Forecast Combination with Entry and Exit of Experts

Carlos Capistrán[1] Allan Timmermann[2]

[1]Banco de México

[2]University of California, San Diego

October, 2007

Overview of this Paper

- Most expert surveys take the form of unbalanced panels as individual forecasters frequently enter and exit from surveys
- Important concern for real-time forecasting
- This paper considers ways to combine expert opinions that work even in the presence of forecasts data that is incomplete
- We study several approaches and propose a new, simple approach that projects the realized value on a constant and the equal-weighted forecast

Overview of this Paper

- Most expert surveys take the form of unbalanced panels as individual forecasters frequently enter and exit from surveys
- Important concern for real-time forecasting
- This paper considers ways to combine expert opinions that work even in the presence of forecasts data that is incomplete
- We study several approaches and propose a new, simple approach that projects the realized value on a constant and the equal-weighted forecast

Overview of this Paper

- Most expert surveys take the form of unbalanced panels as individual forecasters frequently enter and exit from surveys
- Important concern for real-time forecasting
- This paper considers ways to combine expert opinions that work even in the presence of forecasts data that is incomplete
- We study several approaches and propose a new, simple approach that projects the realized value on a constant and the equal-weighted forecast

Overview of this Paper

- Most expert surveys take the form of unbalanced panels as individual forecasters frequently enter and exit from surveys
- Important concern for real-time forecasting
- This paper considers ways to combine expert opinions that work even in the presence of forecasts data that is incomplete
- We study several approaches and propose a new, simple approach that projects the realized value on a constant and the equal-weighted forecast

Motivation

Entry and exit of experts

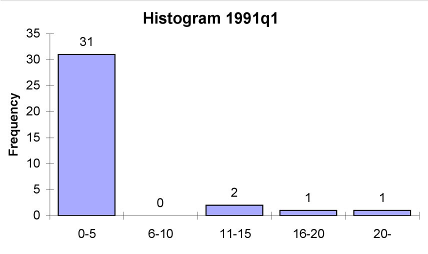
Figure 1: Participants in the Survey of Professional Forecasters (inflation forecasts)

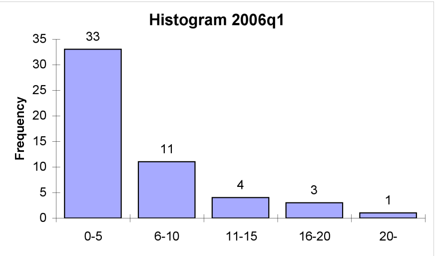
ID	95.1	95.2	95.3	95.4	96.1	96.2	96.3	96.4	97.1	97.2	97.3	97.4	98.1	98.2	98.3	98.4	99.1	99.2	99.3	99.4
481		x	x	x	x				x	x	x	x		x	x	x	x			x
483		x	x				x			x	x	x		x	x		x			x
485		x	x	x		x	x	x	x	x	x	x	x	x		x				
486			x	x			x	x		x		x		x						
487		x	x	x																
488		x	x	x		x	x	x	x	x	x	x	x		x		x		x	x
489		x																		
490				x			x	x	x	x	x									
491							x	x												
494												x								
495														x	x					
496															x	x				
497															x	x	x		x	
498																x	x	x	x	x
499																		x	x	x
500																			x	x
501																		x	x	x
502																		x	x	x
504																		x	x	x
505																		x	x	x

Notes: The ID corresponds to the identification number assigned to each forecaster in the survey. The columns represent the quarter when the survey was taken. The Xs show when a particular forecaster responded to the inflation part of the survey and provided a one-step-ahead forecast.

Motivation

Number of contiguous forecasts per forecaster





Theoretical Results

Y_{t+1} : variable of interest; $\widehat{\mathbf{Y}}_{t+1,t}$: $N \times 1$ vector of forecasts

$$\begin{pmatrix} Y_{t+1} \\ \widehat{\mathbf{Y}}_{t+1,t} \end{pmatrix} \sim \left(\begin{pmatrix} \mu_y \\ \mu \end{pmatrix} \begin{pmatrix} \sigma_y^2 & \sigma'_{y\widehat{\mathbf{y}}} \\ \sigma_{y\widehat{\mathbf{y}}} & \mathbf{\Sigma}_{\widehat{\mathbf{y}}\widehat{\mathbf{y}}} \end{pmatrix} \right).$$

$\mathbf{e}_{t+1,t} = Y_{t+1}\iota - \widehat{\mathbf{Y}}_{t+1,t}$: vector of forecast errors

$$\begin{aligned} \mathbf{\Sigma}_e &= E[\mathbf{e}_{t+1,t}\mathbf{e}'_{t+1,t}] \\ &= (\sigma_y^2 + \mu_y^2)\iota\iota' + \mu\mu' + \mathbf{\Sigma}_{\widehat{\mathbf{y}}\widehat{\mathbf{y}}} - \iota\sigma'_{y\widehat{\mathbf{y}}} - \sigma_{y\widehat{\mathbf{y}}}\iota' - \mu_y\iota\mu' - \mu_y\mu\iota'. \end{aligned}$$

Forecaster's standard problem under squared loss:

$$\min \omega' \mathbf{\Sigma}_e \omega$$
$$s.t.\ \omega'\iota = 1.$$

Constraint ensures unbiasedness of the combination if $\mu = \mu_y \iota$.

Theoretical Results

Standard solution for the optimal weights

$$\omega^* = (\iota' \mathbf{\Sigma}_e^{-1} \iota)^{-1} \mathbf{\Sigma}_e^{-1} \iota.$$

The optimal weights depend on the full covariance matrix, $\mathbf{\Sigma}_e$. Only in very special cases are these reduced to equal weights−the most prominent case being when the forecast errors have identical variance, σ^2, and identical pair-wise correlations, ρ:

$$\begin{aligned} \mathbf{\Sigma}_e^{-1} \iota &= \frac{\iota}{\sigma^2(1 + (N-1)\rho)} \\ (\iota' \mathbf{\Sigma}_e^{-1} \iota)^{-1} &= \frac{N}{\sigma^2(1 + (N-1)\rho)}, \end{aligned}$$

and so the optimal weights are given by:

$$\omega^* = \left(\frac{1}{N} \right) \iota.$$

Combination Methods in Common Use

Equal Weights (EW)

$$\bar{Y}_{t+1,t} = N_t^{-1} \sum_{i=1}^{N_t} \hat{Y}_{t+1,t,i}$$

- The simple average has proven surprisingly difficult to outperform (Clemen (1989); Makridakis and Hibbon (2000); Stock and Watson (2001,2003))

Combination Methods in Common Use

Equal Weights (EW)

- The robustness of the simple average forecast across different data samples and forecasting methods remains a puzzle
- One would expect to find considerable heterogeneity in experts' forecasting ability - this ought to be exploitable by differentiating the weights applied to different forecasts
- In practice, however, individual forecasters' true ability—and consequently the combination weights—are unknown
- Improving on the EW average requires having a procedure for estimating the combination weights which ensures that the sample estimates do not get too far removed from their true but unknown values

Combination Methods in Common Use

Equal Weights (EW)

- The robustness of the simple average forecast across different data samples and forecasting methods remains a puzzle
- One would expect to find considerable heterogeneity in experts' forecasting ability - this ought to be exploitable by differentiating the weights applied to different forecasts
- In practice, however, individual forecasters' true ability—and consequently the combination weights—are unknown
- Improving on the EW average requires having a procedure for estimating the combination weights which ensures that the sample estimates do not get too far removed from their true but unknown values

Combination Methods in Common Use

Equal Weights (EW)

- The robustness of the simple average forecast across different data samples and forecasting methods remains a puzzle
- One would expect to find considerable heterogeneity in experts' forecasting ability - this ought to be exploitable by differentiating the weights applied to different forecasts
- In practice, however, individual forecasters' true ability—and consequently the combination weights—are unknown
- Improving on the EW average requires having a procedure for estimating the combination weights which ensures that the sample estimates do not get too far removed from their true but unknown values

Combination Methods in Common Use

Equal Weights (EW)

- The robustness of the simple average forecast across different data samples and forecasting methods remains a puzzle
- One would expect to find considerable heterogeneity in experts' forecasting ability - this ought to be exploitable by differentiating the weights applied to different forecasts
- In practice, however, individual forecasters' true ability—and consequently the combination weights—are unknown
- Improving on the EW average requires having a procedure for estimating the combination weights which ensures that the sample estimates do not get too far removed from their true but unknown values

Combination Methods in Common Use

Least Squares Approaches

A common approach is to use least squares on the N-vector of forecasts, $\widehat{\mathbf{Y}}_\tau$ using data over the period $\tau = 1, ..., T$:

$$\widehat{\omega}_T = (\sum_{\tau=1}^{T-1} \widehat{\mathbf{Y}}_{\tau+1,\tau} \widehat{\mathbf{Y}}'_{\tau+1,\tau})^{-1} \sum_{\tau=1}^{T-1} \widehat{\mathbf{Y}}_{\tau+1,\tau} Y_{\tau+1}.$$

Different versions of this least squares projection have been proposed. Granger and Ramanathan (1984) consider three regressions:

$$\begin{aligned} (i)\ Y_{t+1} &= \omega_{0t} + \omega_t' \widehat{\mathbf{Y}}_{t+1,t} + \varepsilon_{t+1} \\ (ii)\ Y_{t+1} &= \omega_t' \widehat{\mathbf{Y}}_{t+1,t} + \varepsilon_{t+1} \\ (iii)\ Y_{t+1} &= \omega_t' \widehat{\mathbf{Y}}_{t+1,t} + \varepsilon_{t+1}, \text{ s.t. } \omega_t' \iota = 1. \end{aligned}$$

Combination Methods in Common Use

Least Squares Approaches

- Least-squares procedures require estimating the covariance matrix of the forecast errors, but achieving a precise estimate is difficult due to:
    1. Short and incomplete data samples for individual forecasters
    2. The dimensionality of the problem at hand with a large number of forecasters relative to the length of the time-series
    3. Instability of the covariance matrix reflecting structural breaks, time-varying coefficients or other changes in the underlying data generating process

Combination Methods in Common Use

Least Squares Approaches

- Least-squares procedures require estimating the covariance matrix of the forecast errors, but achieving a precise estimate is difficult due to:
  1. Short and incomplete data samples for individual forecasters
  2. The dimensionality of the problem at hand with a large number of forecasters relative to the length of the time-series
  3. Instability of the covariance matrix reflecting structural breaks, time-varying coefficients or other changes in the underlying data generating process

Combination Methods in Common Use

Least Squares Approaches

- Least-squares procedures require estimating the covariance matrix of the forecast errors, but achieving a precise estimate is difficult due to:
  1. Short and incomplete data samples for individual forecasters
  2. The dimensionality of the problem at hand with a large number of forecasters relative to the length of the time-series
  3. Instability of the covariance matrix reflecting structural breaks, time-varying coefficients or other changes in the underlying data generating process

Combination Methods in Common Use

Least Squares Approaches

- Least-squares procedures require estimating the covariance matrix of the forecast errors, but achieving a precise estimate is difficult due to:
    1. Short and incomplete data samples for individual forecasters
    2. The dimensionality of the problem at hand with a large number of forecasters relative to the length of the time-series
    3. Instability of the covariance matrix reflecting structural breaks, time-varying coefficients or other changes in the underlying data generating process

Combination Methods in Common Use

Shrinkage (e.g., Stock and Watson (2003))

$$\begin{aligned} \omega_{it} &= \psi\hat{\omega}_{it} + (1-\psi)(1/N_t), \\ \psi_t &= \max(0, 1 - \kappa N_t/(T-1-N_t-1)), \end{aligned}$$

- κ regulates the strength of the shrinkage towards equal weights
- As the sample size, T, rises relative to the number of forecasts, N, the least squares estimate gets a larger weight

Combination Methods in Common Use

Shrinkage (e.g., Stock and Watson (2003))

$$
\begin{aligned}
\omega_{it} &= \psi \hat{\omega}_{it} + (1-\psi)(1/N_t), \\
\psi_t &= \max(0, 1 - \kappa N_t/(T - 1 - N_t - 1)),
\end{aligned}
$$

- κ regulates the strength of the shrinkage towards equal weights
- As the sample size, T, rises relative to the number of forecasts, N, the least squares estimate gets a larger weight

Combination Methods in Common Use

Odds matrix approach (Gupta and Wilson (1987))

- Weights are derived from a matrix of pair-wise odds ratios. Each entry in the matrix is interpreted as the odds that forecast i will outperform forecast j
- If the odds matrix is denoted $\mathbf{O}$, then the weight vector, $\mathbf{w}$, is the normalized eigenvector associated with the largest eigenvalue
- The entries of the $\mathbf{O}$ matrix are $o_{ij} = \frac{\pi_{ij}}{\pi_{ji}}$, where $\pi_{ij} = \frac{a_{ij}}{(a_{ij}+a_{ji})}$, and a_{ij} is the number of times forecast i had a smaller absolute error than forecast j in the historical sample. π_{ij} represents the probability that the ith forecast will outperform the jth forecast in the next realization

Combination Methods in Common Use

Odds matrix approach (Gupta and Wilson (1987))

- Weights are derived from a matrix of pair-wise odds ratios. Each entry in the matrix is interpreted as the odds that forecast i will outperform forecast j
- If the odds matrix is denoted $\mathbf{O}$, then the weight vector, $\mathbf{w}$, is the normalized eigenvector associated with the largest eigenvalue
- The entries of the $\mathbf{O}$ matrix are $o_{ij} = \frac{\pi_{ij}}{\pi_{ji}}$, where $\pi_{ij} = \frac{a_{ij}}{(a_{ij}+a_{ji})}$, and a_{ij} is the number of times forecast i had a smaller absolute error than forecast j in the historical sample. π_{ij} represents the probability that the ith forecast will outperform the jth forecast in the next realization

Combination Methods in Common Use

Odds matrix approach (Gupta and Wilson (1987))

- Weights are derived from a matrix of pair-wise odds ratios. Each entry in the matrix is interpreted as the odds that forecast i will outperform forecast j
- If the odds matrix is denoted $\mathbf{O}$, then the weight vector, $\mathbf{w}$, is the normalized eigenvector associated with the largest eigenvalue
- The entries of the $\mathbf{O}$ matrix are $o_{ij} = \frac{\pi_{ij}}{\pi_{ji}}$, where $\pi_{ij} = \frac{a_{ij}}{(a_{ij}+a_{ji})}$, and a_{ij} is the number of times forecast i had a smaller absolute error than forecast j in the historical sample. π_{ij} represents the probability that the ith forecast will outperform the jth forecast in the next realization

Proposed Method

Considerations

1. Individual expert forecasts are often biased and the slope coefficient in a regression of the realized value on individual forecasts often differs from unity (e.g., Zarnowitz (1985); Davies and Lahiri (1995))
2. Bias correction is best done at the level of the combined forecast by including a single intercept and more refined adjustments generally do not lead to improvements
3. Forecasts from data sources such as surveys are generally highly unbalanced which makes standard covariance-based approaches difficult to apply

Proposed Method

Considerations

1. Individual expert forecasts are often biased and the slope coefficient in a regression of the realized value on individual forecasts often differs from unity (e.g., Zarnowitz (1985); Davies and Lahiri (1995))
2. Bias correction is best done at the level of the combined forecast by including a single intercept and more refined adjustments generally do not lead to improvements
3. Forecasts from data sources such as surveys are generally highly unbalanced which makes standard covariance-based approaches difficult to apply

Proposed Method

Considerations

1. Individual expert forecasts are often biased and the slope coefficient in a regression of the realized value on individual forecasts often differs from unity (e.g., Zarnowitz (1985); Davies and Lahiri (1995))
2. Bias correction is best done at the level of the combined forecast by including a single intercept and more refined adjustments generally do not lead to improvements
3. Forecasts from data sources such as surveys are generally highly unbalanced which makes standard covariance-based approaches difficult to apply

Proposed Method

Projection on the Equal-Weighted Mean

$$\tilde{Y}_{t+1,t} = \alpha + \beta \bar{Y}_{t+1,t}$$

- This extension of the EW only requires estimating two parameters, α and β, through least squares regression
- As in the case with EW, information from forecasters with no more than a single data point can be used
- By including a constant, the forecast combination method adjusts for biases that may be present in the individual forecasts as well as in the aggregate
- By allowing for a slope coefficient different from unity the method is likely to work well under a much broader set of scenarios than the simple EW forecast

Proposed Method

Projection on the Equal-Weighted Mean

$$\tilde{Y}_{t+1,t} = \alpha + \beta \bar{Y}_{t+1,t}$$

- This extension of the EW only requires estimating two parameters, α and β, through least squares regression
- As in the case with EW, information from forecasters with no more than a single data point can be used
- By including a constant, the forecast combination method adjusts for biases that may be present in the individual forecasts as well as in the aggregate
- By allowing for a slope coefficient different from unity the method is likely to work well under a much broader set of scenarios than the simple EW forecast

Proposed Method

Projection on the Equal-Weighted Mean

$$\tilde{Y}_{t+1,t} = \alpha + \beta \bar{Y}_{t+1,t}$$

- This extension of the EW only requires estimating two parameters, α and β, through least squares regression
- As in the case with EW, information from forecasters with no more than a single data point can be used
- By including a constant, the forecast combination method adjusts for biases that may be present in the individual forecasts as well as in the aggregate
- By allowing for a slope coefficient different from unity the method is likely to work well under a much broader set of scenarios than the simple EW forecast

Proposed Method

Projection on the Equal-Weighted Mean

$$\tilde{Y}_{t+1,t} = \alpha + \beta \bar{Y}_{t+1,t}$$

- This extension of the EW only requires estimating two parameters, α and β, through least squares regression
- As in the case with EW, information from forecasters with no more than a single data point can be used
- By including a constant, the forecast combination method adjusts for biases that may be present in the individual forecasts as well as in the aggregate
- By allowing for a slope coefficient different from unity the method is likely to work well under a much broader set of scenarios than the simple EW forecast

Proposed Method

Projection on the Equal-Weighted Mean

$$\tilde{Y}_{t+1,t} = \alpha + \beta \bar{Y}_{t+1,t}$$

- This extension of the EW only requires estimating two parameters, α and β, through least squares regression
- As in the case with EW, information from forecasters with no more than a single data point can be used
- By including a constant, the forecast combination method adjusts for biases that may be present in the individual forecasts as well as in the aggregate
- By allowing for a slope coefficient different from unity the method is likely to work well under a much broader set of scenarios than the simple EW forecast

Monte Carlo Experiment

- We compare the real-time forecasting performance of various methods through MC simulations in the context of a common factor model (2 factors) that allows for bias in individual forecasts, dynamic dependencies in the common factors, and heterogeneity in individual forecasters' ability
- All forecasts are one-step-ahead, simulated out-of-sample, and based on recursive parameter estimates using only information available at the time of the forecast

Monte Carlo Experiment

- We compare the real-time forecasting performance of various methods through MC simulations in the context of a common factor model (2 factors) that allows for bias in individual forecasts, dynamic dependencies in the common factors, and heterogeneity in individual forecasters' ability
- All forecasts are one-step-ahead, simulated out-of-sample, and based on recursive parameter estimates using only information available at the time of the forecast

A Common Factor Model

The target variable and the individual forecasts are driven by the following common factor model:

$$
\begin{aligned}
Y_{t+1} &= \mu_y + \beta_{yF}' \mathbf{F}_{t+1} + \varepsilon_{yt+1}, \quad \varepsilon_{yt+1} \sim N(0, \sigma^2_{\varepsilon_Y}) \\
\hat{Y}_{it+1} &= \mu_i + \beta_{iF}' \mathbf{F}_{t+1} + \varepsilon_{it+1}, \;\; \varepsilon_{it+1} \sim N(0, \sigma^2_{\varepsilon_i}), \quad i = 1, ..., N.
\end{aligned}
$$

Cross-sectional heterogeneity in the individual forecasters' performance can be introduced by letting any one of the parameters $(\mu_i, \beta_{iF}, \sigma^2_{\varepsilon_i})$ differ across forecasters

A Common Factor Model

Factor dynamics can be introduced through:

$$\mathbf{F}_t = \mathbf{B}_F \mathbf{F}_{t-1} + \varepsilon_{F_t}, \quad \varepsilon_{F_t} \sim N(\mathbf{0}, \mathbf{D}_{\varepsilon_F}),$$

where $\mathbf{D}_{\varepsilon_F}$ is an $n_f \times n_f$ diagonal matrix with entries:

$$\mathbf{D}_{\varepsilon_F} = \begin{pmatrix} \sigma^2_{F_1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{F_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & & \sigma^2_{F_{n_f}} \end{pmatrix}.$$

This gives the following convenient expressions:

$$\begin{aligned} \sigma^2_y &= \beta'_{yF}(\mathbf{I} - \mathbf{B}^2_F)^{-1}\mathbf{D}_{\varepsilon_F}\beta_{yF} + \sigma^2_{\varepsilon_Y}, \\ \sigma_{y\hat{\mathbf{y}}}[i] &= \beta'_{iF}(\mathbf{I} - \mathbf{B}^2_F)^{-1}\mathbf{D}_{\varepsilon_F}\beta_{yF} \\ \mathbf{\Sigma}_{\hat{\mathbf{y}}\hat{\mathbf{y}}}[i,j] &= \beta'_{iF}(\mathbf{I} - \mathbf{B}^2_F)^{-1}\mathbf{D}_{\varepsilon_F}\beta_{jF} + \mathcal{I}_{\{i=j\}}\sigma^2_{\varepsilon_i}. \end{aligned}$$

Monte Carlo Evidence

Table 1: Simulation results from forecast combinations under factor structure

# of Forecasts	Sample Size	EW	PEW	GR1	GR2	GR3	Shrink 1	Shrink 2	Odds	Previous Best
Experiment 1 : Equal weights summing to one										
4	100	1.000	1.015	1.052	1.046	1.037	1.045	1.043	0.993	1.538
4	1000	1.000	1.002	1.002	1.001	1.000	1.001	1.001	0.999	1.664
20	100	1.000	1.020	1.253	1.236	1.222	1.206	1.129	0.981	4.357
20	1000	1.000	1.002	1.021	1.019	1.019	1.019	1.019	0.998	5.293
Experiment 5: Strong heterogeneity										
4	100	1.000	0.861	0.851	0.843	0.898	0.842	0.841	0.959	0.939
4	1000	1.000	0.840	0.815	0.814	0.882	0.814	0.814	0.966	0.972
20	100	1.000	0.734	0.866	0.851	0.871	0.832	0.799	0.943	0.809
20	1000	1.000	0.705	0.700	0.700	0.725	0.700	0.700	0.951	0.828
Experiment 7: Bias in individual forecasts										
4	100	1.000	0.841	0.872	0.924	1.019	0.923	0.922	0.991	1.123
4	1000	1.000	0.830	0.830	0.890	0.987	0.890	0.890	0.995	1.159
20	100	1.000	0.706	0.871	0.909	1.191	0.887	0.844	0.986	1.084
20	1000	1.000	0.691	0.705	0.742	0.990	0.742	0.742	0.995	1.210

Notes: Results are based on 10,000 simulations. EW: equal-weighted forecast, PEW: projection of actual value on an intercept and EW forecast, GR1: unconstrained OLS, GR2: OLS w/o constant, GR3: OLS w/o constant and weights constrained to add to unity, Shrink1: shrinkage with κ=0.25, Shrink2: shrinkage with κ=1, Odds: Odds ratio approach, Previous Best: forecast from previous best model.

Monte Carlo Evidence

Table 2: Simulation results from forecast combinations under factor structure with survey-like data

# of Forecasts	Sample Size	EW	PEW	GR1	GR2	GR3	Shrink 1	Shrink 2	Odds	Previous Best
Experiment 1 : Equal weights summing to one										
20	100	1.000	1.000	1.040	1.030	1.520	1.030	1.030	1.520	1.540
20	500	1.000	0.986	1.040	1.030	1.510	1.030	1.030	1.510	1.520
20	1000	1.000	0.990	1.030	1.020	1.530	1.020	1.020	1.530	1.550
Experiment 5: Strong heterogeneity										
20	100	1.000	0.563	0.987	0.981	0.976	0.981	0.981	0.976	0.976
20	500	1.000	0.552	0.988	0.983	0.980	0.983	0.983	0.981	0.981
20	1000	1.000	0.557	0.988	0.983	0.977	0.983	0.982	0.977	0.977
Experiment 7: Bias in individual forecasts										
20	100	1.000	0.586	0.987	0.992	0.990	0.992	0.992	0.990	0.991
20	500	1.000	0.564	0.994	0.998	0.997	0.998	0.998	0.996	0.997
20	1000	1.000	0.579	0.995	0.997	0.995	0.997	0.997	0.995	0.995

Notes: Results are based on 10,000 simulations. The minimum number of contiguous observations used by the least squares and shrinkage combinations is 20. EW: equal-weighted forecast, PEW: projection of actual value on an intercept and EW forecast, GR1: unconstrained OLS, GR2: OLS w/o constant, GR3: OLS w/o constant and weights constrained to add to unity, Shrink1: shrinkage with κ=0.25, Shrink2: shrinkage with κ=1, Odds: Odds ratio approach, Previous Best: forecast from previous best model.

Monte Carlo Evidence

The performance of PEW

- The out-of-sample forecasting performance of the projection on the equal-weighted forecast continues to be very good in the unbalanced panel as this approach makes use of the full set of forecasts in the first stage and then adjusts for any biases remaining in the equal-weighted forecasts in the second stage

Empirical Application

Table 3: Empirical application to inflation forecasts from the Survey of Professional Forecasters.[1/]

	1-Step-Ahead		4-Steps-Ahead	
	Revised Data	Real-Time Data	Revised Data	Real-Time Data
RMSE				
EW	0.877	0.903	1.151	1.146
EWc	1.005	0.978	1.008	0.998
PEW, Recursive	0.896	0.899	0.998	0.976
PEW, Rolling	0.788	0.881	0.831	0.926
PEW c	0.903	0.972	0.780	0.996
GR1	2.860	2.921	1.096	1.207
GR2	1.241	1.817	1.120	1.087
GR3	4.499	11.480	1.070	1.113
Shrink 1	1.398	2.918	1.850	1.033
Shrink 2	3.107	7.015	5.503	1.503
Odds	0.892	0.893	1.070	1.092
Previous Best	0.937	0.876	1.053	1.092

1/ Sample: 1979q4-2006q3. Revised data is the last revision as of January 2007, real-time data corresponds to the first revision. The minimum number of contiguous observations required is 10, except for EW, PEW Recursive, and PEW Rolling, where no restriction was imposed. For PEW Rolling a fixed window with 30 observations was used. The number of out-of-sample forecasts equals 77 for 1-step-ahead and 74 for 4-steps-ahead.

Empirical Application

Table 3: Empirical application to inflation forecasts from the Survey of Professional Forecasters.[1/]

	1-Step-Ahead		4-Steps-Ahead	
	Revised Data	Real-Time Data	Revised Data	Real-Time Data
P-Values Giacomini-White Test [2/]				
EW	0.146	0.276	0.000 ***	0.007 ***
EWc	0.021 **	0.079 *	0.000 ***	0.033 **
PEWc	0.015 **	0.265	0.240	0.128
GR1	0.068 *	0.059 *	0.205	0.042 **
GR2	0.001 ***	0.133	0.165	0.301
GR3	0.356	0.366	0.010 **	0.073 *
Shrink 1	0.071 *	0.290	0.308	0.590
Shrink 2	0.210	0.293	0.279	0.174
Odds	0.059 *	0.225	0.000 ***	0.113
Previous Best	0.037 **	0.244	0.000 ***	0.251

* $p<0.10$. ** $p<0.05$. *** $p<0.01$.

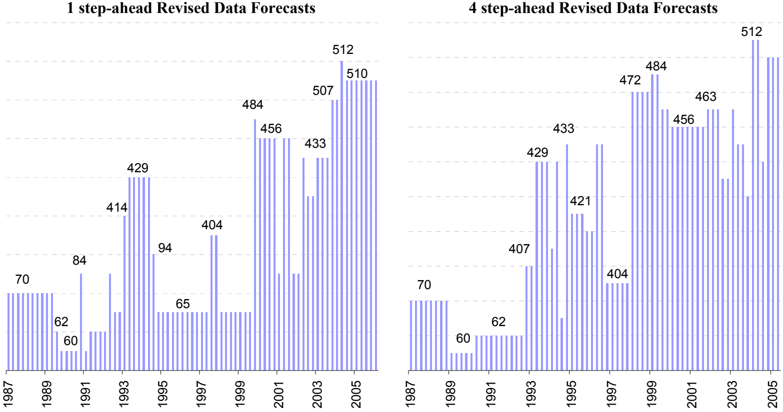
1/ Sample: 1979q4-2006q3. Revised data is the last revision as of January 2007, real-time data corresponds to the first revision. The minimum number of contiguous observations required is 10, except for EW, PEW Recursive, and PEW Rolling, where no restriction was imposed. For PEW Rolling a fixed window with 30 observations was used. The number of out-of-sample forecasts equals 77 for 1-step-ahead and 74 for 4-steps-ahead.

2/ Computed with respect to PEW Rolling. Test is conditional on the (first/fourth) lag of the difference of the losses.

Empirical Application

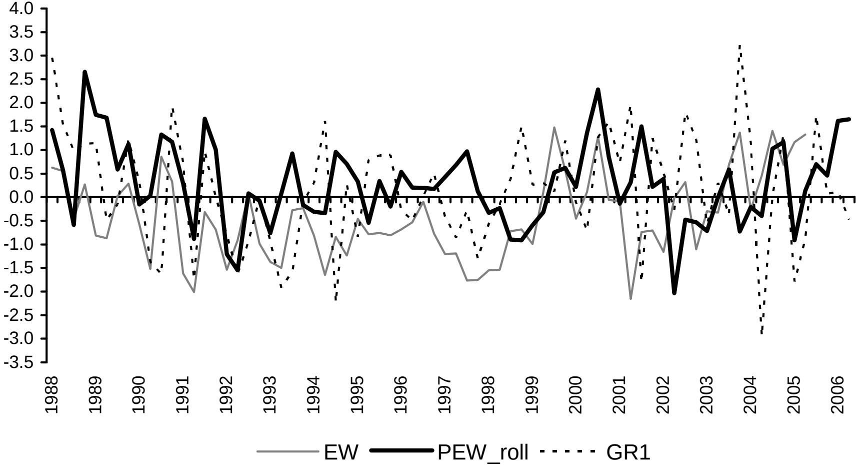
Figure 2: Previous Best Forecaster

a) Time-Series



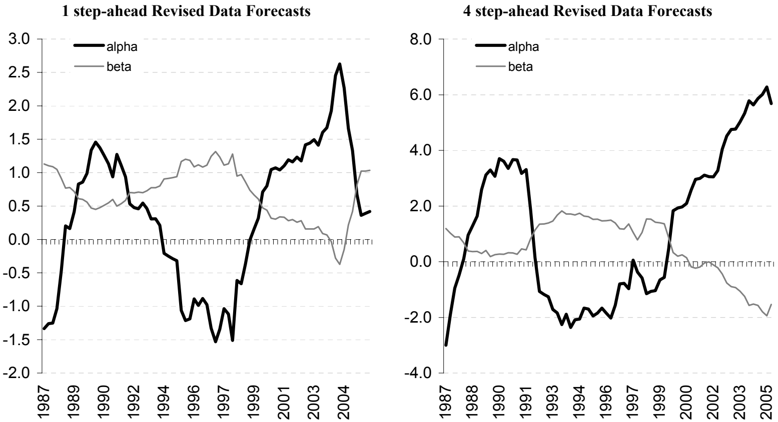
Empirical Application

Figure 2: Forecast Errors, Four-Steps-Ahead, Real-Time Data



Empirical Application

Figure 4: Estimated Parameters of PEW Rolling



Conclusions

- The unbalanced panel structure of survey data means that the real-time performance of combination methods that require estimating the full covariance between the experts' forecasts deteriorates relative to that of more robust methods such as equal-weighting
- Successful schemes for real-time combination of expert forecasts achieve a favorable trade-off between the bias of using sub-optimal combination weights and the effect of parameter estimation error of using estimated combination weights
- We propose a new combination method that projects the outcome variable on a constant and the equal-weighted forecast. This approach uses information in the full set of individual forecasts (incorporated into the equal-weighted average) but then adjusts for possible bias and noise in this aggregate forecast.

Conclusions

- The unbalanced panel structure of survey data means that the real-time performance of combination methods that require estimating the full covariance between the experts' forecasts deteriorates relative to that of more robust methods such as equal-weighting
- Successful schemes for real-time combination of expert forecasts achieve a favorable trade-off between the bias of using sub-optimal combination weights and the effect of parameter estimation error of using estimated combination weights
- We propose a new combination method that projects the outcome variable on a constant and the equal-weighted forecast. This approach uses information in the full set of individual forecasts (incorporated into the equal-weighted average) but then adjusts for possible bias and noise in this aggregate forecast.

Conclusions

- The unbalanced panel structure of survey data means that the real-time performance of combination methods that require estimating the full covariance between the experts' forecasts deteriorates relative to that of more robust methods such as equal-weighting
- Successful schemes for real-time combination of expert forecasts achieve a favorable trade-off between the bias of using sub-optimal combination weights and the effect of parameter estimation error of using estimated combination weights
- We propose a new combination method that projects the outcome variable on a constant and the equal-weighted forecast. This approach uses information in the full set of individual forecasts (incorporated into the equal-weighted average) but then adjusts for possible bias and noise in this aggregate forecast.