it} - \overline{x}_t)^2$,

where all the limits are taken as $T \to \infty$.

Proof. The proof is given in appendix 2. ■

Notice that all limits are taken as $T \to \infty$. This means that the bivariate correlations are better estimated with longer historical forecasts. Nothing is said about the number of forecasts at each point of time. Notice also that the principal components are defined for each time t. Therefore, although the sample eigenvectors $\mathbf{v}_i^T$ converge in probability to the population eigenvectors $\boldsymbol{\nu}_i$ as $T \to \infty$, they also define the sample principal components at each time t. Analogously, the population eigenvectors define the population principal components at each time t.

The previous results provide necessary conditions to interpret the first common factor as proportional to the mean of the forecasts, but it also states the relationship between the disagreement and the remaining factors. These particular correlation matrixes among the forecasters for the different variables are showing two important facts relevant in this forecasting context: i) all the forecasters agree in the good or bad future behavior of the economy (because of the positive correlations) and ii) all the forecasters share only part of their knowledge with the rest, and consequently do provide relevant information to the final forecast.

Based on the previous proposition, as regards the number of factors, one should choose either only one factor or retain all of them. This is due to the fact that the explained variability given by factor i is given by

$$\frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i}$$

where λ_i is the corresponding ordered (from the largest to the smallest) i-th eigenvalue of the population correlation matrix among forecasters. It is straightforward to show that $\lambda_1 = (N-1)a + 1$ and $\lambda_i = 1 - a$, for $i = 2, .., N$. Therefore the explained variability provided by the first common factor should be $\frac{(N-1)a+1}{N}$ and each of the remaining ones explains $\frac{1-a}{N}$ of the total variability of the data. Recall that from the second factor onwards, all of them explain the same amount of variability. If one chooses the criterium of amount of variance explained by each of the common factors to select the number of factors, there would be no reason to choose just a few of them. Moreover, from part (ii) of Proposition 1, just the squared of the second common factor does not explain the variance, but in the case of $N = 2$ individuals.

4.1 Simulations

To check the validity of the previous results for finite samples, we have performed 1000 replications of two sets of four time series of 50 and 200 observations, respectively, with a correlation matrix of the form:

$$\begin{pmatrix} 1 & & & \\ a & 1 & & \\ a & a & 1 & \\ a & a & a & 1 \end{pmatrix}$$

for $a = 0.2, 0.4, 0.6$ and 0.8.

The simulation results are given in tables 4 and 5. Table 4 shows the variance explained by each factor for $T = 50$ and 200 respectively. By rows, we have the mean of the variance explained by each factor for the 1000 replications and by columns we have the same result by different values of a simulated for the two sample sizes considered. As it can be seen on Table 4, the percentage of variance explained by the first common factor grows with a, as expected, and the difference on the percentage of variance explained by the remaining factors is smaller for T=200, than for T=50. This might be an indication of the asymptotic result given in Proposition 1, where it is found that on the limit the share of the variance explained by the second factor onwards should be the same.

Table 4: % of the total variance explained by each factor for sample sizes $T = 50$ and 200 for different values of the bivariate correlations a.

	$a = 0.2$		$a = 0.4$		$a = 0.6$		$a = 0.8$	
	T=50	T=200	T=50	T=200	T=50	T=200	T=50	T=200
1^{st} factor	0.43	0.41	0.56	0.56	0.70	0.70	0.85	0.85
2^{nd} factor	0.25	0.23	0.19	0.17	0.13	0.12	0.07	0.06
3^{rd} factor	0.19	0.20	0.14	0.15	0.10	0.10	0.05	0.05
4^{th} factor	0.14	0.17	0.11	0.13	0.07	0.09	0.04	0.04

Table 5 shows the average and standard deviation (in parenthesis) results for the 1000 simulated replications for the two sample sizes considered $T = 50$ and 200. The first two rows indicate the bivariate correlation a and the sample sizes $T = 50$ and 200, respectively. On the third row, it is shown the average of the correlations between the first common factor and the mean of the forecasts at each time t; on the fourth row, the average of the correlations between the sum of the remaining squared factors and the variance of the forecasts at each time t; and on the fifth and sixth rows, the average of the correlations between the squared second factor and the sum of the second and third squared factors and the variance of the forecasts at each time t, respectively.

As it can be seen in table 5, the correlations between the first common factor and the mean of the forecasts at each time t are always very high (greater than 0.9) and grow with a, as expected. The same happens with the correlations between the sum of the squared of the remaining factors and the variance of the forecasts at each time t. If one only chooses only the squared of the second common factor to explain the variance (instead of the sum of the squared of the second factor onwards, as Proposition 1 states), it can be seen that although quite large (around 0.6 for the values simulated), it is not as large as the sum of them. The differences among the factors are smaller for T=200, than for T=50. This might be an indication of the asymptotic result given in Proposition 1, where it is found that on the limit all the share of the variance explained by the second factor onwards should be the same. Nevertheless, due to sample variability, the eigenvalues (from the second one onwards) are not the same, and therefore not each of these remaining factor weights the same on the sample replications.

Table 5: Average correlations (and standard deviation in parenthesis) between (i) the first common factor and the average forecast, (ii) the sum of the squared of the remaining factors and the variance among forecasters, (iii) the squared of the second factor and the variance among forecasters, and (iv) the sum of the squared of the second and third factors and the variance among forecasters for sample sizes $T = 50$ and 200 for different values of the bivariate correlations a among N=4 panelists.

	$a = 0.2$		$a = 0.4$		$a = 0.6$		$a = 0.8$	
N=4	T=50	T=200	T=50	T=200	T=50	T=200	T=50	T=200
$corr(\widehat{f}_{1t},\overline{x}_t)$	.87(.39)	.99(.01)	.99(.01)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)
$corr(\sum_{i=2}^{4}\widehat{f}_{it}^2,var_t)$	.90(.10)	.98(.02)	.96(.04)	.99(.01)	.97(.03)	.99(.01)	.98(.02)	.99(.01)
$corr(\widehat{f}_{2t}^2,var_t)$	.63(.15)	.64(.06)	.69(.10)	.65(.06)	.69(.10)	.65(.06)	.70(.10)	.65(.06)
$corr(\widehat{f}_{2t}^2+\widehat{f}_{3t}^2,var_t)$	.81(.11)	.85(.04)	.87(.06)	.86(.03)	.88(.05)	.86(.03)	.89(.05)	.86(.03)

As in the real case, we ended up with 14 panelists from the SPF, we also repeated the simulations with 14 series. The results are shown in table 6.

Table 6: Average correlations (and standard deviation in parenthesis) between (i) the first common factor and the average forecast, (ii) the sum of the squared of the remaining factors and the variance among forecasters, (iii) the squared of the second factor and the variance among forecasters, (iv) the sum of the squared of the second and third factors and the variance among forecasters, and (v) the sum of the squared of the second, third and fourth factors and the variance among forecasters,for sample sizes $T = 50$ and 200 for different values of the bivariate correlations a among N=14 panelists.

	$a = 0.2$		$a = 0.4$		$a = 0.6$		$a = 0.8$	
$N = 14$	T=50	T=200	T=50	T=200	T=50	T=200	T=50	T=200
$corr(\widehat{f}_{1t},\overline{x}_t)$	.98(.11)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)	1.00(.00)
$corr(\sum_{i=2}^{14} \widehat{f}_{it}^2, var_t)$	.96(.03)	.99(.00)	.97(.02)	.99(.00)	.97(.01)	.99(.00)	.98(.01)	1.00(.00)
$corr(\widehat{f}_{2t}^2, var_t)$	.47(.12)	.39(.06)	.48(.12)	.39(.07)	.48(.12)	.39(.06)	.49(.12)	.39(.06)
$corr(\widehat{f}_{2t}^2 + \widehat{f}_{3t}^2, var_t)$	.63(.10)	.53(.05)	.63(.10)	.53(.06)	.64(.10)	.53(.06)	.64(.09)	.53(.06)
$corr(\sum_{i=2}^{4} \widehat{f}_{it}^2, var_t)$	.72(.08)	.62(.05)	.73(.08)	.63(.05)	.73(.08)	.63(.05)	.73(.07)	.62(.05)

We wonder how robust is the previous result to deviations of equal correlation between each pair of individuals. In order to check this, we perform a second set of simulations (1000 replications) where the correlation between each two forecasters is given by $a + \xi_{ij}$, $i, j = 1, ..., N$, $i \neq j$, where $\xi_{ij} \sim N(0, \sigma)$ for $\sigma = 0.1$. This value is chosen to mimic the average of the observed bivariate correlations in the SPF. The cases where the realization of $a + \xi_{ij}$ was greater or equal to one were set to 0.9999 in the simulations.

We do not impose any dynamics on the time series. The results are shown in table 7. The structure of the table is the same as tables 5 and 6.

Table 7: Average correlations (and standard deviation in parenthesis) between (i) the first common factor and the average forecast, (ii) the sum of the squared of the remaining factors and the variance among forecasters, (iii) the squared of the second factor and the variance among forecasters, and (iv) the sum of the squared of the second and third factors and the variance among forecasters for sample sizes $T = 50$ *and 200 for stochastic bivariate correlations* $a + \xi_{ij}$ *among N=14 panelists.*

N=14	$a = 0.2$		$a = 0.4$		$a = 0.6$		$a = 0.8$	
$\sigma = 0.1$	T=50	T=200	T=50	T=200	T=50	T=200	T=50	T=200
$corr(\widehat{f}_{1t},\overline{x}_t)$	0.98(0.02)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.01)
$corr(\sum_{i=2}^{14}\widehat{f}_{it}^2,var_t)$	0.94(0.04)	0.99(0.01)	0.97(0.02)	0.99(0.00)	0.97(0.02)	0.99(0.00)	0.97(0.02)	0.77(0.18)
$corr(\widehat{f}_{2t}^2,var_t)$	0.49(0.12)	0.45(0.06)	0.53(0.12)	0.47(0.06)	0.55(0.12)	0.51(0.07)	0.57(0.11)	0.49(0.17)
$corr(\widehat{f}_{2t}^2+\widehat{f}_{3t}^2,var_t)$	0.65(0.10)	0.59(0.05)	0.69(0.09)	0.63(0.05)	0.71(0.08)	0.66(0.05)	0.73(0.08)	0.60(0.18)

Results shown in table 7 are compatible with those observed with the real forecasts of the Survey of Professional Forecasters.

5 Empirical results

Now we go back to the Survey of Professional Forecasters and test whether the results hold for the one period ahead forecasts. In order to check if the results given in Proposition 1 hold, we first test the equality of the correlations between each two forecasters by means of the statistical tests developed by Brien *et al.* (1984) and Jennrich (1970). In the last case we have used a given by the test of Brien *et al.* (1984), for the null hypothesis.

Table 8: p-values for the equal correlation tests for the variables considered in the SPF.

Variable	Brien et al. (1984)	Jennrich (1970)
$\Delta \log HOUSING$	0.2921	0.3700
CPI	0.0722	0.1130
$\Delta \log NGDP$	0.0002	0.0074
$\Delta BOND$	0.0001	0.0062
$\Delta \log UNEMP$	0.0000	0.0000
$\Delta \log TBILL$	0.0000	0.0000

For $\Delta \log HOUSING$ and CPI the null hypothesis of equal correlation between forecasters cannot be rejected, while for the rest it will be difficult to accept. So, for the two variables where we cannot reject the null hypothesis, table 9 summarizes the information on the explained variance given by each one of the factors.

Table 9: % of the variance explained by each factor for CPI and $\Delta \log HOUSING$.

Factor nr.	$\Delta \log HOUSING$	CPI
1	0.62	0.44
2	0.11	0.15
3	0.06	0.10
4	0.04	0.07
5	0.03	0.05
6	0.03	0.04
	...	...
	...	...
14	0.01	0.01

Table 10 shows the very high correlation between the first factor and the mean of the forecasts for each t and the sum of the squared of the remaining factors and the variance for each t as Proposition 1 stated. However, due to sample variability factors 2 to 14 are not accounting for the same amount of variance and just the squared of factor 2 shows a very high correlation with the variance of the individuals. Of course, this correlation increases as we consider more factors. Just with a few of them we can obtain a high correlation with the variance.

Table 10: Correlations between (i) the first common factor and the average forecast, (ii) the sum of the squared of the remaining factors and the variance among forecasters, (iii) the squared of the second factor and the variance among forecasters, (iv) the sum of the squared of the second and third factors and the variance among forecasters, and (v) the sum of the squared of the second, third and fourth factors and the variance among forecasters, for $\Delta \log HOUSING$ *and* CPI.

	$\Delta \log HOUSING$	CPI
$corr(\widehat{f}_{1t},\overline{x}_t)$	1.00	1.00
$corr(\sum_{i=2}^{14} \widehat{f}_{it}^2,var_t)$	0.93	0.98
$corr(\widehat{f}_{2t}^2,var_t)$	0.88	0.85
$corr(\widehat{f}_{2t}^2 + \widehat{f}_{3t}^2,var_t)$	0.91	0.94
$corr(\sum_{i=2}^{4} \widehat{f}_{it}^2,var_t)$	0.96	0.96

For the remaining variables the null hypothesis of equal correlations is rejected but in the simulation exercise we have seen that the results are robust to some deviations of the null hypothesis. Tables 11 and 12 confirm the previous conjecture, showing a very high correlation between the first common factor and the average of the panelists (0.99 or higher) and a very high correlation between the sum of the rest of the squared common factors with the variance among the forecasters at each time t (correlation 0.86 or higher), as well.

Table 11: % of the variance explained by each factor for the remaining variables.

Factor nr.	$\Delta BOND$	$\Delta \log NGDP$	$\Delta \log TBILL$	$\Delta \log UNEMP$
1	0.47	0.46	0.70	0.49
2	0.11	0.13	0.09	0.13
3	0.09	0.09	0.06	0.09
4	0.07	0.07	0.05	0.06
5	0.06	0.06	0.04	0.05
6	0.05	0.04	0.02	0.04
14	0.01	0.00	0.00	0.01

Table 12: Correlations between the first common factor and the average forecast, and the sum of the squared of the remaining factors and the variance among forecasters, for the remaining variables.

	Δ $BOND$	Δ $\log NGDP$	Δ $\log TBILL$	Δ $\log UNEMP$
$corr(\widehat{f}_{1t},\overline{x}_t)$	0.99	1.00	0.99	0.99
$corr(\sum_{i=2}^{14}\widehat{f}_{it}^2,var_t)$	0.92	0.96	0.86	0.87

6 Concluding Remarks

Factor analysis seems a reasonable alternative to combine forecasts from the Survey of Professional Forecasters. Although just one factor is enough to beat the average of the forecasters, except for CPI (see Poncela and Senra, 2006, relating the need for the second factor to Friedman's hypothesis), a second factor also might improve the forecast accuracy, measured in terms of the RMSE.

We have also found empirical evidence on high correlation between the first factor and the average of the forecasts and the second squared factor and the variance of them, which would allow for an interpretation of the factors as measures related to consensus and disagreement among the forecasters. In particular, we have developed necessary conditions based on the structure of the correlation matrix among the panelists, in order to obtain a related interpretation of the factors. If the panelists are equally correlated, we have proven that the first common factor is proportional to the average of forecasts at each time t and the sum of the remaining squared common factors is also proportional to the variance at each time t. A simulation study has shown that the result is robust to small deviations of the equal correlation hypothesis.

We have checked the conjecture of the equal correlation between the panelists and it cannot be rejected for CPI and $\Delta \log HOUSING$. We have also found that the first common factor accounts for most of the total variance and that due to the sample variability each of the remaining common factors explain slightly decreasing percentage of variance (instead of all of the the same). We have also seen that at each time t the first common factor is the average of the panelists (correlation 0.99 or higher between the first common factor and the average) and that the sum of the rest of the squared common factors is very highly correlated with the variance among the forecasters at each time t (correlation 0.93 or higher).

We have also found that although the equal correlation hypothesis among each pair of forecasters cannot be accepted for the remaining variables ($\Delta BOND$, $\Delta\ \log NGDP$, $\Delta\ \log TBILL$ and $\Delta\ \log UNEMP$), the simulation performed showed that the results were robust to some deviations from this null hypothesis. In fact, we still see that at each time t the first common factor is the average of the panelists (correlation 1 between the first common factor and the average) and that the sum of the rest of the squared common factors is very highly correlated with the variance among the forecasters at each time t (correlation 0.86 or higher).

We would like to point out that the results hold for this type of data which consists of survey of forecasts. These data seem not too far from the equal correlation hypothesis. Most of the forecasters agree about the future behavior of the economy (it seems that the market does not pay for disagreement). We do not know if the same type of results would apply, for instance, to real data. In the case of real data, the hypothesis of equal correlation among measure variables seems hard to meet. Nevertheless, the same type of results could apply to other surveys of forecasts.

Acknowledgments: This research has been sponsored by the IIF-SAS Grant 2005. Results were presented at the ISF2006 in Santander, Spain.

7 Appendix 1

Table A.1: Factor loadings for the first common factor for all the variables.

	Δ logNGDP	Δ logUNEMP	Δ logHOUS	CPI	Δ logTBILL	ΔBOND
1	0.021800	0.177700	0.333200	0.347800	0.013300	0.027600
2	0.255400	0.287700	0.232400	0.220300	0.328600	0.227900
3	0.254400	0.263300	0.454500	0.273700	0.310900	0.253900
4	0.407300	0.358600	0.154500	0.318500	0.373900	0.425700
5	0.203500	0.163300	0.340900	0.267100	0.056700	0.121200
6	0.148200	0.188300	0.169700	0.237400	0.313000	0.158900
7	0.225500	0.293700	0.307100	0.312300	0.310900	0.239900
8	0.324300	0.269000	0.197900	0.206300	0.184000	0.310100
9	0.297700	0.155600	0.213700	0.205000	0.177800	0.304200
10	0.239400	0.102400	0.154900	0.250400	0.260800	0.220600
11	0.289900	0.202100	0.251300	0.268100	0.298400	0.314100
12	0.282700	0.275300	0.323000	0.171200	0.320000	0.290500
13	0.249800	0.340000	0.181100	0.283100	0.199200	0.255500
14	0.334800	0.225100	0.236900	0.316800	0.304900	0.344100

Table A.2: Factor loadings for the second common factor for all the variables.

	Δ logNGDP	Δ logUNEMP	Δ logHOUS	CPI	Δ logTBILL	ΔBOND
1	0.113700	0.213800	-0.664900	-0.365100	0.045700	-0.181700
2	-0.038300	0.067500	0.306100	-0.182700	0.224600	-0.002600
3	-0.067700	-0.107300	0.457600	-0.158000	0.082300	-0.056700
4	0.142800	-0.181700	0.054500	-0.036800	-0.020600	-0.042400
5	-0.870600	0.432500	-0.306300	0.063500	-0.458000	0.823900
6	0.351900	-0.179800	0.086200	0.844900	0.124900	-0.227000
7	0.045000	-0.157900	0.121200	-0.058000	0.129300	-0.003800
8	0.041600	0.021700	0.112600	-0.094500	-0.054700	-0.099700
9	-0.095000	0.153500	-0.026100	-0.111200	-0.627500	0.214100
10	0.100900	0.685700	0.211600	-0.010800	0.241000	-0.376600
11	0.232800	-0.085000	-0.251700	0.036500	-0.105600	-0.068800
12	-0.016400	-0.220600	0.031800	-0.093500	0.032500	-0.059000
13	-0.012500	-0.031000	-0.108200	0.000800	-0.471800	0.140400
14	0.036400	0.024400	0.040700	0.232800	0.092500	0.078300

Table A.3: Bivariate correlations summary

	CPI	$\Delta BOND$	$\Delta \log HOUSING$	$\Delta \ \log NGDP$	$\Delta \ \log TBILL$	$\Delta \ \log UNEMP$
Average	0.41	0.36	0.54	0.37	0.54	0.45
Variance	0.02	0.05	0.01	0.04	0.09	0.04

8 Appendix 2

Proof of Proposition 1:

Proof of part (i):

If $r_{ij}^T \xrightarrow{p} a, \forall i \neq j$, the limiting correlation matrix $\boldsymbol{\rho}$ is if the form

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & & & & \\ a & 1 & & & \\ a & a & 1 & & \\ \vdots & \vdots & & \ddots & \\ a & a & a & \cdots & 1 \end{pmatrix}.$$

Let $\boldsymbol{\nu}_i$ and λ_i the corresponding ordered (from the largest to the smallest) i-th eigenvector and eigenvalue of the populations matrix $\boldsymbol{\rho}$, respectively, and $\mathbf{v}_i^T$ and l_i^T those of the sample correlation matrix $\mathbf{R}^T$. It is easy to show that $\boldsymbol{\nu}_1 = (1/\sqrt{N}, ..., 1/\sqrt{N})$ and $\lambda_1 = (N-1)a + 1$ and $\boldsymbol{\nu}_i = (\frac{1}{\sqrt{i(i-1)}}, \frac{1}{\sqrt{i(i-1)}}, ..., \frac{1}{\sqrt{i(i-1)}}, -\underset{\substack{\uparrow \\ i}}{\frac{(i-1)}{\sqrt{i(i-1)}}}, 0, ..., 0)$ and $\lambda_i = 1 - a$, for $i = 2, ..., N$. By the Slutsky theorem, $l_i^T \xrightarrow{p} \lambda_i$ and $\mathbf{v}_i^T \xrightarrow{p} \boldsymbol{\nu}_i$ as $T \to \infty, \forall i = 1, ..., N$.

Define $\widehat{f}_{it}^T = \left(\widehat{\mathbf{v}}_i^T\right)' \mathbf{x}_{.t}$. If $\boldsymbol{\nu}_1 = (1/\sqrt{N}, ..., 1/\sqrt{N})$, then $\widehat{f}_{1t} = \left(\mathbf{v}_1^T\right)' \mathbf{x}_{.t} \xrightarrow{p} \boldsymbol{\nu}_1' \mathbf{x}_{.t} = \frac{\sum_{i=1}^N x_{it}}{\sqrt{N}} = \sqrt{N}\overline{x}_t$.

Proof of part (ii):

Let $\widehat{f}_{it}^T$ be defined as in (i) and consider

$$\left(\widehat{f}_{2t}^T\right)^2 + \left(\widehat{f}_{3t}^T\right)^2 + \cdots + \left(\widehat{f}_{Nt}^T\right)^2 \xrightarrow{p} f_{2t}^2 + f_{3t}^2 + \cdots + f_{Nt}^2$$

where the probability convergence is obtained applying that $l_i^T \xrightarrow{p} \lambda_i$ and $\mathbf{v}_i^T \xrightarrow{p} \boldsymbol{\nu}_i$ as $T \to \infty, \forall i = 2, ..., N$ and the Slutsky theorem. Now we are going to show, by induction, that

$$f_{2t}^2 + f_{3t}^2 + \cdots + f_{Nt}^2 = \sum_{i=1}^N (x_{it} - \overline{x}_t)^2 \tag{3}$$

so

$$\left(\widehat{f}_{2t}^{T}\right)^2 + \left(\widehat{f}_{3t}^{T}\right)^2 + \cdots + \left(\widehat{f}_{Nt}^{T}\right)^2 \xrightarrow{p} \sum_{i=1}^{N}(x_{it} - \overline{x}_t)^2$$

and the proof is completed.

To show by induction that (3) holds, first let us prove it for $N = 2$

$$f_{2t}^2 = (\boldsymbol{\nu}_2'\mathbf{x}_{.t})^2 = \left((\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0, ..., 0)\mathbf{x}_{.t}\right)^2 = \frac{1}{2}(x_{1t}^2 + x_{2t}^2 - 2x_{1t}x_{2t})^2 = \sum_{i=1}^{2}(x_{it} - \overline{x}_t)^2$$

where $\overline{x}_t = \frac{\sum_{i=1}^{2} x_{it}}{2}$. Now assume that it is true for $N-1$ and let us show it for N, that is if $f_{2t}^2 + f_{3t}^2 + \cdots + f_{N-1t}^2 = \sum_{i=1}^{N-1}(x_{it} - \overline{x}_t^{(N-1)})^2$, (where the superindex in the average $\overline{x}_t^{(N-1)} = \frac{\sum_{i=1}^{N-1} x_{it}}{N-1}$ indicates the number of elements used to compute the average), then let us show that (3) holds. First, notice that

$$\begin{aligned}
f_{Nt}^2 &= (\boldsymbol{\nu}_N'\mathbf{x}_{.t})^2 = \left((\frac{1}{\sqrt{N(N-1)}}, \frac{1}{\sqrt{N(N-1)}}, ..., \frac{1}{\sqrt{N(N-1)}}, \frac{-(N-1)}{\sqrt{N(N-1)}})\mathbf{x}_{.t}\right)^2 \\
&= \frac{1}{N(N-1)}\left(\sum_{i=1}^{N-1} x_{it}\right)^2 + \frac{(N-1)}{N}x_{Nt}^2 - \frac{2}{N}\left(\sum_{i=1}^{N-1} x_{it}\right)x_{Nt} \\
&= \frac{N-1}{N}\left(\left(\overline{x}_t^{(N-1)}\right)^2 + x_{Nt}^2 - 2\overline{x}_t^{(N-1)}x_{Nt}\right) \\
&= \frac{N-1}{N}\left(\overline{x}_t^{(N-1)} - x_{Nt}\right)^2 \qquad (4)
\end{aligned}$$

In a similar way to $\overline{x}_t^{(N-1)}$, denote by $\overline{x}_t^{(N)} = \frac{\sum_{i=1}^{N} x_{it}}{N}$ the sample average at time t computed with N elements. Then

$$\begin{aligned}
\sum_{i=1}^{N}(x_{it}-\overline{x}_t^{(N)})^2-\sum_{i=1}^{N-1}(x_{it}-\overline{x}_t^{(N-1)})^2 &= \sum_{i=1}^{N} x_{it}^2 - N\left(\overline{x}_t^{(N)}\right)^2-\sum_{i=1}^{N-1} x_{it}^2+(N-1)\left(\overline{x}_t^{(N-1)}\right)^2 \\
&= x_{Nt}^2 - N\left(\frac{\sum_{i=1}^{N} x_{it}}{N}\right)^2+(N-1)\left(\frac{\sum_{i=1}^{N-1} x_{it}}{N-1}\right)^2 \\
&= x_{Nt}^2 - \frac{1}{N}\left(\sum_{i=1}^{N-1} x_{it}+x_{Nt}\right)^2+\frac{1}{N-1}\left(\frac{\sum_{i=1}^{N-1} x_{it}}{N-1}\right)^2 \\
&= x_{Nt}^2 - \frac{1}{N}\left(\left(\sum_{i=1}^{N-1} x_{it}\right)^2+x_{Nt}^2+2x_{Nt}\sum_{i=1}^{N-1} x_{it}\right)^2+\frac{1}{N-1}\left(\frac{\sum_{i=1}^{N-1} x_{it}}{N-1}\right)^2 \\
&= \frac{N-1}{N}\left(x_{Nt}^2+\left(\overline{x}_t^{(N-1)}\right)^2-2x_{Nt}\overline{x}_t^{(N-1)}\right) \\
&= \frac{N-1}{N}\left(\overline{x}_t^{(N-1)}-x_{Nt}\right)^2. \qquad (5)
\end{aligned}$$

From (4) and (5),

$$\sum_{i=1}^{N}(x_{it}-\overline{x}_t^{(N)})^2=\sum_{i=1}^{N-1}(x_{it}-\overline{x}_t^{(N-1)})^2+f_{Nt}^2$$

and since by hypothesis the result is true for $N-1$, that is $f_{2t}^2+f_{3t}^2+\cdots+f_{N-1t}^2=\sum_{i=1}^{N-1}(x_{it}-\overline{x}_t^{(N-1)})^2$, then

$$\sum_{i=1}^{N}(x_{it}-\overline{x}_t^{(N)})^2=f_{2t}^2+f_{3t}^2+\cdots+f_{N-1t}^2+f_{Nt}^2$$

which completes the proof.

9 References

Anderson, T.W. (1963) The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika,* 28, 1-25.

Bates, J.M. and C.W.J. Granger (1969) The combination of forecasts, *Operations Research Quarterly*, 20, 451-468.

Box, G., Tiao, G. (1977) A canonical analysis of multiple time series. *Biometrika,* 64, 355-65.

Brien, C. J., W. N. Venables, A. T. James and O. Mayo (1984) An Analysis of Correlation Matrices: Equal Correlations, *Biometrika*, 71, 545-554.

Brillinger, D.R. (1981) *Time series Data Analysis and Theory*, expanded edition. Holden-Day, San Francisco.

Chan, Y.L., J. Stock, and M. Watson (1999) A dynamic factor model framework for forecast combination, *Spanish Economic Review*, 1, 91-121.

Clemen, R.T. (1989) Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting,* 5, 559-583.

Diebold, F.X. and J.A. López (1996) Forecast Evaluation and Combination, *Handbook of Statistics,* vol. 14, 241-268 (North-Holland, Amsterdam).

Favero, C., Marcellino, M, and Neglia, F. (2004) Principal components at work: the empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, forthcoming.

Figlewski, S. (1983) Optimal Price Forecasting Using Survey Data, *The Review of Economics and Statistics*, 65, 13-21.

Figlewski, S. and T. Urich (1983) Optimal Aggregation of Money Supply Forecasts: Accuracy, Profitability and Market Efficiency, *The Journal of Finance,* 28, 695-710.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000) The generalized factor model: Identification and estimation, *Review of Economic and Statistics,* 82, 540-554.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005) The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, Vol. 100, No. 471, 830-840 .

Geweke, J.F., Singleton K.J. (1981) Maximum Likelihood Confirmatory Analysis of Economic Time Series. *International Economic Review,* 22, 37-54.

Harvey, D. and Newbold, P. (2003) The non-normality of some macroeconomic forecast errors *International Journal of Forecasting*, Volume 19, Issue 4, 635-653.

Jenrich, R. I. (1970). An asymptotic χ^2 test for the equality of two correlation matrices. *Journal of the American Statistical Association*, 65, 904-912.

Newbold, P. and C.W.J. Granger (1974) Experience with forecasting univariate time series and the combination of forecasts, *Journal of the Royal Statistical Society A,* 137, 131-146.

Newbold, P. and D.I. Harvey (2001) Forecasting combination and encompassing in *A companion to economic forecasting* (Eds.) M. Clements and D. Hendry, Blackwells, Oxford, pp. 268-283.

Peña, D., Box, G. (1987). Identifying a simplifying structure in time series. *J. Am. Statist. Ass.,* 82, 836-43.

Peña, D. and Poncela, P. (2004). Forecasting with nonstationary dynamic factor models. *Journal of Econometrics* 119, 291-321.

Peña, D. and Poncela, P. (2006) Nonstationary dynamic factor analysis. *Journal of Statistical Ploanning and Inference,* 136, 1237-1257.

Poncela, P. and Senra, E. (2006) A two factor model to forecast US inflation. *Applied Economics,* Volume 38, Issue 18, 2191-2197.

Priestley, M.B., Rao, T.S., Tong, J. (1974) Applications of principal component analysis and factor analysis in the identification of multivariable systems. *IEEE Trans. Autom. Control*, 19, 703-4.

Stock, J.H., M. Watson (2002a) Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association,* 97, 1167-1179.

Stock, J.H., M. Watson (2002b) Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics*, 20, 147-163

Stock, J.H., M. Watson (2004) Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23, 405-430.

Velu, R.P., Reinsel, G.C., Wichern, D.W. (1986) Reduced rank models for multiple time series. *Biometrika,* 73, 105-18.

Zarnowitz, V. and P. Braun (1993) Twenty-two years of the NBER-ASA quarterly economic Outlook Survey: Aspects and Comparisons of Forecasting Performance, in *Business Cycles, Indicators and Forecasting* (Eds.) J. Stock and M. Watson, The University of Chicago Press, Chicago, pp. 11-93.

Zaher, F. (2007) Evaluating factor forecasts for the UK: The role of asset prices *International Journal of Forecasting* (In Press).