Span. Econ. Rev. 1, 91–121 (1999)

Spanish Economic Review

A dynamic factor model framework for forecast combination

Yeung Lewis Chan[1], **James H. Stock**[2], **Mark W. Watson**[3]

[1] Department of Economics, Harvard University, Cambridge, MA 02138, USA
[2] Kennedy School of Government and NBER, 79 John F. Kennedy Street, Harvard University, Cambridge, MA 02138, USA
[3] Woodrow Wilson School, Princeton University, Princeton, NJ 08544, USA

Abstract. A panel of ex-ante forecasts of a single time series is modeled as a dynamic factor model, where the conditional expectation is the single unobserved factor. When applied to out-of-sample forecasting, this leads to combination forecasts that are based on methods other than OLS. These methods perform well in a Monte Carlo experiment. These methods are evaluated empirically in a panel of simulated real-time computer-generated univariate forecasts of U.S. macroeconomic time series.

JEL classification: C32, C22

Key words: Combination forecasts, principal component regression, James-Stein estimation

1 Introduction

A provocative finding in the literature on combination forecasts is that, in empirical situations with real forecasts, simple averages of forecasts often provide substantial improvements over individual forecasts and, even more strikingly, simple averages often outperform more sophisticated statistical combination methods; see for example Clemen and Winkler (1986), Guerard and Clemen (1989), Diebold and Pauly (1990), and the review articles by Granger (1989), Clemen (1989) and Diebold and Lopez (1995). This poses a puzzle, because the theoretical work on combination forecasts by Bates and Granger (1969), Granger and Ramanathan (1984), and others argues that statistical methods such as ordinary least squares (OLS) regression should improve upon simple averaging.

The authors thank Frank Diebold, Bruce Hansen, Serena Ng, and Norm Swanson for helpful suggestions and discussions. This research was supported in part by National Science Foundation grants no. SBR-9409629 and SBR-9730489.

This paper provides some new theoretical results that shed some light on this puzzle and which suggest alternative statistical methods for improving combination forecasts. These results follow from modeling the panel of individual forecasts as an approximate dynamic factor model, where the unobserved factor is the true conditional expectation. The dynamic factor structure has important implications for estimation of the forecast combination weights. Assuming that the loss is the squared out-of-sample forecast error (MSFE), we show that the optimal estimator of the linear combination weights minimizes a particular weighted average of the elements of the variance-covariance matrix of the estimator. Examination of this expression suggests that this weighted average will be similar to the trace of the covariance matrix, at least in some limiting cases. This suggests that, under this loss function, OLS is no longer admissable when there are at least three forecasts to be combined. Instead, alternative estimation methods, such as James-Stein (1961) estimation, ridge regression, and methods that explicitly exploit the dynamic factor structure can in theory provide significant improvements over simple average and, especially, OLS combination forecasts. Clemen and Winkler (1986) and Diebold and Pauly (1990) have appealed to Bayesian arguments to justify the use of shrinkage-type estimators for combination forecasting. In contrast, our results provide a theoretical justification grounded in classical statistical theory for the use of such estimators.[1]

In Sect. 3, these implications are examined in a Monte Carlo experiment. The results are largely consistent with the theoretical predictions. We find that a wide range of methods produce large improvements over OLS combinations, particularly when the number of forecasts to be combined is large. Motivated by the dynamic single factor model and theoretical results in Stock and Watson (1998b), principal components regression is also considered, and in this Monte Carlo experiment it is found to produce most frequently the best combination forecasts of all methods considered. Not surprisingly, a robust combination estimator, the median forecast, has lower risk than the simple average when the distribution of forecasts has some large outliers.

These theoretical predictions are examined in a panel of forecasts in Sect. 4. This dataset consists of univariate forecasts for 215 U.S. monthly macroeconomic time series. The dataset is taken from Stock and Watson (1998a), where it was developed for the purpose of comparing linear vs. nonlinear univariate forecasting methods. For each series, the panel consists of 49 monthly forecasts, computed in a way to simulate real-time forecasting (ex-ante forecasts). The best combination forecasts provide considerable improvements over the best individual forecasts, and many of the combination methods provide similar good performance. However, some puzzles remain.

[1] Figlewski (1983) and Figlewski and Urich (1983) proposed using a single-factor model for the combination of forecasts. Our formulation differs from theirs in several ways. Most importantly, they model the forecast errors by a factor model while we focus on the forecasts themselves. Our focus is also different as we use the model to simplify the risk function to suggest improvements to OLS combining methods.

2 Theoretical framework

2.1 Risk for a panel of forecasts

Notation and assumptions: Let y_t, $t = 1, \ldots, T$ be a strictly stationary univariate time series with unconditional mean zero. Let F_s^t denote the information set consisting of $\{y_s, y_{s+1}, \ldots, y_t, Z_s, Z_{s+1}, \ldots, Z_t\}$, where Z_t is a multiple time series available for forecasts made at date t. If the forecasts are univariate, $\{Z_t\}$ is absent from this information set.

We focus on mean squared forecast error loss, and for notational convenience, this section only considers one-step ahead forecasting. The loss from using a forecast f_t to predict y_{t+1} is,

$$L(f_t, y_{t+1}) = (y_{t+1} - f_t)^2 \tag{2.1}$$

and the risk is,

$$R(f_t) = E(y_{t+1} - f_t)^2 \ . \tag{2.2}$$

Under (2.2), the optimal forecast of y_{t+1} given $F_{-\infty}^t$ is its conditional expectation, $E(y_{t+1}|F_{-\infty}^t) \equiv \mu_t$, and y_{t+1} can be written,

$$y_{t+1} = \mu_t + \epsilon_{t+1} \ , \tag{2.3}$$

where $E(\epsilon_{t+1}|F_{-\infty}^t) = 0$.

We consider combination forecasts based on a panel of m individual one-step ahead forecasts of y_{t+1}, the i-th of which is denoted $y_{t+1|t,i}$. We make the following assumptions on these forecasts: (1) the forecasts $y_{t+1|t,i}$ are assumed to be true ex-ante forecasts so that they are a function only of F_1^t; (2) the forecasts can be written,

$$y_{t+1|t,i} = g_i(F_{t-p}^t, \hat{\theta}_i) \ , \tag{2.4}$$

where $p < t$ and $\hat{\theta}_i$ is a finite-dimensional parameter vector that is either imposed *a-priori* or is estimated using F_1^t; (3) the forecasts are unconditionally unbiased, that is, $Ey_{t+1|t,i} = Ey_{t+1}$; and (4) neither the data nor the forecasts are subject to revision.

These assumptions are not very restrictive. Assumption (1) is just an assumption about the timing of forecasts.

Assumption (2) is general enough to include model-based forecasts with econometrically estimated parameters, model-based forecasts that incorporate data-based selection criteria such as the AIC, model-based forecasts with judgmentally imposed parameters, and partially or entirely judgmental forecasts. Importantly, it is not assumed that g_i is correctly specified, so that in general g_i will not generate the optimal forecast. The most restrictive aspect of (2.4) is that the forecasting method (the function g_i) is assumed to be constant over time.

Assumption (3) is readily relaxed, at the cost of additional notation, by allowing $E g_i \neq 0$ and including intercept terms in the relevant expressions. Relaxing assumption (4) would greatly increase the notational complexity but would not change the basic argument of this section.

Given these assumptions, it is possible to decompose the forecast error of the i-th forecast into three terms. Let θ_i^0 denote the value of θ_i that minimizes (2.2) when f_t has the form $g_i(F_{t-p}^t, \theta_i)$. This can be thought of as the true value of θ_i in the sense that it is the unknown value of θ that provides the forecast with the lowest risk (expected MSFE) among all forecasts with this functional form. With this notation, we can write,

$$y_{t+1|t,i} = \mu_t + \zeta_{it} + \nu_{it} \,, \tag{2.5}$$

where $\zeta_{it} = g_i(F_{t-p}^t, \theta_i^0) - \mu_t$ and $\nu_{it} = g_i(F_{t-p}^t, \hat{\theta}_i) - g_i(F_{t-p}^t, \theta_i^0)$.

The expression (2.5) represents the difference between the i-th forecast and the optimal forecast as the sum of two components. The first, ζ_{it}, represents model specification error; if the i-th family of models contains the true conditional expectation function, then $\zeta_{it} = 0$, but if there is model specification error (which in practice there will be), $\zeta_{it} \neq 0$ for at least some t. The second component, ν_{it}, represents estimation error. If θ^0 is a parameter that is consistently estimated by nonlinear least squares, and if g_i has a continuous first derivative, then $\text{corr}(\zeta_{it}, \nu_{it}) = 0$ to first order as a consequence of the first order conditions for the estimation problem.

In general, μ_t will be correlated with $\zeta_{it} + \nu_{it}$. For example, if the i-th forecast tends to be less than the optimal forecast when the optimal forecast is large, either because of specification or estimation error, then this correlation will be negative. Because of this correlation, $E(y_{t+1|t,i}|\mu_t) \neq \mu_t$. If $(\mu_t, \zeta_{it} + \nu_{it})$ are jointly normally distributed, then this conditional expectation is linear, and (2.5) implies,

$$y_{t+1|t,i} = \lambda_i \mu_t + e_{it} \,, \tag{2.6}$$

where $\lambda_i = 1 + \text{cov}(\zeta_{it} + \nu_{it}, \mu_t)/\text{var}(\mu_t)$ and where e_{it} is normally distributed and is independent of μ_t. If $(\mu_t, \zeta_{it} + \nu_{it})$ are not jointly normal, then (2.6) still holds with $\text{corr}(\mu_t, e_{it}) = 0$, but $\lambda_i \mu_t$ is interpreted as the linear projection of $y_{t+1|t}$ on μ_t rather than the conditional expectation and e_{it} is not normally distributed. It is convenient to write (2.6) in matrix notation:

$$Y_{t+1|t} = \Lambda \mu_t + e_t \,, \tag{2.7}$$

where $Y_{t+1|t}$ denotes the m-vector of forecasts with i-th element $y_{t+1|t,i}$, Λ is a m-vector with i-th element λ_i, and $e_t = (e_{1t}, \ldots, e_{mt})'$.

Risk for linearly combined forecasts: The primary focus of this paper is on the construction of out-of-sample forecast for y_{T+1} using a combination of the individual forecasts. Here we consider the linear combination forecast,

$$\phi_T(\beta) = \beta' Y_{T+1|T} \,. \tag{2.8}$$

Because y_t is assumed to have mean zero, the intercept is excluded from (2.8), although in practice one could be added.

Two questions arise in considering (2.8): given the structure (2.7), what is the optimal population value of β, and what is the best way to choose β when one has a panel of forecasts?

It is straightforward to derive the optimal value of β, β_0, using the representation (2.7). Under squared error loss the risk (2.2) is

$$\begin{aligned} R(\phi_T(\beta)) &= E(y_{T+1} - \beta' Y_{T+1|T})^2 \\ &= E[\epsilon_{T+1} + \mu_T - \beta'(\Lambda\mu_T + e_T)]^2 \\ &= \sigma_\epsilon^2 + (\beta'\Lambda - 1)^2\sigma_\mu^2 + \beta'\Sigma_e\beta \ , \end{aligned} \tag{2.9}$$

where $\sigma_\mu^2 = E\mu_t^2$ and $\Sigma_e = Ee_te_t'$, and where the final expression follows because ϵ_{t+1}, μ_t, and e_t are all mutually uncorrelated. Solving the minimization problem yields,

$$\beta_0 = [\Sigma_e + \sigma_\mu^2\Lambda\Lambda']^{-1}\sigma_\mu^2\Lambda \tag{2.10}$$

as the vector of weights that minimizes the risk (2.2) among all combination forecasts of the form (2.8). This result confirms the general finding of the forecast combination literature that all forecasts, even poor ones, in general receive some weight in optimal combination forecasts.

Because β_0 is unknown, the linear combination forecast based on β_0 is infeasible. Let $\tilde{\beta}$ be a feasible weighting vector (either a function of data through date T or a constant vector imposed a-priori) so that $\tilde{\beta}'Y_{T+1|T}$ is a feasible linear combination forecast. Now

$$\begin{aligned} R(\tilde{\beta}'Y_{T+1|T}) &= R(\phi_T(\beta_0)) \\ &\quad + E[(\tilde{\beta} - \beta_0)'Y_{T+1|T}Y_{T+1|T}'(\tilde{\beta} - \beta_0)] \\ &\quad - 2\mathrm{cov}[(\tilde{\beta} - \beta_0)'Y_{T+1|T}, y_{T+1} - \beta_0'Y_{T+1|T}] \ . \end{aligned} \tag{2.11}$$

The term $R(\phi_t(\beta_0))$ does not depend on $\tilde{\beta}$. When $\tilde{\beta}$ is estimated from past data, for example if $\tilde{\beta}$ is estimated by OLS, then the correlation between $\tilde{\beta}$ and both y_{T+1} and $Y_{T+1|T}$ can be expected to be small. If these correlations are zero, the last term on the right hand side of (2.11) vanishes and the expression simplifies to,

$$R(\tilde{\beta}'Y_{T+1|T}) = \mathrm{tr}(V_{\tilde{\beta}}EY_{T+1|T}Y_{T+1|T}') + \gamma \ , \tag{2.12}$$

where γ does not depend on $\tilde{\beta}$, $V_{\tilde{\beta}} = E(\tilde{\beta} - \beta_0)(\tilde{\beta} - \beta_0)'$ and $\mathrm{tr}(\bullet)$ denotes the trace. The representation (2.7) permits additional simplification of (2.12),

$$R[\tilde{\beta}'Y_{T+1|T}] = \Lambda'V_{\tilde{\beta}}\Lambda\sigma_\mu^2 + \mathrm{tr}(V_{\tilde{\beta}}\Sigma_e) + \gamma \ . \tag{2.13}$$

2.2 Approximate factor model of the forecasts

Expression (2.7) represents each forecast as having two components, the unobserved conditional mean (times the weight λ_i) and an idiosyncratic component, e_{it}. If some additional restrictions hold on the distribution of these idiosyncratic components, then this representation constitutes a dynamic factor model for the panel of forecasts. Because (2.7) was obtained under weak conditions, these stronger, factor model restrictions need not hold for any particular panel of forecasts. Nevertheless, because factor analytical methods provide a convenient way to handle large numbers of series that have a common element (which is the case here), it is useful to explore the statistical consequences of (2.7) satisfying these further restrictions so that the panel of forecasts has a factor model representation.

Two such sets of restrictions appear in the literature on dynamic factor models. Under the stronger of the two, the idiosyncratic errors $\{e_{it}\}$ are mutually uncorrelated at all leads and lags ($Ee_{it}e_{js} = 0, i \neq j$, all t, s) and are uncorrelated with the factor ($Ee_{it}\mu_s = 0$, all i, t, s). This yields the exact dynamic factor model studied by Geweke (1977) and Sargent and Sims (1977). In the forecasting environment studied here, there is no reason that these strong conditions should be satisfied. However, when the number of forecasts (m) is large, many factor analytic methods continue to be useful under weaker conditions. These weaker conditions essentially permit some cross-sectional dependence among the idiosyncratic errors, and between these terms and lags of μ_t, and conditions like these characterize so-called approximate factor models; see Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988, 1993), Forni and Reichlin (1995), Stock and Watson (1998b). This suggests that even if the exact dynamic factor model assumptions do not strictly hold, that model may still yield valuable insights into the forecast combination problem studied here.

The expressions for the optimal linear combination (2.10) and the risk of a feasible linear combination forecast simplify when the panel of forecasts has an exact dynamic factor representation. If the idiosyncratic errors have the same variances, so $\Sigma_e = \sigma_e^2 I_m$ (where I_m is the $m \times m$ identity matrix), then (2.10) and (2.13) become,

$$\beta_0 = [\sigma_e^2 I_m + \Lambda\Lambda'\sigma_\mu^2]^{-1}\sigma_\mu^2\Lambda \tag{2.14}$$

$$R[\tilde{\beta}' Y_{T+1|T}] = \sigma_\mu^2 \Lambda' V_{\tilde{\beta}} \Lambda + \sigma_e^2 tr V_{\tilde{\beta}} + \gamma\ . \tag{2.15}$$

In the special case that $\Lambda = (1, \dots, 1)' \equiv \iota'$, all forecasts are equally accurate and $\beta_0 = [\sigma_e^2 I_m + \iota\iota'\sigma_\mu^2]^{-1}\sigma_\mu^2\iota$, so that all forecasts receive equal weight. Even in this case, however, the elements of β_0 in general differ from $1/m$, and approach $1/m$ only in the limit that $\sigma_e^2/\sigma_\mu^2 \to 0$ (so that all forecasts equal the conditional expectation).

2.3 Implications for the estimation of β

The expressions for the risk of the feasible linear combination forecast provide some guidance about the optimal estimation of $\tilde{\beta}$. In the general case (2.13), the risk is weighted average of the elements of the variance covariance matrix of $\tilde{\beta}$, where the weights are a function of the unknown parameters of the covariance structure of the forecasts. Although this risk function depends on unknown nuisance parameters, when m is large this risk is *not* the risk that leads to estimation by ordinary least squares β (cf. Lehmann [1983]). Thus there is no reason to expect OLS estimation of β to provide good combination forecasts for m sufficiently large.[2]

The special case (2.15) provides more concrete guidance. If σ_e^2 is sufficiently large, the expression is dominated by $\mathrm{tr}V_{\tilde{\beta}}$. When $m \geq 3$ and the risk is $\mathrm{tr}V_{\tilde{\beta}}$, OLS is inadmissable. Intuitively, the inadmissability of OLS arises because, for $m \geq 3$, the loss function $\mathrm{tr}V_{\tilde{\beta}}$ allows for a tradeoff between bias and variance. Among unbiased linear estimators, OLS remains efficient in the usual sense of multivariate regression. However, allowing estimators that shrink towards particular parameter values, and thus have bias but smaller variance, can reduce $\mathrm{tr}V_{\tilde{\beta}}$. The OLS estimator is dominated by James-Stein estimators (Stein 1955; James and Stein 1961). A related technique is ridge regression, which is similar to Bayesian shrinkage estimation. A common concern with these techniques in practice is that they entail shrinkage towards a prespecified parameter vector, which is in general unknown and unavailable. In the current application, (2.15) and the subsequent discussion suggests shrinkage towards a vector of equal weights. This provides a formal justification for this intuitive notion that is prevalent in the forecast combination literature (cf. Diebold and Lopez 1995).

The other implication of (2.15) for the estimation of β is that improved estimates might be obtained by exploiting the potential factor structure of the forecasts. Stock and Watson (1998b) consider the estimation of the factors in approximate dynamic factor models. An estimator for which they provide theoretical results, principal components regression, is discussed in the next section. Principal components regression is typically motivated as an ad-hoc device for solution of multicollinearity. With this as motivation, latent root regression (essentially principal components regression) was applied to forecast combination by Guerard and Clemen (1989). The factor analytic structure (2.7) provides a formal reason to expect multicollinearity in the panel of forecasts. However, the forecasts could still have the factor analytic structure, and thus principal components regression could apply, even if the forecasts are themselves only moderately correlated. Thus multicollinearity per se is not relevant to our motivation for this method. Moreover, the factor model representation and the results in Stock and Watson (1998b) indicate that only the first principal component is needed for forecasting, at least when m is large.

Because the risk function depends on the unknown parameters of the factor model, it appears that the performance of various estimators must be evaluated numerically for a particular set of parameters. Such an evaluation is conducted using Monte Carlo techniques in the next section.

[2] A similar point is made by Amemiya and Morimune (1974) in their study of order determination in distributed lag models. They consider a risk function which is essentially (2.11) and discuss how this provides motivation for considering estimators other than OLS; also see the discussion in Hendry et al. (1984), p. 1063.

3 Monte Carlo analysis

3.1 Design

The objectives of the Monte Carlo analysis reported in this section are to evaluate the risk of various linear combination forecasting methods with Gaussian errors, to consider the effect of heavy tails in the error distribution on this risk, and to examine the performance of a nonlinear forecast combination method that is robust to heavy tailed error distributions (the median). This is done by examining the performance of various feasible and infeasible combination forecasts within a dynamic factor model. In most cases, the dynamic factor model is exact. In some trials, however, the sensitivity of these results to the exact dynamic factor specification is explored by allowing the factor loadings to change over time.

Design. The variable to be forecast, y_t, and the panel of forecasts are taken to follow a factor model with time-varying factor loadings,

$$y_{t+1} = \mu_t + \epsilon_{t+1} \tag{3.1}$$

$$y_{t+1|t,i} = \lambda_{it}\mu_t + e_{it} \tag{3.2}$$

$$\lambda_{it} = \lambda_{it-1} + \zeta_{it} \tag{3.3}$$

where ϵ_{t+1} is i.i.d. $N(0,1)$, μ_t is i.i.d. $N(0,\sigma_\mu^2)$, ζ_{it} is i.i.d. $N(0,\sigma_\zeta^2)$, and e_{it} has the mixture of normals distribution, where e_{it} is distributed $N(0,\sigma_e^2)$ with probability $1-\pi$ and is distributed $N(0,25\sigma_e^2)$ with probability π. The random variables $(\mu_t, e_{it}, \zeta_{it}, \epsilon_{t+1})$ are mutually independent.

The initial factor loadings $\{\lambda_{i0}\}$ vary across series. To generate this dispersion, $\{\lambda_{i0}\}$ were drawn independently from a $N(\bar{\lambda},\sigma_\lambda^2)$ distribution, and are distributed independently of $(\mu_t, e_{it}, \zeta_{it}, \epsilon_{t+1})$. In most cases, $\sigma_\zeta^2 = 0$, so the factor loadings are constant over time ($\lambda_{it} = \lambda_{i0}$, all t); in this case, the panel of forecasts follows an exact factor model.

The free parameters which vary in the Monte Carlo design are the sample size T, the number of forecasts m, $\bar{\lambda}$, σ_λ^2, σ_e^2, σ_μ^2, σ_ζ^2, and π.

The risk of a particular estimator, say $\tilde{\beta}$, was computed as follows: (i) $\{\lambda_{it}\}$ were drawn, $i = 1,\ldots,m$, $t = 0,\ldots,T+r$; (ii) realizations of y_{t+1} and $y_{t+1|t,i}$ were drawn for $t = 1,\ldots,T+r$; (iii) $\tilde{\beta}$ was computed using the first T of these realizations; (iv) $\phi_t(\tilde{\beta}) = \tilde{\beta}'Y_{t+1|t}$ was computed for $t = T+1,\ldots,T+r$; and (v) a Monte Carlo realization of the risk $R(\tilde{\beta})$ was computed as $r^{-1}\sum_{T+1}^{T+r}[y_{t+1}-\phi_t(\tilde{\beta})]^2$. For each estimator and set of parameter values, steps (i)–(v) were repeated for 10,000 Monte Carlo repetitions with $r = 10$.

3.2 Combination forecasts

Six forecast combination schemes were considered: equal weighting; OLS; the James-Stein estimator; ridge regression; a method based on preliminary estimation of the factor; and the median. In addition, as benchmarks two infeasible combination forecasts were computed: the conditional mean, μ_t, and the optimal linear combination forecast.

Optimal infeasible linear combination. The optimal infeasible linear combination forecast is $\beta_{0t}' Y_{T+1|T}$, where β_{0t} is obtained from (2.14) with $\Lambda_t = (\lambda_{1t}, \ldots, \lambda_{mt})'$ replacing Λ_t. This is the minimum MSFE linear combination forecast when Λ_t is known. Because Λ_t is known in the Monte Carlo design, but is unknown in applications, this combination forecast is infeasible in applications.

Equal weighting. For the equal-weighting estimator, the weight vector is fixed a-priori to be $\beta^{\text{equal}} = (1/m, \ldots, 1/m)$.

OLS. Another combination method that has received considerable attention is estimation of the weights by OLS. The OLS weighting vector, $\hat{\beta}^{\text{OLS}}$, is the OLS estimator of β in the regression,

$$y_{t+1} = \beta' Y_{t+1|t} + u_{t+1} \ , \tag{3.4}$$

where u_{t+1} is an error term.

James-Stein estimator. The discussion of the risk function in Sect. 2.2 pointed toward alternative estimation methods, one of which is James-Stein estimation. The specific James-Stein estimator considered is the optimal (minimum risk) James-Stein estimator, shrunk toward equal weighting (e.g. Judge et al. 1980, p. 68):

$$\hat{\beta}^{\text{JS}} = \beta^{\text{equal}} + \{1 - [(m-2)/(T-m+2)]W^{-1}\} \left(\hat{\beta}^{\text{OLS}} - \beta^{\text{equal}}\right) \ , \tag{3.5}$$

where $W = (\hat{\beta}^{\text{OLS}} - \beta^{\text{equal}})'(\sum_{t=1}^{T} Y_{t+1|t} Y_{t+1|t}')(\hat{\beta}^{\text{OLS}} - \beta^{\text{equal}}) / \sum_{t=1}^{T}(y_{t+1} - \hat{\beta}^{\text{OLS}\prime} Y_{t+1|t})^2$.

Ridge regression. The ridge regression estimator, shrunk towards equal weighting, is,

$$\hat{\beta}^{\text{RR}} = \left(cI_m + \sum_{t=1}^{T} Y_{t+1|t} Y_{t+1|t}'\right)^{-1} \left(\sum_{t=1}^{T} Y_{t+1|t} y_{t+1} + c\beta^{\text{equal}}\right) . \tag{3.6}$$

The parameter c was set to $c = k\,\text{tr}(\sum_{t=1}^{T} Y_{t+1|t} Y_{t+1|t}')/m$, where $k = 0.1$, 0.5, and 1.0. A value of $k = 1$ can be considered large in the sense that the weight on the identity matrix equals the average diagonal element in the regressor moment matrix.

Principal component (factor analytic) forecasts. The factor analytic forecasts exploit the factor structure of the panel of forecasts first to compute an estimator of μ_t. Stock and Watson (1998b) show that, under general conditions that include time varying factor loadings, the principal component of the forecast second moment matrix will be a uniformly consistent estimator of μ_t as the sample size and m get large, and OLS regression with this principal component will produce asymptotically efficient forecasts. The principal component forecast is produced by first estimating Λ by the eigenvector corresponding to the largest eigenvalue of $T^{-1}\sum_{t=1}^{T} Y_{t+1|t}Y'_{t+1|t}$; call this $\hat{\Lambda}$. Then the factor μ_t is estimated as $\hat{\mu}_t = \hat{\Lambda}'Y_{t+1|t}$. The next step is to use OLS to estimate α in the regression,

$$y_{t+1} = \alpha\hat{\mu}_t + \text{error}_{t+1} \tag{3.7}$$

over the period $t = 1, \ldots, T-1$. The factor analytic linear combination is thus given by,

$$\hat{\beta}^{\text{PC}} = \hat{\alpha}\hat{\Lambda}\ . \tag{3.8}$$

Median. With nongaussian errors, linear combination estimators no longer need be optimal, and in particular for distributions with sufficiently heavy tails robust estimators such as the median are more efficient than the simple average. The final combination estimator considered is therefore the median forecast at date t.

3.3 Results

The risks of the various combination forecasts for different sets of design parameters are summarized in Table 1. The risks are relative to the variance of ϵ_t, which is the risk of the infeasible forecast based on the conditional mean μ_t; thus the efficient forecast has a relative risk of 1.00.

First consider the results for $T = 100$ and constant factor loadings ($\sigma_\zeta = 0$). These results are generally consistent with the theoretical analysis of the previous section. When $m = 2$, OLS is typically as efficient as the most efficient forecast, or nearly so. However, the relative performance of OLS deteriorates sharply as the number of forecasts in the panel grows. For $T = 100$ and $m = 20$, for example, the relative risk of the OLS forecast is never less than 1.31, while the relative risk of the other forecasts is typically 1.1 or less.

For $m \geq 10$, the equal-weighted forecasts do very well overall in this comparison. Indeed, the equal-weighted forecast performs well for large m even when $\bar{\lambda} = 0.6$ and $\sigma_\lambda = 0.15$; this is surprising because as $\bar{\lambda}$ decreases, the optimal weight vector (β_0) is far from $1/m$. For small m, the estimation methods such as OLS have fewer parameters to estimate, and for $m = 2$ equal-weighting typically does much worse than any of the estimation methods.

The risk of the median is nearly as low as the risk of the mean in the cases that $\pi = 0$, and as expected the gap decreases as π increases. Indeed, the median has the smallest risk of all combination forecasts in the heavy-tailed case ($\pi = 0.05$) for $m \geq 5$.

The ridge regression and James-Stein forecasts also have low risks, especially for large m. The ridge regression forecast with the greatest shrinkage towards equal weighting ($k = 1.0$) generally has a smaller risk than the ridge regression forecasts with less shrinkage.

The PC forecast performs very well, for both large and small m; in the cases considered in Table 1, the PC forecast most frequently has the smallest risk. Its performance is relatively best when the factor weights are not close to one

Table 1. Monte Carlo results: Relative risks of various combination forecasts

T	m	$\bar{\lambda}$	σ_λ	σ_e	σ_μ	π	σ_ζ	InfeasLC	Equal	OLS	J-S	RR(0.1)	RR(0.5)	RR(1.0)	PC	Median
100	2	1.00	0.00	1.00	1.00	0.00	0.00	1.342	1.510	1.368	1.368	1.363	1.363	1.375	1.357*	1.510
100	5	1.00	0.00	1.00	1.00	0.00	0.00	1.172	1.204	1.234	1.210	1.216	1.193	1.186*	1.186	1.291
100	10	1.00	0.00	1.00	1.00	0.00	0.00	1.089	1.098*	1.218	1.124	1.177	1.126	1.110	1.101	1.138
100	20	1.00	0.00	1.00	1.00	0.00	0.00	1.047	1.050*	1.307	1.082	1.206	1.103	1.076	1.059	1.072
100	30	1.00	0.00	1.00	1.00	0.00	0.00	1.031	1.032*	1.476	1.071	1.274	1.109	1.070	1.042	1.049
100	2	1.00	0.15	1.00	1.00	0.00	0.00	1.334	1.504	1.361	1.361	1.356	1.355	1.368	1.349*	1.504
100	5	1.00	0.15	1.00	1.00	0.00	0.00	1.168	1.207	1.233	1.210	1.215	1.192	1.185	1.181*	1.294
100	10	1.00	0.15	1.00	1.00	0.00	0.00	1.088	1.102	1.212	1.125	1.172	1.123	1.109	1.101*	1.142
100	20	1.00	0.15	1.00	1.00	0.00	0.00	1.049	1.052*	1.316	1.086	1.212	1.106	1.079	1.059	1.074
100	30	1.00	0.15	1.00	1.00	0.00	0.00	1.033	1.035*	1.482	1.078	1.277	1.110	1.070	1.042	1.053
100	2	1.00	0.15	1.00	2.00	0.00	0.00	1.444	1.538	1.473	1.473	1.464	1.461	1.465	1.460*	1.538
100	5	1.00	0.15	1.00	2.00	0.00	0.00	1.184	1.212	1.249	1.220	1.219	1.200	1.197	1.197*	1.308
100	10	1.00	0.15	1.00	2.00	0.00	0.00	1.097	1.110	1.221	1.132	1.152	1.116	1.111	1.109*	1.150
100	20	1.00	0.15	1.00	2.00	0.00	0.00	1.048	1.053*	1.312	1.085	1.146	1.073	1.063	1.058	1.078
100	30	1.00	0.15	1.00	2.00	0.00	0.00	1.033	1.037*	1.478	1.076	1.174	1.067	1.052	1.044	1.055
100	2	0.90	0.15	1.00	1.00	0.00	0.00	1.386	1.524	1.415	1.415	1.409	1.408	1.417	1.404*	1.524
100	5	0.90	0.15	1.00	1.00	0.00	0.00	1.198	1.216	1.263	1.232	1.246	1.222	1.214	1.214*	1.302
100	10	0.90	0.15	1.00	1.00	0.00	0.00	1.111	1.114*	1.235	1.140	1.197	1.147	1.131	1.122	1.155
100	20	0.90	0.15	1.00	1.00	0.00	0.00	1.058	1.062*	1.325	1.095	1.227	1.120	1.091	1.069	1.086
100	30	0.90	0.15	1.00	1.00	0.00	0.00	1.038	1.044*	1.489	1.083	1.295	1.125	1.083	1.050	1.061
100	2	0.80	0.15	1.00	1.00	0.00	0.00	1.443	1.553	1.472	1.472	1.467	1.464	1.471	1.463*	1.553
100	5	0.80	0.15	1.00	1.00	0.00	0.00	1.237	1.246*	1.303	1.270	1.287	1.262	1.253	1.253	1.331
100	10	0.80	0.15	1.00	1.00	0.00	0.00	1.134	1.143*	1.259	1.167	1.224	1.173	1.156	1.146	1.184
100	20	0.80	0.15	1.00	1.00	0.00	0.00	1.068	1.087	1.342	1.119	1.248	1.138	1.105	1.079*	1.111
100	30	0.80	0.15	1.00	1.00	0.00	0.00	1.049	1.075	1.505	1.113	1.319	1.144	1.097	1.060*	1.094

Table 1 (continued)

T	m	$\bar{\lambda}$	σ_λ	σ_e	σ_μ	π	σ_ζ	InfeasLC	Equal	OLS	J-S	RR(0.1)	RR(0.5)	RR(1.0)	PC	Median
100	2	0.60	0.15	1.00	1.00	0.00	0.00	1.587	1.679	1.618	1.618	1.612	1.608*	1.614	1.616	1.679
100	5	0.60	0.15	1.00	1.00	0.00	0.00	1.349	1.363*	1.418	1.384	1.403	1.377	1.368	1.371	1.450
100	10	0.60	0.15	1.00	1.00	0.00	0.00	1.207	1.257	1.342	1.264	1.308	1.255	1.237	1.225*	1.296
100	20	0.60	0.15	1.00	1.00	0.00	0.00	1.118	1.208	1.405	1.215	1.319	1.205	1.167	1.132*	1.232
100	30	0.60	0.15	1.00	1.00	0.00	0.00	1.080	1.192	1.549	1.209	1.380	1.196	1.142	1.094*	1.207
100	2	1.00	0.15	1.00	1.00	0.02	0.00	1.436	1.743	1.466	1.466	1.460	1.467	1.494	1.457*	1.743
100	5	1.00	0.15	1.00	1.00	0.02	0.00	1.230	1.300	1.308	1.285	1.286	1.256	1.249	1.244*	1.308
100	10	1.00	0.15	1.00	1.00	0.02	0.00	1.133	1.153	1.277	1.179	1.230	1.172	1.155	1.146*	1.150
100	20	1.00	0.15	1.00	1.00	0.02	0.00	1.067	1.074*	1.369	1.110	1.253	1.134	1.102	1.077	1.078
100	30	1.00	0.15	1.00	1.00	0.02	0.00	1.043	1.047*	1.542	1.092	1.320	1.136	1.091	1.055	1.051
100	2	1.00	0.15	1.00	1.00	0.05	0.00	1.594	2.095	1.577	1.577	1.572	1.598	1.652	1.568*	2.095
100	5	1.00	0.15	1.00	1.00	0.05	0.00	1.330	1.442	1.407	1.388	1.382	1.348	1.342	1.342	1.326*
100	10	1.00	0.15	1.00	1.00	0.05	0.00	1.195	1.231	1.349	1.253	1.300	1.236	1.216	1.205	1.162*
100	20	1.00	0.15	1.00	1.00	0.05	0.00	1.100	1.110	1.427	1.150	1.306	1.177	1.139	1.110	1.082*
100	30	1.00	0.15	1.00	1.00	0.05	0.00	1.068	1.074	1.617	1.125	1.383	1.179	1.125	1.080	1.055*
100	2	1.00	0.00	1.00	1.00	0.00	0.05	1.349	1.640	1.398	1.398	1.398	1.414	1.439	1.386*	1.640
100	5	1.00	0.00	1.00	1.00	0.00	0.05	1.154	1.256	1.238	1.223	1.222	1.205	1.203	1.187*	1.363
100	10	1.00	0.00	1.00	1.00	0.00	0.05	1.080	1.124	1.216	1.137	1.176	1.128	1.115	1.101*	1.176
100	20	1.00	0.00	1.00	1.00	0.00	0.05	1.040	1.065	1.323	1.094	1.214	1.107	1.081	1.059*	1.091
100	30	1.00	0.00	1.00	1.00	0.00	0.05	1.025	1.042	1.501	1.087	1.286	1.114	1.074	1.041*	1.062
100	2	1.00	0.00	1.00	1.00	0.00	0.10	1.318	2.039	1.438	1.438	1.450	1.514	1.585	1.434*	2.039
100	5	1.00	0.00	1.00	1.00	0.00	0.10	1.115	1.406	1.241	1.241	1.228	1.227	1.241	1.191*	1.587
100	10	1.00	0.00	1.00	1.00	0.00	0.10	1.055	1.208	1.226	1.172	1.182	1.139	1.133	1.106*	1.292
100	20	1.00	0.00	1.00	1.00	0.00	0.10	1.025	1.102	1.352	1.122	1.227	1.114	1.088	1.060*	1.151
100	30	1.00	0.00	1.00	1.00	0.00	0.10	1.018	1.069	1.535	1.111	1.295	1.120	1.081	1.050*	1.105

Table 1 (continued)

T	m	$\bar{\lambda}$	σ_λ	σ_e	σ_μ	π	σ_ζ	InfeasLC	Equal	OLS	J-S	RR(0.1)	RR(0.5)	RR(1.0)	PC	Median
100	2	1.00	0.00	1.00	1.00	0.00	0.20	1.193	3.576	1.423	1.423	1.464	1.676	1.914	1.413*	3.576
100	5	1.00	0.00	1.00	1.00	0.00	0.20	1.055	2.042	1.257	1.260	1.246	1.287	1.344	1.195*	2.499
100	10	1.00	0.00	1.00	1.00	0.00	0.20	1.024	1.523	1.250	1.215	1.192	1.163	1.173	1.120*	1.717
100	20	1.00	0.00	1.00	1.00	0.00	0.20	1.012	1.261	1.414	1.196	1.236	1.124	1.107	1.088*	1.387
100	30	1.00	0.00	1.00	1.00	0.00	0.20	1.006	1.172	1.629	1.184	1.293	1.120	1.090	1.076*	1.262
200	2	1.00	0.00	1.00	1.00	0.00	0.00	1.330	1.493	1.343	1.343	1.341	1.345	1.359	1.338*	1.493
200	5	1.00	0.00	1.00	1.00	0.00	0.00	1.162	1.195	1.193	1.185	1.185	1.173	1.171	1.169*	1.280
200	10	1.00	0.00	1.00	1.00	0.00	0.00	1.092	1.100	1.152	1.111	1.134	1.110	1.103	1.099*	1.138
200	20	1.00	0.00	1.00	1.00	0.00	0.00	1.047	1.050*	1.164	1.063	1.124	1.076	1.062	1.053	1.074
200	30	1.00	0.00	1.00	1.00	0.00	0.00	1.032	1.033*	1.213	1.047	1.146	1.071	1.051	1.037	1.049
200	50	1.00	0.00	1.00	1.00	0.00	0.00	1.021	1.021*	1.362	1.040	1.216	1.083	1.050	1.026	1.031
200	2	1.00	0.15	1.00	1.00	0.00	0.00	1.335	1.507	1.349	1.349	1.347	1.352	1.367	1.343*	1.507
200	5	1.00	0.15	1.00	1.00	0.00	0.00	1.168	1.207	1.198	1.190	1.190	1.179	1.177	1.174*	1.297
200	10	1.00	0.15	1.00	1.00	0.00	0.00	1.089	1.102	1.146	1.109	1.129	1.106	1.099	1.095*	1.141
200	20	1.00	0.15	1.00	1.00	0.00	0.00	1.045	1.049*	1.162	1.061	1.122	1.074	1.060	1.051	1.074
200	30	1.00	0.15	1.00	1.00	0.00	0.00	1.031	1.034*	1.213	1.047	1.146	1.071	1.052	1.037	1.050
200	50	1.00	0.15	1.00	1.00	0.00	0.00	1.019	1.021*	1.363	1.039	1.216	1.082	1.050	1.024	1.031
200	2	0.80	0.15	1.00	1.00	0.00	0.00	1.433	1.542	1.448	1.448	1.446	1.448	1.458	1.443*	1.542
200	5	0.80	0.15	1.00	1.00	0.00	0.00	1.234	1.242*	1.264	1.250	1.257	1.246	1.242	1.242	1.333
200	10	0.80	0.15	1.00	1.00	0.00	0.00	1.134	1.145	1.195	1.154	1.179	1.155	1.147	1.140*	1.185
200	20	0.80	0.15	1.00	1.00	0.00	0.00	1.069	1.091	1.190	1.100	1.154	1.105	1.089	1.076*	1.115
200	30	0.80	0.15	1.00	1.00	0.00	0.00	1.047	1.074	1.232	1.083	1.173	1.096	1.072	1.052*	1.089
200	50	0.80	0.15	1.00	1.00	0.00	0.00	1.027	1.057	1.369	1.072	1.239	1.102	1.064	1.032*	1.066

Notes: The entries are the risk of the combination forecast in the column heading, relative to the risk of the infeasible forecast using the unobserved conditional mean. The design and estimators are discussed in the text. InfeasLC = infeasible optimal linear combination; Equal = simple average of $\{y_{t+1|t,i}\}$; OLS = OLS-weighted; J-S = James-Stein estimator; RR(k) = ridge regression with specified value of k; PC = principal component regression (factor model) estimator; median = median of $\{y_{t+1|t,i}\}$.

* Denotes the feasible combination forecast with the lowest risk for this set of parameter values (i.e. in that row).

($\bar{\lambda}$ = 0.6 and 0.8): in this case, the James-Stein and ridge regression forecasts shrink towards weights of $1/m$, but this weighting is not optimal for small $\bar{\lambda}$ and this evidently is a disadvantage for these estimators. This emphasizes the point that the risk of the James-Stein, ridge regression and equal-weighting estimators depend sensitively on the true parameter values because they favor a particular weighting vector.

A noteworthy feature of these results is that, for large m, the risks of the best combination forecasts approach the risk of the infeasible optimal linear combination forecast based on the theoretical weight vector β_0.

When the factor loadings vary over time ($\sigma_\zeta > 0$), the performance of all the forecasts deteriorates, relative to the unobserved conditional mean. In these trials, the PC forecast tends to have the lowest MSFE. This is somewhat surprising, because the factor structure that the PC forecast exploits is no longer present. However, it was shown theoretically in Stock and Watson (1998b) that principal components continues to provide consistent (as $m \to \infty$ and $T \to \infty$) estimates of the factor even with time-varying factor loadings for small or moderate time variation, and these Monte Carlo results provide some confirmation of this theoretical prediction.

The results for T = 200 are consistent with the results for T = 100. For m = 2, the estimation methods have similar risks. As m increases to 10, the PC forecast has relatively better performance. For large m, however, the equal-weighted forecast slightly outperforms the PC forecast, except in the case that $\bar{\lambda}$ = 0.8. For all m, the relative risk of the PC forecast is close to the bound given by the infeasible optimal linear combination forecast.

In summary, four conclusions emerge from these Monte Carlo results. First, the relative risks of the PC, James-Stein, RR, equal-weighting and median forecast all decrease as the number of forecasts increase, whereas the relative risk of the OLS forecast has a U-shape in m. The reason for the decreasing risk in m for the equal-weighted average, the median, and the PC forecast is that the variability in the idiosyncratic term is averaged out by the estimator so that the averaging produces an estimate of μ_t that improves as m increases. It appears that a similar phenomenon is occuring for the James-Stein and ridge regressions, because they shrink towards equal weighting. For OLS the increase in precision associated with the use of more series is eventually outweighed by the increase in sampling error associated with estimating the combining weights.

Second, the equal-weighting estimator generally works well unless m is small or the optimal weights are far from $1/m$.

Third, the PC forecast is most often the best forecast and is always close to the best forecast. Although the PC estimator does not involve shrinkage, it exploits the strong restriction of the single factor model. Presumably this is the source of its good forecasting performance.

Fourth, the median works well for $m \geq 10$, and is the best of those considered here when there are some large outliers in the forecasts. This provides support for the common practice of reporting the median forecast produced by a panel of forecasters.

4 Application: Forecasts of U.S. macroeconomic time series

4.1 Forecast data set

We now turn to an empirical investigation of combination forecasts using a dataset of simulated real-time univariate forecasts of 215 monthly macroeconomic time series for the United States, constructed and initially analyzed in Stock and Watson (1998a). The details of the construction of this panel are rather involved, and only a summary is provided here. The details are given in Stock and Watson (1998a).

Data. The series were chosen to be representative of the main groups of U.S. macroeconomic and financial time series data of interest to macroeconomists. For most series, the data span 1959:1–1996:12, although some series have later start dates and thus shorter spans. The series are listed in the Appendix. The only preliminary transformation used was taking logarithms of some of the data. In general, series in rates (interest rates, unemployment rates) and series with negative values were left in native units, and logarithms were taken otherwise.

Primitive models and forecasting procedures. For each time series, a total of 105 primitive models were estimated. These primitive models included simple models (exponential smoothing and no change), linear models (autoregressions), and nonlinear models (neural networks and logistic smooth transition autoregressive [LSTAR] models).

From these 105 primitive models, forecasts were produced by 49 forecasting methods. These methods are listed in Table 2. For some methods, the forecasts were those produced by one of the 105 primitive models. For example, the forecasting method in Table 2 produces a forecast using an autoregression specified in levels (or logarithms) of the data with a constant term and four lags. For other methods, these forecasts consisted of model selection using an information criterion, either the AIC or the BIC. Other forecasting methods incorporated unit root pretests, with the unit root test based on the Elliott-Rothenberg-Stock (1996) DF-GLS procedure.

Estimation methods and out-of-sample forecasts. The forecasts were computed as simulated ex-ante forecasts, that is, they were computed in a way to simulate real-time forecasting. Thus, for each series at each date for each horizon, each model was estimated using data up to that date, say date t. The forecasting procedure then made subsequent unit root and model selection decisions using the forecasts and data through date t. This entire process was repeated at date $t + 1$. For series with no seasonal adjustment and no revisions, the forecasts thus constructed are out-of-sample. For series with seasonal adjustment or other revisions, these would be out-of-sample were the final revisions known at date t; since the revisions are unknown, these deviate from true out of sample forecasts in this sense.

Forecasts were made at three horizons: 1, 6 and 12 months. The multistep forecasts were computed by linear or nonlinear least squares regression of the

Table 2. Summary of 49 forecasting methods used to construct the forecast database

Code	Description
A. Linear methods	
AR(p,u,d)	Autoregressive methods
	p = number of lags = 4, A (AIC, $0 \leq p \leq 12$), or B (BIC, $0 \leq p \leq 12$)
	u = method of handling possible unit root
	= L (levels), D (differences), or P (unit root pretest:
	DF-GLS$^\mu$ if d=C, DF-GLS$^\tau$ if d=T)
	d = deterministic components included
	= C (constant only) or T (constant and linear time trend)
EX1	Single exponential smoothing
EX2	Double exponential smoothing
EXP	DF-GLS$^\mu$ pretest between EX1 and EX2
B. Nonlinear methods	
NN(p,u,n_1,n_2)	Artificial neural net methods
	p = number of lags = 3, A (AIC, p=1,3), or B (BIC, p=1,3)
	(same number number of lags in each hidden unit)
	u = L (levels), D (differences), or P (DF-GLS$^\mu$ unit root pretest)
	n_1 = number of hidden units in first hidden layer
	= 2, A (AIC, $1 \leq n_1 \leq 3$), or B (BIC, $1 \leq n_1 \leq 3$)
	n_2 = number of hidden units in second hidden layer
	= 0 (only one hidden layer), 1 or 2
LS(p,u,ξ)	LSTAR methods
	p = number of lags = 3, A (AIC, p=1,3,6), or B (BIC, p=1,3,6)
	u = L (levels), D (differences), or P (DF-GLS$^\mu$ unit root pretest)
	ξ = switching variable
	= L ($\xi_t = y_t$), D ($\xi_t = \Delta y_t$), M (either L or D depending on unit root pretest), D6
	($\xi_t = y_t - y_{t-6}$), A (AIC over $\xi_t = \{y_t, y_{t-2}, y_{t-5}, y_t - y_{t-6}$, and $y_t - y_{t-12}\}$
	if levels specification, or $\xi_t = \{\Delta y_t, \Delta y_{t-2}, \Delta y_{t-5}, y_t - y_{t-6}$, and $y_t - y_{t-12}\}$
	if differences specification), or B (BIC, same set as AIC).
C. No change	
NOCHANGE	$y_{t+h\|t} = y_t$

Source: Stock and Watson (1998a).

h-step ahead value of the series on current and past vaues of the series. For the linear models, estimation was by OLS. For the nonlinear models, estimation was by nonlinear least squares using a combined random search and quadratic hill-climbing method.

The dataset for the analysis of combination forecasts thus consists of a panel of forecasts for each of these 215 series at horizons h = 1, 6, 12. Each panel has 49 forecasts (the 49 procedures in Table 2). For each series, the panel is balanced, but the length of the panel differs from series to series depending on data availability. For most series, the panel of forecasts commences in 1970:1, although for series that start after 1959:1, the panel of forecasts commences 120 periods after the first period available for use in estimating the primitive models. In all, this dataset consists of approximately 9 850 000 forecasts.

Two variants of this data set were used. The first consists of the unadjusted forecasts as just described. A feature of these forecasts is that there are rare, extremely large forecast errors, where the forecast errors can be five to ten standard deviations larger than a typical change in the series. The second variant therefore incorporates an additional adjustment to the forecasts by truncating the largest outlier forecasts. The truncation rule was that if a forecasted change exceeded the largest change in the series (at the relevant horizon) in the full historical data set, then the forecasted change was truncated to equal the largest historical change. This mechanical procedure can be viewed as incorporating, in real time, the effect of human intervention when one of the 49 forecasting methods goes badly awry.

Table 3. Summary of combination forecast methods used in the empirical analysis

Mnemonic	Description
Equal and MSE-weighted averages	
C(0,REC,Group)	Simple average over Group
C(1,TW,Group)	Inverse MSFE weighted average over Group
	TW = number of observations in rolling window to compute MSFEs
	= REC (recursive – all past forecasts used), 120, 60
Ridge regression	
R(k,Group)	Ridge regression, with ridge regression parameter k, by Group.
Principal components (factor model)	
F(r,Group)	PC regression forecast, r factors, by Group
Median and trimmed means	
Med(Group)	Median combination forecasts by Group
TM(α,Group)	Trimmed mean, trimming parameter α (in percent) by Group
Predictive least squares	
PLS(TW,Group)	Predictive least squares combination forecasts
	TW = REC, 120, 60

Notes: Group = A (linear), B (nonlinear), or A-C (all).

4.2 Combination forecasting methods

Five sets of combination forecasts were considered: weighted averages; ridge regression; principal components regression; median and trimmed mean; and predictive least squares (PLS). These combination forecasts are summarized in Table 3. The OLS and James-Stein methods were infeasible for this data set because for some series in some subperiods, one or more of the 49 methods produced identical forecasts (e.g. the AIC and BIC model selection methods selected the same primitive model and thus produced the same forecast). Thus the second moment matrix of the forecasts was singular for some series in some subperiods and these methods, which involve inverses of this matrix, could not be computed without further modifications. A ridge regresssion estimator with very small k was however included to approximate the performance of OLS weighting.[3]

Equal weighting and inverse MSE-*weighted averages.* These methods compute weighted averages of $\{y_{t+h|t,i}\}$, $i = 1, \ldots, 49$, respectively by their simple average or by weighting by inverse MSE as suggested by Bates and Granger (1969). For the inverse-MSE weighted average, the weight on the i-th forecast is $\beta_i = \mathrm{MSE}_{i,t}^{-1} / \sum_{j=1}^{m}(\mathrm{MSE}_{j,t}^{-1})$, where $\mathrm{MSE}_{i,t}$ is the mean squared forecast error over the historical period over which forecasts have been computed, $(t - h - T_1 + 1)^{-1} \sum_{s=T_1}^{t-h}(y_{s+h} - y_{s+h|s,i})^2$, where T_1 is the first date at which simulated out-of-sample forecasts are computed for the 49 methods.

Ridge regression. The ridge regression combination forecasts are computed as described in Sect. 2.3, with k = 0.001, 0.1, 0.5, 1. The ridge regression forecast for k = 0.001 is a numerical approximation to the OLS forecast.

Principal components (factor model). Two principal components forecasts were computed. The first was computed as described in Sect. 2.2. The factor $\hat{\mu}_t(h)$ was estimated by computing the ordered eigenvectors of the $m \times m$ moment matrix of the h-step ahead forecasts. The second stage regression (3.7) was modified so that an intercept was included in the regression of y_{t+h} on $\hat{\mu}_t$, and this intercept was added to the linear combination forecast $\hat{\beta}^{\mathrm{PC}\prime} Y_{t+h|t}$. The second principal component forecast was computed using two estimated factors. If the exact single factor model applies to these data, this second factor is redundant and should not help forecasting ability. The second factor was included in the event that the single factor model provides a poor approximation to these forecasts.

Median and trimmed means. Because some of the forecast errors are very large for some methods in some series at some dates, two robust averages were also computed, the median and the α-trimmed mean of $\{y_{t+h|t,i}\}$. The α-trimmed mean is computed by dropping the smallest 100α% and largest 100α% of the values of $\{y_{t+h|t,i}\}$ and computing a simple average of the remaining 100(1–2α)% forecasts.

PLS. The predictive least squares model selection procedure was included as another benchmark against which to judge the forecast combination methods. The PLS criterion selects the single method which has produced the best forecasts (lowest MSFE) to date, among the group at hand. The PLS is applied to histories that include the entire prediction period, the previous 120 months only, and the previous 60 months only, to yield three PLS forecasts.

Because some of these methods require initial observations on the 49 methods, the combination forecasts were computed starting in 1972:1 (or, for series with start dates after 1959:1, 24 months after the first out-of-sample primitive forecast was produced).

[3] The results for the equal-weighted and inverse MSE-weighted averages, the median, and the PLS forecasts were originally reported for this dataset in Stock and Watson (1998a).

4.3 Results

The performance of the various forecasts is assessed by their mean squared forecast error loss, averaged over all simulated out-of-sample time periods and over all 215 series. This average loss of the h-step ahead forecast produced by forecasting method j is,

$$\text{Loss}_{j,h} = \frac{1}{215} \sum_{\text{series}\{y\}} \frac{1}{T_3 - T_2 + 1} \sum_{t=T_2}^{T_3} \left[(y_{t+h} - \hat{y}_{t+h|t,j}) / \hat{\sigma}_h \right]^2 \tag{4.1}$$

where T_2 is the first out of sample period, T_3 is the final forecast period, and $\hat{\sigma}_h$ is the sample standard deviation of $y_{t+h} - y_t$. This average loss is the empirical counterpart to the theoretical risk (2.2) and its Monte Carlo estimates tabulated in the previous section. Because the units of the series generally differ, the risk for each series is normalized by the variance of its h-period change.

To facilitate comparisons of this loss, it is useful to report the loss relative to a benchmark forecast. The benchmark forecast used here is the forecast from an AR(4) in levels with a constant term. The performance of the combination forecasts is assessed for three different sets of forecasting methods: all 49 methods, the 21 linear methods (group A in Table 2), and the 27 nonlinear methods (group B in Table 2).

Table 4. Performance of combination forecasts, U.S. macroeconomic forecast dataset (Groups A-C (all 49 forecasts per series per month), with outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	C(1,60,A-C)	0.9452	C(0,REC,A-C)	0.8741	C(0,REC,A-C)	0.7749
2	C(1,120,A-C)	0.9456	C(1,120,A-C)	0.8763	TM(5,A-C)	0.7819
3	C(0,REC,A-C)	0.9458	TM(5,A-C)	0.8775	C(1,120,A-C)	0.7844
4	C(1,REC,A-C)	0.9458	C(1,REC,A-C)	0.8777	TM(10,A-C)	0.7861
5	TM(5,A-C)	0.9476	C(1,60,A-C)	0.8778	C(1,REC,A-C)	0.7873
6	TM(10,A-C)	0.9496	TM(10,A-C)	0.8796	C(1,60,A-C)	0.7914
7	TM(20,A-C)	0.9533	TM(20,A-C)	0.8847	TM(20,A-C)	0.7944
8	TM(30,A-C)	0.9558	TM(30,A-C)	0.8883	TM(30,A-C)	0.7993
9	Med(A-C)	0.9589	Med(A-C)	0.8943	Med(A-C)	0.8070
10	R(1,A-C)	0.9659	R(1,A-C)	0.9339	R(1,A-C)	0.8677
11	F(1,A-C)	0.9748	R(0.5,A-C)	0.9653	R(0.5,A-C)	0.9050
12	F(2,A-C)	0.9770	F(1,A-C)	0.9781	PLS(REC,A-C)	0.9874
13	R(0.5,A-C)	0.9822	F(2,A-C)	1.0068	PLS(120,A-C)	1.0034
14	PLS(60,A-C)	1.0048	PLS(REC,A-C)	1.0164	F(1,A-C)	1.0138
15	PLS(120,A-C)	1.0060	PLS(120,A-C)	1.0258	R(0.1,A-C)	1.0376
16	PLS(REC,A-C)	1.0095	PLS(60,A-C)	1.0630	PLS(60,A-C)	1.0445
17	R(0.1,A-C)	1.0556	R(0.1,A-C)	1.0938	F(2,A-C)	1.0475
18	R(0.001,A-C)	1.8050	R(0.001,A-C)	2.0246	R(0.001,A-C)	1.8645

Notes: "Rel. Loss" is the mean squared error loss relative to AR(4) with a constant in levels, horizon h = 1, 6, and 12.

The performance of the combination forecasts are summarized in Tables 4 (all 49 methods), 5 (group A), and 6 (group B) for the adjusted forecasts, in which the largest forecasts are truncated to simulate forecaster involvement. For comparison

Table 5. Performance of individual and combination forecasts, U.S. macroeconomic forecast dataset (Group A (21 linear forecasts per series per month), with outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	C(0,REC,A)	0.9776	C(0,REC,A)	0.9052	C(0,REC,A)	0.8131
2	C(1,60,A)	0.9788	C(1,120,A)	0.9129	C(1,120,A)	0.8278
3	C(1,120,A)	0.9789	C(1,60,A)	0.9137	TM(10,A)	0.8288
4	C(1,REC,A)	0.9792	C(1,REC,A)	0.9139	TM(5,A)	0.8293
5	TM(5,A)	0.9808	TM(5,A)	0.9141	C(1,REC,A)	0.8299
6	TM(10,A)	0.9816	TM(10,A)	0.9142	TM(20,A)	0.8301
7	R(1,A)	0.9820	TM(20,A)	0.9163	TM(30,A)	0.8324
8	TM(20,A)	0.9832	TM(30,A)	0.9182	C(1,60,A)	0.8329
9	TM(30,A)	0.9841	Med(A)	0.9242	Med(A)	0.8424
10	R(0.5,A)	0.9853	R(1,A)	0.9321	R(1,A)	0.8592
11	Med(A)	0.9880	AR(B,P,C)	0.9452	AR(A,P,C)	0.8615
12	AR(4,P,C)	0.9961	R(0.5,A)	0.9466	AR(A,D,C)	0.8636
13	AR(4,P,T)	0.9977	AR(B,P,T)	0.9498	AR(A,P,T)	0.8667
14	AR(4,D,C)	0.9982	AR(B,D,C)	0.9502	AR(B,P,C)	0.8679
15	AR(4,L,C)	1.0000	AR(A,P,C)	0.9531	AR(B,D,C)	0.8725
16	PLS(120,A)	1.0019	AR(A,D,C)	0.9551	AR(B,P,T)	0.8748
17	PLS(REC,A)	1.0020	AR(A,P,T)	0.9565	R(0.5,A)	0.8797
18	AR(B,P,C)	1.0021	AR(4,P,C)	0.9641	AR(4,P,C)	0.8909
19	R(0.1,A)	1.0027	AR(4,D,C)	0.9694	AR(4,D,C)	0.8960
20	AR(B,P,T)	1.0041	AR(4,P,T)	0.9701	AR(4,P,T)	0.8987
21	AR(B,D,C)	1.0048	AR(B,L,C)	0.9842	AR(A,L,C)	0.9352
22	AR(B,L,C)	1.0059	PLS(REC,A)	0.9904	R(0.1,A)	0.9487
23	PLS(60,A)	1.0075	AR(A,L,C)	0.9928	AR(B,L,C)	0.9502
24	F(1,A)	1.0098	PLS(120,A)	0.9998	PLS(120,A)	0.9610
25	AR(4,D,T)	1.0116	AR(4,L,C)	1.0000	PLS(REC,A)	0.9642
26	AR(A,P,C)	1.0148	F(1,A)	1.0047	AR(4,L,C)	1.0000
27	AR(A,D,C)	1.0163	R(0.1,A)	1.0101	PLS(60,A)	1.0051
28	AR(4,L,T)	1.0164	PLS(60,A)	1.0384	AR(A,D,T)	1.0363
29	AR(A,P,T)	1.0170	AR(B,D,T)	1.0402	F(1,A)	1.0478
30	AR(A,L,C)	1.0171	AR(4,D,T)	1.0452	EXP	1.0610
31	AR(B,D,T)	1.0191	AR(A,D,T)	1.0508	AR(B,D,T)	1.0630
32	AR(B,L,T)	1.0248	F(2,A)	1.0514	AR(4,D,T)	1.0673
33	AR(A,D,T)	1.0291	AR(4,L,T)	1.0939	EX2	1.0792
34	AR(A,L,T)	1.0351	AR(B,L,T)	1.1008	F(2,A)	1.1099
35	EXP	1.0505	EXP	1.1058	AR(B,L,T)	1.2035
36	F(2,A)	1.0527	AR(A,L,T)	1.1100	AR(A,L,T)	1.2094
37	EX2	1.0528	EX2	1.1177	AR(4,L,T)	1.2299
38	R(0.001,A)	1.1118	R(0.001,A)	1.4774	R(0.001,A)	1.3450
39	EX1	1.5218	EX1	1.9416	EX1	1.8426

Notes: See the notes to Table 4.

purposes, the 21 original (noncombined) linear forecasting methods are included along with the combination forecasts in Table 5, and the 27 nonlinear methods are included in Table 6. The performance of these forecasts for the unadjusted data set, which has some large outlier forecasts, particularly among the nonlinear models, are reported in Tables 7, 8 and 9. Because of the outliers, some relative losses are very large, so these final three tables only report results for procedures with relative losses less than 2.0.

Table 6. Performance of individual and combination forecasts, U.S. macroeconomic forecast dataset (Group B (27 nonlinear forecasts per series per month), with outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	C(1,60,B)	0.9376	C(1,120,B)	0.8757	C(1,120,B)	0.7809
2	C(1,120,B)	0.9378	C(1,REC,B)	0.8771	C(1,REC,B)	0.7838
3	C(1,REC,B)	0.9379	C(1,60,B)	0.8771	C(0,REC,B)	0.7845
4	C(0,REC,B)	0.9393	TM(5,B)	0.8793	TM(5,B)	0.7850
5	TM(5,B)	0.9394	TM(10,B)	0.8794	TM(10,B)	0.7859
6	TM(10,B)	0.9397	C(0,REC,B)	0.8800	C(1,60,B)	0.7879
7	TM(20,B)	0.9430	TM(20,B)	0.8845	TM(20,B)	0.7935
8	TM(30,B)	0.9458	TM(30,B)	0.8917	TM(30,B)	0.8056
9	Med(B)	0.9508	Med(B)	0.9011	Med(B)	0.8200
10	R(1,B)	0.9617	R(1,B)	0.9254	R(1,B)	0.8553
11	F(1,B)	0.9698	R(0.5,B)	0.9484	R(0.5,B)	0.8860
12	F(2,B)	0.9745	NN(3,P,2,0)	0.9733	NN(3,P,2,0)	0.9059
13	R(0.5,B)	0.9758	NN(3,P,2,1)	0.9762	LS(3,P,D6)	0.9071
14	NN(3,L,2,1)	0.9866	NN(3,P,2,2)	0.9769	NN(3,D,2,0)	0.9108
15	NN(3,P,2,0)	0.9883	F(1,B)	0.9778	NN(3,P,2,2)	0.9110
16	NN(3,D,2,0)	0.9915	NN(3,D,2,0)	0.9785	NN(3,P,2,1)	0.9116
17	NN(3,P,2,1)	0.9919	NN(3,D,2,1)	0.9806	L(3,D,D6)	0.9118
18	NN(3,L,2,2)	0.9921	NN(3,D,2,2)	0.9817	NN(A,P,A,0)	0.9140
19	NN(3,D,2,1)	0.9945	LS(3,P,D6)	0.9838	NN(3,D,2,1)	0.9159
20	PLS(120,B)	1.0077	NN(A,P,A,0)	0.9845	NN(3,D,2,2)	0.9163
21	PLS(60,B)	1.0102	LS(3,D,D6)	0.9877	NN(B,P,B,0)	0.9176
22	NN(3,P,2,2)	1.0113	NN(A,D,A,0)	0.9890	NN(A,D,A,0)	0.9198
23	PLS(REC,B)	1.0118	F(2,B)	0.9893	NN(B,D,B,0)	0.9214
24	NN(3,D,2,2)	1.0168	NN(B,P,B,0)	0.9919	LS(B,P,B)	0.9224
25	NN(A,P,A,0)	1.0170	NN(B,D,B,0)	0.9966	LS(B,D,B)	0.9238
26	NN(B,P,B,0)	1.0197	LS(B,P,B)	1.0108	LS(A,P,A)	0.9275
27	NN(A,L,A,0)	1.0200	LS(B,D,B)	1.0118	LS(A,D,A)	0.9298
28	NN(A,D,A,0)	1.0223	LS(3,P,P)	1.0197	LS(3,P,P)	0.9468
29	NN(B,D,B,0)	1.0229	LS(3,D,D)	1.0249	LS(3,D,D)	0.9521
30	LS(3,P,D6)	1.0251	PLS(REC,B)	1.0250	PLS(REC,B)	0.9661
31	NN(3,L,2,0)	1.0261	PLS(120,B)	1.0313	PLS(120,B)	0.9724
32	NN(B,L,B,0)	1.0267	LS(A,D,A)	1.0435	F(1,B)	0.9986
33	LS(3,D,D6)	1.0275	LS(A,P,A)	1.0436	R(0.1,B)	1.0061
34	LS(3,L,D6)	1.0378	R(0.1,B)	1.0455	NN(3,L,2,1)	1.0178
35	R(0.1,B)	1.0380	NN(3,L,2,1)	1.0605	F(2,B)	1.0255
36	LS(3,P,P)	1.0483	NN(3,L,2,2)	1.0664	PLS(60,B)	1.0267
37	LS(3,D,D)	1.0515	PLS(60,B)	1.0669	LS(3,L,D6)	1.0651
38	LS(B,P,B)	1.0571	LS(3,L,D6)	1.0693	NN(3,L,2,2)	1.1257
39	LS(3,L,L)	1.0616	LS(3,L,L)	1.2237	NN(3,L,2,0)	1.2911
40	LS(B,D,B)	1.0618	NN(3,L,2,0)	1.2408	LS(3,L,L)	1.2958
41	LS(B,L,B)	1.0978	NN(A,L,A,0)	1.2778	NN(B,L,B,0)	1.3737
42	LS(A,P,A)	1.0989	NN(B,L,B,0)	1.2828	LS(B,L,B)	1.3922
43	LS(A,D,A)	1.1040	LS(B,L,B)	1.3808	NN(A,L,A,0)	1.3962
44	LS(A,L,A)	1.1190	LS(A,L,A)	1.3921	LS(A,L,A)	1.4146
45	R(0.001,B)	1.5122	R(0.001,B)	1.6456	R(0.001,B)	1.6261

Notes: See the notes to Table 4.

Four conclusions are apparent. First, in all cases considered, combining forecasts improves on the performance of any individual method. For example, for the adjusted forecasts at horizon h = 12, among all 49 methods, the method with the lowest loss is AR(A,P,C), which has a relative mean squared error of 0.8615 (Table 5), while the best combination forecast is the simple average of the 49 forecasts, which has a relative loss of 0.7749, an improvement of 9%. In fact, at this horizon for the adjusted forecasts, trimmed means, medians, simple averages, inverse MSE-weighted averages, and ridge regression forecasts all outperform the best individual method. This is of course consistent with other results in the forecast combination literature.

Second, in the unadjusted datasets that include nonlinear forecasts, which contain some very large outliers, the only combination methods to perform well are the robust procedures, the median and the trimmed means. Overall, the trimmed means provide a good alternative to the highly robust median, which is inefficient unless the tails are very heavy, and the simple average. The simple average combination and ridge regression combinations deterioriate very sharply in the full unadjusted dataset. However, strikingly the trimmed mean forecasts deteriorate only slightly. For example, the 20% trimmed mean has a relative loss of 0.9533 at h = 1 in the adjusted data set (all 49 methods), and this becomes 0.9532 in the unadjusted dataset; at h = 6, these relative losses are 0.8847 and 0.8922; and at h = 12, they are 0.7944 and 0.8326. Even though the combination methods incorporate some extremely poor forecasts in the unadjusted dataset, the trimmed means substantially outperform the best individual method in every case.

Third, the ridge regression estimator that approximates OLS (with k = 0.001) performs very poorly in all cases. This is consistent with the theoretical and Monte Carlo results which suggest that OLS is a poor choice for combining many forecasts.

Fourth, the biggest discrepency between the empirical results and the Monte Carlo analysis arises in the performance of the ridge regression and, especially, the principal component forecasts. In the Monte Carlo analysis, the principal component forecasts were most often the best, but in the empirical analysis, they are among the worst of the combination methods in most cases, with relative losses in the unadjusted case hovering near 1.0 (although it should be noted that the principal component forecasts far outperform the ridge regression forecast that approximates OLS). Adding an additional factor to the principal component forecasts tends to increase their relative loss. Similarly, the ridge regression forecasts, which worked very well in the simulation exercise, typically perform worse than the simple average, median, and trimmed mean forecasts.

5 Discussion and conclusions

A common motivation for forecast combination is that the individual forecasts are based on different information sets and reflect expert judgments; thus combining the forecasts effectively broadens the information set used by any one forecaster,

Table 7. Performance of combination forecasts, U.S. macroeconomic forecast dataset (Groups A-C (all 49 forecasts per series per month), no outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	TM(20,A-C)	0.9532	TM(10,A-C)	0.8877	TM(10,A-C)	0.8261
2	TM(30,A-C)	0.9558	TM(20,A-C)	0.8922	TM(20,A-C)	0.8326
3	Med(A-C)	0.9587	TM(30,A-C)	0.8955	TM(30,A-C)	0.8369
4	TM(10,A-C)	0.9869	Med(A-C)	0.9000	Med(A-C)	0.8438
5	PLS(REC,A-C)	1.2505	TM(5,A-C)	1.3532	PLS(REC,A-C)	1.0363
6	TM(5,A-C)	1.5509	F(1,A-C)	1.4146	PLS(120,A-C)	1.0635
7	F(1,A-C)	1.7984	.	.	TM(5,A-C)	1.4621

Notes: Only methods with relative loss < 2 are reported. See the notes to Table 4.

and moreover combines the expert judgments of the individuals on the panel. Neither of these features are present in the dataset analyzed here: all the forecasts are univariate, and thus have the same information set (up to differences in lags), and the forecasts were all generated by computer so that there is no expert opinion involved. It is therefore all the more striking that in this dataset combination forecasts improved on individual forecasts, in some cases by a considerable margin. In this dataset, the deviations of forecasts from conditional expectations arise solely from specification error and estimation error.

Consistent with the theoretical and Monte Carlo results, the simple average forecasts and the robust forecasts (in particular the trimmed means) worked especially well in this dataset. A concrete recommendation which emerges from this empirical analysis is that the trimmed mean with 10% or 20% trimming produces robust yet efficient combination forecasts. Also, consistent with the theoretical results, the OLS combination forecast (as approximated by the ridge regression forecast with the least shrinkage) has very poor performance empirically.

Still, there are some discrepencies between the theoretical predictions and the actual performance of some of the forecasting methods. In particular, the ridge regression with the most shrinkage and principal component forecasts performed well in the Monte Carlo simulation, but not in the empirical analysis. This finding is similar to Guerard and Clemen's (1989) finding that simple averaging of forecasts outperformed latent root regression combination forecasts in a panel of real-time forecasts of U.S. GDP. One can speculate why this might be. For example, the quality of a particular forecast might change episodically, which would have the effect of inducing time-varying factor loadings. In this case the time-invariant single factor model would be inapplicable and the principal component forecasts might perform less well. However, the principal components forecast performed well in the Monte Carlo trials with time varying factor loadings, so time variation alone seems an unlikely explanation of the empirical findings. Alternatively, the idiosyncratic error covariance matrix Σ_e might have large correlations, so that the risk function is not well approximated by the trace of the variance-covariance matrix of the estimated weights. If so, the theoretical argument for James-Stein estimation, ridge regression, and principal component forecasts would be less compelling.

Table 8. Performance of individual and combination forecasts, U.S. macroeconomic forecast dataset (Group A (21 linear forecasts per series per month), no outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	C(0,REC,A)	0.9779	C(0,REC,A)	0.9134	C(0,REC,A)	0.8489
2	C(1,60,A)	0.9791	C(1,120,A)	0.9216	TM(10,A)	0.8645
3	C(1,120,A)	0.9792	C(1,REC,A)	0.9224	TM(20,A)	0.8656
4	C(1,REC,A)	0.9794	TM(10,A)	0.9224	C(1,120,A)	0.8658
5	TM(5,A)	0.9811	TM(5,A)	0.9225	TM(5,A)	0.8662
6	TM(10,A)	0.9819	C(1,60,A)	0.9226	TM(30,A)	0.8668
7	R(1,A)	0.9820	TM(20,A)	0.9241	C(1,REC,A)	0.8676
8	TM(20,A)	0.9836	TM(30,A)	0.9252	C(1,60,A)	0.8723
9	TM(30,A)	0.9845	Med(A)	0.9306	Med(A)	0.8765
10	R(0.5,A)	0.9853	R(1,A)	0.9368	R(1,A)	0.8820
11	Med(A)	0.9884	AR(B,P,C)	0.9506	AR(A,P,C)	0.8958
12	AR(4,P,C)	0.9961	R(0.5,A)	0.9515	AR(A,D,C)	0.8980
13	AR(4,P,T)	0.9977	AR(B,P,T)	0.9547	AR(A,P,T)	0.9012
14	AR(4,D,C)	0.9982	AR(B,D,C)	0.9556	AR(B,P,C)	0.9029
15	AR(4,L,C)	1.0000	AR(A,P,C)	0.9588	R(0.5,A)	0.9031
16	PLS(120,A)	1.0020	AR(A,D,C)	0.9608	AR(B,D,C)	0.9078
17	PLS(REC,A)	1.0020	AR(A,P,T)	0.9624	AR(B,P,T)	0.9102
18	R(0.1,A)	1.0027	AR(4,P,C)	0.9702	AR(4,P,C)	0.9258
19	AR(B,P,C)	1.0027	AR(4,D,C)	0.9755	AR(4,D,C)	0.9311
20	AR(B,P,T)	1.0047	AR(4,P,T)	0.9758	AR(4,P,T)	0.9341
21	AR(B,D,C)	1.0054	AR(B,L,C)	0.9904	AR(A,L,C)	0.9487
22	AR(B,L,C)	1.0065	PLS(REC,A)	0.9917	AR(B,L,C)	0.9669
23	PLS(60,A)	1.0077	AR(A,L,C)	0.9966	PLS(REC,A)	0.9912
24	F(1,A)	1.0098	AR(4,L,C)	1.0000	AR(4,L,C)	1.0000
25	AR(4,D,T)	1.0116	PLS(120,A)	1.0034	PLS(120,A)	1.0024
26	AR(A,P,C)	1.0151	F(1,A)	1.0081	PLS(60,A)	1.0535
27	AR(4,L,T)	1.0164	R(0.1,A)	1.0181	AR(A,D,T)	1.0775
28	AR(A,D,C)	1.0166	PLS(60,A)	1.0468	AR(B,D,T)	1.0851
29	AR(A,P,T)	1.0173	AR(B,D,T)	1.0501	F(1,A)	1.0893
30	AR(A,L,C)	1.0177	AR(4,D,T)	1.0545	AR(4,D,T)	1.0945
31	AR(B,D,T)	1.0197	AR(A,D,T)	1.0600	EXP	1.1711
32	AR(B,L,T)	1.0253	F(2,A)	1.0984	EX2	1.1903
33	AR(A,D,T)	1.0295	AR(4,L,T)	1.1036	F(2,A)	1.2490
34	AR(A,L,T)	1.0358	EXP	1.1152	AR(B,L,T)	1.2548
35	EXP	1.0506	AR(B,L,T)	1.1153	AR(A,L,T)	1.2617
36	F(2,A)	1.0527	AR(A,L,T)	1.1245	AR(4,L,T)	1.2868
37	EX2	1.0528	EX2	1.1272	EX1	1.9407
38	R(0.001,A)	1.1121	R(0.001,A)	1.5592	.	.
39	EX1	1.5218	EX1	1.9596	.	.

Notes: Only methods with relative loss < 2 are reported. See the notes to Table 4.

Although further work on these issues remains, these results provide a framework for understanding the often-good performance of simple average forecasts and median forecasts in empirical studies of combination forecasts. The instances of disparities between the theory and the empirical results suggest challenges for future research.

Table 9. Performance of individual and combination forecasts, U.S. macroeconomic forecast dataset (Group B (27 nonlinear forecasts per series per month), no outlier adjustment)

	h = 1		h = 6		h = 12	
Rank	Method	Rel.Loss	Method	Rel.Loss	Method	Rel.Loss
1	TM(30,B)	0.9459	TM(20,B)	0.8935	TM(20,B)	0.8369
2	Med(B)	0.9509	TM(30,B)	0.8995	TM(30,B)	0.8477
3	TM(20,B)	0.9966	Med(B)	0.9087	Med(B)	0.8620
4	NN(3,D,2,1)	1.0485	NN(3,P,2,1)	0.9867	NN(3,P,2,1)	0.9551
5	NN(3,P,2,1)	1.0734	NN(3,D,2,1)	0.9938	NN(3,D,2,1)	0.9596
6	NN(3,P,2,2)	1.2297	NN(3,L,2,1)	1.1884	LS(3,L,L)	1.9261
7	NN(3,D,2,2)	1.2322	F(1,B)	1.6681	.	.
8	.	.	LS(3,L,D6)	1.6882	.	.

Notes: Only methods with relative loss < 2 are reported. See the notes to Table 4.

Appendix: Data description

The time series used in this analysis are listed below. The data were obtained from the DRI BASIC Economics Database (creation date 9/97). The format for each series is its DRI BASIC mnemonic; a brief description; and the first date used (in brackets). A series that was preliminarily transformed by taking its logarithm is denoted by "log" in parentheses; otherwise, the series was used without preliminary transformation. Abbreviations: sa=seasonally adjusted; saar=seasonally adjusted at an annual rate ; nsa=not seasonally adjusted.

IP industrial production: total index (1992=100,sa) [1959:1] (log)

IPP industrial production: products, total (1992=100,sa) [1959:1] (log)

IPF industrial production: final products (1992=100,sa) [1959:1] (log)

IPC industrial production: consumer goods (1992=100,sa) [1959:1] (log)

IPCD industrial production: durable consumer goods (1992=100,sa) [1959:1] (log)

IPCN industrial production: nondurable condsumer goods (1992=100,sa) [1959:1] (log)

IPE industrial production: business equipment (1992=100,sa) [1959:1] (log)

IPI industrial production: intermediate products (1992=100,sa) [1959:1] (log)

IPM industrial production: materials (1992=100,sa) [1959:1] (log)

IPMD industrial production: durable goods materials (1992=100,sa) [1959:1] (log)

IPMND industrial production: nondurable goods materials (1992=100,sa) [1959:1] (log)

IPMFG industrial production: manufacturing (1992=100,sa) [1959:1] (log)

IPD industrial production: durable manufacturing (1992=100,sa) [1959:1] (log)

IPN industrial production: nondurable manufacturing (1992=100,sa) [1959:1] (log)

IPMIN industrial production: mining (1992=100,sa) [1959:1] (log)

IPUT industrial production: utilities (1992-=100,sa) [1959:1] (log)

IPX capacity util rate: total industry (% of capacity,sa)(frb) [1967:1]

IPXMCA capacity util rate: manufacturing,total(% of capacity,sa)(frb) [1959:1]

IPXDCA capacity util rate: durable mfg (% of capacity,sa)(frb) [1967:1]

IPXNCA capacity util rate: nondurable mfg (% of capacity,sa)(frb) [1967:1]

IPXMIN capacity util rate: mining (% of capacity,sa)(frb) [1967:1]

IPXUT capacity util rate: utilities (% of capacity,sa)(frb) [1967:1]

LHEL index of help-wanted advertising in newspapers (1967=100;sa) [1959:1]

LHELX employment: ratio; help-wanted ads:no. unemployed clf [1959:1]
LHEM civilian labor force: employed, total (thous.,sa) [1959:1] (log)
LHNAG civilian labor force: employed, nonagric.industries (thous.,sa) [1959:1] (log)
LHUR unemployment rate: all workers, 16 years & over (%,sa) [1959:1]
LHU680 unemploy.by duration: average(mean)duration in weeks (sa) [1959:1]
LHU5 unemploy.by duration: persons unempl.less than 5 wks (thous.,sa) [1959:1] (log)
LHU14 unemploy.by duration: persons unempl.5 to 14 wks (thous.,sa) [1959:1] (log)
LHU15 unemploy.by duration: persons unempl.15 wks + (thous.,sa) [1959:1] (log)
LHU26 unemploy.by duration: persons unempl.15 to 26 wks (thous.,sa) [1959:1] (log)
LHU27 unemploy.by duration: persons unempl.27 wks + (thous,sa) [1959:1] (log)
LHCH average hours of work per week (household data)(sa) [1959:1]
LPNAG employees on nonag. payrolls: total (thous.,sa) [1959:1] (log)
LP employees on nonag payrolls: total, private (thous,sa) [1959:1] (log)
LPGD employees on nonag. payrolls: goods-producing (thous.,sa) [1959:1] (log)
LPMI employees on nonag. payrolls: mining (thous.,sa) [1959:1] (log)
LPCC employees on nonag. payrolls: contract construction (thous.,sa) [1959:1] (log)
LPEM employees on nonag. payrolls: manufacturing (thous.,sa) [1959:1] (log)
LPED employees on nonag. payrolls: durable goods (thous.,sa) [1959:1] (log)
LPEN employees on nonag. payrolls: nondurable goods (thous.,sa) [1959:1] (log)
LPSP employees on nonag. payrolls: service-producing (thous.,sa) [1959:1] (log)
LPTU employees on nonag. payrolls: trans. & public utilities (thous.,sa) [1959:1] (log)
LPT employees on nonag. payrolls: wholesale & retail trade (thous.,sa) [1959:1] (log)
LPFR employees on nonag. payrolls: finance,insur.&real estate (thous.,sa [1959:1] (log)
LPS employees on nonag. payrolls: services (thous.,sa) [1959:1] (log)
LPGOV employees on nonag. payrolls: government (thous.,sa) [1959:1] (log)
LW avg. weekly hrs. of prod. wkrs.: total private (sa) [1964:1]
LPHRM avg. weekly hrs. of production wkrs.: manufacturing (sa) [1959:1]
LPMOSA avg. [1959:1]
LEH avg hr earnings of prod wkrs: total private nonagric ($,sa) [1964:1] (log)
LEHCC avg hr earnings of constr wkrs: construction ($,sa) [1959:1] (log)
LEHM avg hr earnings of prod wkrs: manufacturing ($,sa) [1959:1] (log)
LEHTU avg hr earnings of nonsupv wkrs: trans & public util($,sa) [1964:1] (log)
LEHTT avg hr earnings of prod wkrs:wholesale & retail trade(sa) [1964:1] (log)
LEHFR avg hr earnings of nonsupv wkrs: finance,insur,real est($,sa) [1964:1] (log)
LEHS avg hr earnings of nonsupv wkrs: services ($,sa) [1964:1] (log)
HSFR housing starts:nonfarm(1947-58);total farm&nonfarm(1959-)(thous.,sa [1959:1] (log)
HSNE housing starts:northeast (thous.u.)s.a. [1959:1] (log)
HSMW housing starts:midwest(thous.u.)s.a. [1959:1] (log)
HSSOU housing starts:south (thous.u.)s.a. [1959:1] (log)
HSWST housing starts:west (thous.u.)s.a. [1959:1] (log)
HSBR housing authorized: total new priv housing units (thous.,saar) [1959:1] (log)
HSBNE houses authorized by build. permits:northeast(thou.u.)s.a [1960:1] (log)
HSBMW houses authorized by build. permits:midwest(thou.u.)s.a. [1960:1] (log)
HSBSOU houses authorized by build. permits:south(thou.u.)s.a. [1960:1] (log)
HSBWST houses authorized by build. permits:west(thou.u.)s.a. [1960:1] (log)

HNS new 1-family houses sold during month (thous,saar) [1963:1] (log)
HNSNE one-family houses sold:northeast(thou.u.,s.a.) [1973:1] (log)
HNSMW one-family houses sold:midwest(thou.u.,s.a.) [1973:1] (log)
HNSSOU one-family houses sold:south(thou.u.,s.a.) [1973:1] (log)
HNSWST one-family houses sold:west(thou.u.,s.a.) [1973:1] (log)
HNR new 1-family houses, month's supply @ current sales rate(ratio) [1963:1]
HMOB mobile homes: manufacturers' shipments (thous.of units,saar) [1959:1] (log)
CONTC construct.put in place:total priv & public 1987$(mil$,saar) [1964:1] (log)
CONPC construct.put in place:total private 1987$(mil$,saar) [1964:1] (log)
CONQC construct.put in place:public construction 87$(mil$,saar) [1964:1] (log)
CONDO9 construct.contracts: comm'l & indus.bldgs(mil.sq.ft.floor sp.;sa) [1959:1] (log)
MSMTQ manufacturing & trade: total (mil of chained 1992 dollars)(sa) [1959:1] (log)
MSMQ manufacturing & trade:manufacturing;total(mil of chained 1992 dollars)(sa) [1959 (log):1]
MSDQ manufacturing & trade:mfg; durable goods (mil of chained 1992 dollars)(sa) [1959 (log):1]
MSNQ manufact. & trade:mfg;nondurable goods (mil of chained 1992 dollars)(sa) [1959:1 (log)]
WTQ merchant wholesalers: total (mil of chained 1992 dollars)(sa) [1959:1] (log)
WTDQ merchant wholesalers:durable goods total (mil of chained 1992 dollars)(sa) [1959 (log):1]
WTNQ merchant wholesalers:nondurable goods (mil of chained 1992 dollars)(sa) [1959:1] (log)
RTQ retail trade: total (mil of chained 1992 dollars)(sa) [1959:1] (log)
RTDQ retail trade:durable goods total (mil.87$)(s.a.) [1959:1] (log)
RTNQ retail trade:nondurable goods (mil of 1992 dollars)(sa) [1959:1] (log)
IVMTQ manufacturing & trade inventories:total (mil of chained 1992)(sa) [1959:1] (log)
IVMFGQ inventories, business, mfg (mil of chained 1992 dollars, sa) [1959:1] (log)
IVMFDQ inventories, business durables (mil of chained 1992 dollars, sa) [1959:1] (log)
IVMFNQ inventories, business, nondurables (mil of chained 1992 dollars, sa) [1959:1] (log)
IVWRQ manufacturing & trade inv:merchant wholesalers (mil of chained 1992 dollars)(s (log)[1959:1]
IVRRQ manufacturing & trade inv:retail trade (mil of chained 1992 dollars)(sa) [1959: (log)1]
IVSRQ ratio for mfg & trade: inventory/sales (chained 1992 dollars, sa) [1959:1]
IVSRMQ ratio for mfg & trade:mfg;inventory/sales (87$)(s.a.) [1959:1]
IVSRWQ ratio for mfg & trade:wholesaler;inventory/sales(87$)(s.a.) [1959:1]
IVSRRQ ratio for mfg & trade:retail trade;inventory/sales(87$)(s.a.) [1959:1]
PMI purchasing managers' index (sa) [1959:1]
PMP napm production index (percent) [1959:1]
PMNO napm new orders index (percent) [1959:1]
PMDEL napm vendor deliveries index (percent) [1959:1]
PMNV napm inventories index (percent) [1959:1]
PMEMP napm employment index (percent) [1959:1]
PMCP napm commodity prices index (percent) [1959:1]
MOCMQ new orders (net) - consumer goods & materials, 1992 dollars (bci) [1959:1] (log)
MDOQ new orders, durable goods industries, 1992 dollars (bci) [1959:1] (log)
MSONDQ new orders, nondefense capital goods, in 1992 dollars (bci) [1959:1] (log)
MO mfg new orders: all manufacturing industries, total (mil$,sa) [1959:1] (log)
MOWU mfg new orders: mfg industries with unfilled orders(mil$,sa) [1959:1] (log)
MDO mfg new orders: durable goods industries, total (mil$,sa) [1959:1] (log)
MDUWU mfg new orders:durable goods indust with unfilled orders(mil$,sa) [1959:1] (log)

MNO mfg new orders: nondurable goods industries, total (mil$,sa) [1959:1] (log)
MNOU mfg new orders: nondurable gds ind.with unfilled orders(mil$,sa) [1959:1] (log)
MU mfg unfilled orders: all manufacturing industries, total (mil$,sa) [1959:1] (log)
MDU mfg unfilled orders: durable goods industries, total (mil$,sa) [1959:1] (log)
MNU mfg unfilled orders: nondurable goods industries, total (mil$,sa) [1959:1] (log)
MPCON contracts & orders for plant & equipment (bil$,sa) [1959:1] (log)
MPCONQ contracts & orders for plant & equipment in 1992 dollars (bci) [1959:1] (log)
FM1 money stock: m1(curr,trav.cks,dem dep,other ck'able dep)(bil$,sa) [1959:1] (log)
FM2 money stock:m2(m1+o'nite rps,euro$,g/p&b/d mmmfs&sav&sm time dep(bil$, [1959:1] (log)
FM3 money stock: m3(m2+lg time dep,term rp's&inst only mmmfs)(bil$,sa) [1959:1] (log)
FML money stock:l(m3 + other liquid assets) (bil$,sa) [1959:1] (log)
FM2DQ money supply - m2 in 1992 dollars (bci) [1959:1] (log)
FMFBA monetary base, adj for reserve requirement changes(mil$,sa) [1959:1] (log)
FMBASE monetary base, adj for reserve req chgs(frb of st.louis)(bil$,sa) [1959:1] (log)
FMRRA depository inst reserves:total,adj for reserve req chgs(mil$,sa) [1959:1] (log)
FMRNBA depository inst reserves:nonborrowed,adj res req chgs(mil$,sa) [1959:1] (log)
FMRNBC depository inst reserves:nonborrow+ext cr,adj res req cgs(mil$,sa) [1959:1] (log)
FMFBA monetary base, adj for reserve requirement changes(mil$,sa) [1959:1] (log)
FCLS loans & sec @ all coml banks: total (bils,sa) [1973:1] (log)
FCSGV loans & sec @ all coml banks: U.S.govt securities (bil$,sa) [1973:1] (log)
FCLRE loans & sec @ all coml banks: real estate loans (bil$,sa) [1973:1] (log)
FCLIN loans & sec @ all coml banks: loans to individuals (bil$,sa) [1973:1] (log)
FCLNBF loans & sec @ all coml banks: loans to nonbank fin inst(bil$,sa) [1973:1] (log)
FCLNQ commercial & industrial loans oustanding in 1992 dollars (bci) [1959:1] (log)
FCLBMC wkly rp lg com'l banks:net change com'l & indus loans(bil$,saar) [1959:1]
CCI30M consumer instal.loans: delinquency rate,30 days & over, (%,sa) [1959:1]
CCINT net change in consumer instal cr: total (mil$,sa) [1975:1]
CCINV net change in consumer instal cr: automobile (mil$,sa) [1975:1]
FSNCOM nyse common stock price index: composite (12/31/65=50) [1959:1] (log)
FSNIN nyse common stock price index: industrial (12/31/65=50) [1966:1] (log)
FSNTR nyse common stock price index: transportation (12/31/65=50) [1966:1] (log)
FSNUT nyse common stock price index: utility (12/31/65=50) [1966:1] (log)
FSNFI nyse common stock price index: finance (12/31/65=50) [1966:1] (log)
FSPCOM s&p's common stock price index: composite (1941-43=10) [1959:1] (log)
FSPIN s&p's common stock price index: industrials (1941-43=10) [1959:1] (log)
FSPCAP s&p's common stock price index: capital goods (1941-43=10) [1959:1] (log)
FSPTR s&p's common stock price index: transportation (1970=10) [1970:1] (log)
FSPUT s&p's common stock price index: utilities (1941-43=10) [1959:1] (log)
FSPFI s&p's common stock price index: financial (1970=10) [1970:1] (log)
FSDXP s&p's composite common stock: dividend yield (% per annum) [1959:1] (log)
FSPXE s&p's composite common stock: price-earnings ratio (%,nsa) [1959:1] (log)
FSNVV3 nyse mkt composition:reptd share vol by size,5000+ shrs,% [1959:1] (log)
FYFF interest rate: federal funds (effective) (% per annum,nsa) [1959:1]
FYCP interest rate: commercial paper, 6-month (% per annum,nsa) [1959:1]
FYGM3 interest rate: U.S.treasury bills,sec mkt,3-mo.(% per ann,nsa) [1959:1]

FYGM6 interest rate: U.S.treasury bills,sec mkt,6-mo.(% per ann,nsa) [1959:1]
FYGT1 interest rate: U.S.treasury const maturities,1-yr.(% per ann,nsa) [1959:1]
FYGT5 interest rate: U.S.treasury const maturities,5-yr.(% per ann,nsa) [1959:1]
FYGT10 interest rate: U.S.treasury const maturities,10-yr.(% per ann,nsa) [1959:1]
FYAAAC bond yield: moody's aaa corporate (% per annum) [1959:1]
FYBAAC bond yield: moody's baa corporate (% per annum) [1959:1]
FWAFIT weighted avg foreign interest rate(%,sa) [1959:1]
FYFHA secondary market yields on fha mortgages (% per annum) [1959:1]
EXRUS united states;effective exchange rate(merm)(index no.) [1973:1] (log)
EXRGER foreign exchange rate: germany (deutsche mark per U.S.$) [1973:1] (log)
EXRSW foreign exchange rate: switzerland (swiss franc per U.S.$) [1973:1] (log)
EXRJAN foreign exchange rate: japan (yen per U.S.$) [1973:1] (log)
EXRUK foreign exchange rate: united kingdom (cents per pound) [1973:1] (log)
EXRCAN foreign exchange rate: canada (canadian $ per U.S.$) [1973:1] (log)
HHSNTN u. of mich. index of consumer expectations(bcd-83) [1959:1]
F6EDM U.S.mdse exports: [1964:1] (log)
FTMC6 U.S.mdse imports: crude materials & fuels (mil$,nsa) [1964:1] (log)
FTMM6 U.S.mdse imports: manufactured goods (mil$,nsa) [1964:1] (log)
PWFSA producer price index: finished goods (82=100,sa) [1959:1] (log)
PWFCSA producer price index:finished consumer goods (82=100,sa) [1959:1] (log)
PWIMSA producer price index:intermed mat.supplies & components(82=100,sa) [1959:1] (log)
PWCMSA producer price index:crude materials (82=100,sa) [1959:1] (log)
PWFXSA producer price index: finished goods,excl. foods (82=100,sa) [1967:1] (log)
PW160A producer price index: crude materials less energy (82=100,sa) [1974:1] (log)
PW150A producer price index: crude nonfood mat less energy (82=100,sa) [1974:1] (log)
PW561 producer price index: crude petroleum (82=100,nsa) [1959:1] (log)
PWCM producer price index: construction materials (82=100,nsa) [1959:1] (log)
PWXFA producer price index: all commodities ex.farm prod (82=100,nsa) [1959:1] (log)
PSM99Q index of sensitive materials prices (1990=100)(bci-99a) [1959:1] (log)
PUNEW cpi-u: all items (82-84=100,sa) [1959:1] (log)
PU81 cpi-u: food & beverages (82-84=100,sa) [1967:1] (log)
PUH cpi-u: housing (82-84=100,sa) [1967:1] (log)
PU83 cpi-u: apparel & upkeep (82-84=100,sa) [1959:1] (log)
PU84 cpi-u: transportation (82-84=100,sa) [1959:1] (log)
PU85 cpi-u: medical care (82-84=100,sa) [1959:1] (log)
PUC cpi-u: commodities (82-84=100,sa) [1959:1] (log)
PUCD cpi-u: durables (82-84=100,sa) [1959:1] (log)
PUS cpi-u: services (82-84=100,sa) [1959:1] (log)
PUXF cpi-u: all items less food (82-84=100,sa) [1959:1] (log)
PUXHS cpi-u: all items less shelter (82-84=100,sa) [1959:1] (log)
PUXM cpi-u: all items less midical care (82-84=100,sa) [1959:1] (log)
PSCCOM spot market price index:bls & crb: all commodities(67=100,nsa) [1959:1] (log)
PSCFOO spot market price index:bls & crb: foodstuffs (67=100,nsa) [1959:1] (log)
PSCMAT spot market price index:bls & crb: raw industrials(67=100,nsa) [1959:1] (log)
PZFR prices received by farmers: all farm products (1977=100,nsa) [1975:1] (log)

PCGOLD commodities price:gold,london noon fix,avg of daily rate,$ per oz [1975:1] (log)
GMDC pce,impl pr defl:pce (1987=100) [1959:1] (log)
GMDCD pce,impl pr defl:pce; durables (1987=100) [1959:1] (log)
GMDCN pce,impl pr defl:pce; nondurables (1987=100) [1959:1] (log)
GMDCS pce,impl pr defl:pce; services (1987=100) [1959:1] (log)
GMPYQ personal income (chained) (series #52) (bil 92$,saar) [1959:1] (log)
GMYXPQ personal income less transfer payments (chained) (#51) (bil 92$,saar) [1959:1] (log)
GMCQ personal consumption expend (chained) - total (bil 92$,saar) [1959:1] (log)
GMCDQ personal consumption expend (chained) - total durables (bil 92$,saar) [1959:1] (log)
GMCNQ personal consumption expend (chained) - nondurables (bil 92$,saar) [1959:1] (log)
GMCSQ personal consumption expend (chained) - services (bil 92$,saar) [1959:1] (log)
GMCANQ personal cons expend (chained) - new cars (bil 92$,saar) (log)

References

1. Amemiya, T., Morimune, K. (1974) Selecting the Optimal Order of Polynomial in the Almon Distributed Lag. *Review of Economics and Statistics* 56: 378–386
2. Bates, J.M., Granger, C.W.J. (1969) The Combination of Forecasts. *Operations Research Quarterly* 20: 451–468
3. Chamberlain, G., Rothschild, M. (1983) Arbitrage Factor Stucture, and Mean-Variance Analysis of Large Asset Markets. *Econometrica* 51: No. 5
4. Clemen, R.T. (1989) Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting* 5: 559–583
5. Clemen, R.T., Winkler, R.L. (1986) Combining Economic Forecasts. *Journal of Business and Economic Statistics* 4: 39–46
6. Connor, G., Korajczyk, R.A. (1986) Performance Measurement with the Arbitrage Pricing Theory. *Journal of Financial Economics* 15: 373–394
7. Connor, G., Korajczyk, R.A. (1988) Risk and Return in an Equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics* 21: 255–289
8. Connor, G., Korajczyk, R.A. (1993) A Test for the Number of Factors in an Approximate Factor Model. *Journal of Finance* 48: No. 4
9. Diebold, F.X., Pauly, P. (1990) The Use of Prior Information in Forecast Combination. *International Journal of Forecasting* 6: 503–508
10. Diebold, F.X., Lopez, J.A. (1995) Forecast Evaluation and Combination. In: Maddala, G.S., Rao, C.R. (eds.) *Handbook of Statistics*, vol. 14, pp. 241–268
11. Elliott, G., Rothenberg, T.J. Stock, J.H. (1996) "Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64: 813–836
12. Figlewski, S. (1983) Optimal Price Forecasting Using Survey Data. *Review of Economics and Statistics* 65: 13–21
13. Figlewski, S., Urich, T. (1983) Optimal Aggregation of Money Supply Forecasts: Accuracy, Profitability and Market Efficiency. *The Journal of Finance* 28: 695–710
14. Forni, M., Reichlin, L. (1995) Lets Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle. CEPR Discussion Paper no. 1244
15. Geweke, J. (1977) The Dynamic Factor Analysis of Economic Time Series. In: Aigner, D.J., Goldberger, A.S. (eds.) *Latent Variables in Socio-Economic Models*. North-Holland, Amsterdam, Chap. 19
16. Granger, C.W.J. (1989) Combining Forecasts – Twenty Years Later. *Journal of Forecasting* 8: 167–173
17. Granger, C.W.J., Ramanathan, R. (1984) Improved Methods of Combining Forecasting. *Journal of Forecasting* 3: 197–204
18. Guerard, J.B. Jr., Clemen, R.T. (1989) Collinearity and the Use of Latent Root Regression for Combining GNP Forecasts. *Journal of Forecasting* 8: 231–238

19. Hendry, D.F., Pagan, A.R. Sargan, J.D. (1984) Dynamic Specification. In: Griliches, Z., Intriligator, M.D. (eds.) *The Handbook of Econometrics*, vol. 2, pp. 1023–1100. North-Holland, Amsterdam
20. James, W., Stein, C. (1961) Estimation with Quadratic Loss. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, pp. 361–379
21. Judge, G.G., Griffiths, W.E. Hill, R.C., Lee, T.C. (1980). *The Theory and Practice of Econometrics*, 1st edn. John Wiley and Sons, New York
22. Lehmann, E.L. (1983). *Theory of Point Estimation*. John Wiley & Sons, New York
23. Sargent, T.J, Sims, C.A. (1977) Business Cycle Modeling without Pretending to have Too Much a-priori Economic Theory. In: Sims, C., et al. (eds.) *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis, Minneapolis
24. Stein, C. (1955) Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, pp. 197–206
25. Stock, J.H., Watson, M.W. (1998a) A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. NBER Working Paper
26. Stock, J.H., Watson, M.W. (1998b) Diffusion Indexes. NBER Working Paper