# Regression Approach to the Linear Combination of Multivariate Forecasts

Sven-Oliver Troschke
*Department of Statistics, University of Dortmund, 44221 Dortmund, Germany*
*troschke@statistik.uni-dortmund.de*

**Abstract:** In Troschke (2002) the author introduces a linear approach to the scalar mean square error optimal combination of forecasts for a vector random variable. In this paper it is shown how the optimal combination parameters can be obtained with the help of linear regression. Thus the application of these combination methods to empirical data is facilitated. An empirical example illustrating the performance of the new methods is given. These methods are compared to the classical univariate treatment of the respective variables.

**Keywords:** Combination of forecasts, multivariate forecasts, linear combination, linear regression.

**AMS 2000 Subject Classification:** 62M20, 62J05, 62P20

## 1 Introduction

Suppose that we are given $k$ forecasts $\mathbf{f}_1, \dots, \mathbf{f}_k$ for an $l$-dimensional random vector variable $\mathbf{y}$. The idea is nearby to combine the information contained in the individual forecasts $\mathbf{f}_i = (f_{i,1}, \dots, f_{i,l})^{\mathrm{T}}$ in order to obtain a single improved forecast for the target variable $\mathbf{y} = (y_1, \dots, y_l)^{\mathrm{T}}$.

The literature on the combination of forecasts almost exclusively deals with the combination of forecasts for a *univariate* random variable $y$, i.e. $l = 1$. Compare e.g. Clemen (1989) or Thiele (1993) for good overviews on the topic.

Troschke (2002) presents approaches for the linear combination of *multivariate* forecasts, i.e. $l > 1$. The single forecasts are stacked in the forecast vector $\mathbf{f} =$

$(\mathbf{f}_1^{\mathrm{T}}, \dots, \mathbf{f}_k^{\mathrm{T}})^{\mathrm{T}} \sim (kl \times 1)$. Then the linear combinations under study are of the form $\mathbf{f}_{\mathrm{comb}} = \mathbf{B}\mathbf{f} + \mathbf{c}$, where $\mathbf{B}$ is an $(l \times kl)$ matrix and $\mathbf{c}$ is an $(l \times 1)$ vector.

The approaches investigated in TROSCHKE (2002) varied with respect to the restrictions imposed on the combination parameters $\mathbf{B}$ and $\mathbf{c}$. We refer to Section 3 for details on the employed restrictions and the classes of multivariate linear combinations evolving from them. Assuming knowledge of the first and second order moments of the joint distribution of $\mathbf{y}$ and $\mathbf{f}$ the respective optimal choices for $\mathbf{B}$ and $\mathbf{c}$ obeying the respective restrictions were derived. These choices were made optimal in the sense of the scalar mean square prediction error

$$\mathrm{SMSPE}(\widetilde{\mathbf{f}}, \mathbf{y}) = \mathrm{E}[(\mathbf{y} - \widetilde{\mathbf{f}})^{\mathrm{T}}(\mathbf{y} - \widetilde{\mathbf{f}})] \tag{1.1}$$

of a forecast $\widetilde{\mathbf{f}}$ for a target variable $\mathbf{y}$. The optimal SMSPE-values for the respective combinations were derived. For an exemplary set of realistic first and second order moments the potential of the multivariate methods was compared to the univariate treatment of each variable involved.

For practical purposes one cannot expect the mentioned moments to be known and thus they have to be estimated from past data on the variable $\mathbf{y}$ and the forecasts $\mathbf{f}_i$. In this paper it is shown that instead of estimating the combination parameters on the basis of the corresponding sample moments one may equivalently perform appropriate least squares regressions.

This facilitates the application of multivariate combinations. Furthermore, the interpretation of multivariate combinations as linear regression problems constitutes an analogy to the univariate case. Here GRANGER and RAMANATHAN (1984) have introduced the regression approach as an alternative view on and enhancement of the classical methods based on sample moments.

Section 2 shortly reviews the regression methods in the case of univariate linear combinations, whereas the appropriate regression models in the multivariate case are identified in Section 3. In TROSCHKE (2002) also the case of $k = 1$ forecast is investigated, which results in multivariate linear adjustments of single forecasts. The linear regression approaches cover these adjustments as well. We will compare the performance of the various multivariate linear adjustments and combinations of forecasts in a small empirical example (see Section 4). Section 5 concludes the paper.

## 2 Univariate linear combinations

We write univariate linear forecast combinations in the form $f_{\mathbf{b},c} = \mathbf{b}^{\mathrm{T}}\mathbf{f} + c$ with $\mathbf{f} = (f_1, \dots, f_k)^{\mathrm{T}} \sim (k \times 1)$, $c \in \mathbb{R}$ and $\mathbf{b} = (b_1, \dots, b_k)^{\mathrm{T}} \in \mathbb{R}^k$, i.e. we consider the special case of univariate individual forecasts $f_i$, $\mathbf{B} = \mathbf{b}^{\mathrm{T}}$ and $\mathbf{c} = c$ of our general notation.

Depending on the situation it may be appropriate to impose restrictions on the combination parameters $\mathbf{b}$ and $c$. The restrictions that are dealt with in this paper are $\mathbf{b}^{\mathrm{T}}\mathbf{1}_k = 1$, i.e. the weights for the single forecasts sum up to one, and $c = 0$, i.e. no constant term is included. Each of the two restrictions may or may not be imposed resulting in a total of four different univariate linear combinations. Here $\mathbf{1}_k$ denotes the $k$-dimensional vector of ones.

If each of the single forecasts is unbiased the combined forecast will also be unbiased if the combination parameters satisfy $\mathbf{b}^{\mathrm{T}}\mathbf{1}_k = 1$ and $\mathbf{c} = \mathbf{0}$. Since unbiasedness is often assumed for the individual forecasts this combination is a standard approach in the literature. The theoretical considerations in Troschke (2002) showed that inclusion of a constant term guarantees unbiasedness of the combined forecast if the constant term is chosen adequately.

Granger and Ramanathan (1984) showed that these linear combinations are closely related to certain linear regression models as summarized in Table 1. Obviously, the unrestricted OLSCO (= ordinary least squares employing a constant term) combination provides the best fit in the sense of least squares regression, but as indicated above it might be reasonable to utilize restrictions on the combination parameters. For empirical data the following two lead to the same results: a) Performing an ordinary least squares regression of the target variable $y$ on the forecasts $f_i$; b) Inserting the simple sample moments for the true moments of the joint distribution of $y$ and $\mathbf{f}$ in the formulae for the theoretically optimal combination parameters derived in Troschke (2002).

The special case of $k = 1$ forecast results in adjustments of that forecast as described in Troschke (2002). The performance of an individual forecast can be improved by such an adjustment. Naturally, the adjustment parameters $b$ and $c$ are now real numbers. The corresponding regression models are also shown in Table 1. Granger (1989, p. 169) emphasizes the usefulness of the OLSCO adjustment and bias correction is a popular means as well.

The regressions for ERLSCO or ERLS (= equality restricted least squares) combinations may alternatively be carried out as regression of $y - f_1$ on $f_2 - f_1, \dots, f_k - f_1$ with or without a constant term. Thus the parameters $b_2, \dots, b_k$ are obtained while

| | Method | Notation | Regression |
|---|---|---|---|
| Single forecast | | $f_i,\ i \in \{1,\dots,k\}$ | |
| Adjustments | OLSCO-Adj<br>OLS-Adj<br>ERLSCO-Adj | $(f_i)_{b,c} = bf_i + c$<br>$(f_i)_b = bf_i$<br>$(f_i)_{1,c} = f_i + c$ | $y$ on $f_i$ with constant term<br>$y$ on $f_i$<br>bias correction |
| Linear combinations | AM | $f_{\mathrm{am}} = (1/k)\mathbf{1}_k^{\mathrm{T}}\mathbf{f}$ | arithmetic mean |
| | OLSCO | $f_{\mathbf{b},c} = \mathbf{b}^{\mathrm{T}}\mathbf{f} + c$ | $y$ on $f_1,\dots,f_k$ with constant term |
| | OLS | $f_{\mathbf{b}} = \mathbf{b}^{\mathrm{T}}\mathbf{f}$ | $y$ on $f_1,\dots,f_k$ |
| | ERLSCO | $f_{\mathbf{b},c,\mathrm{rest}} = \mathbf{b}^{\mathrm{T}}\mathbf{f} + c$ | $y$ on $f_1,\dots,f_k$ with constant term under the restriction $\mathbf{b}^{\mathrm{T}}\mathbf{1}_k = 1$ |
| | ERLS | $f_{\mathbf{b},\mathrm{rest}} = \mathbf{b}^{\mathrm{T}}\mathbf{f}$ | $y$ on $f_1,\dots,f_k$ under the restriction $\mathbf{b}^{\mathrm{T}}\mathbf{1}_k = 1$ |

Table 1: Univariate linear adjustment and combination methods

$b_1$ results from $b_1 = 1 - \sum_{i=2}^{k} b_i$. ERLSCO may be interpreted as the combination of the bias corrected forecasts subject to the then reasonable restriction of the weights $b_i$ summing up to unity, cf. TROSCHKE (2002). For the forecast $f_i$ the corresponding bias corrected forecast is given as $f_i - \mu_i + \mu_0$, where $\mu_i$ is the expectation of $f_i$ and $\mu_0$ is the expectation of $y$. In empirical applications these expectations will be replaced by the corresponding sample means.

Another important linear combination is the arithmetic mean of the individual forecasts:

$$f_{\mathrm{am}} = \frac{1}{k}\sum_{i=1}^{k} f_i = \frac{1}{k}\mathbf{1}_k^{\mathrm{T}}\mathbf{f}\ . \tag{2.1}$$

Here no regression is necessary, since the combination parameters are fixed as $b_i = 1/k$, $i = 1,\dots,k$ and $c = 0$. In spite of (or maybe because of) being so simple the arithmetic mean proves to be very powerful in empirical applications.

We now turn to the linear approaches to the combination and adjustment of *multivariate* forecasts. By considering interactions between the different components of $\mathbf{y}$ and the $\mathbf{f}_i$ additional information is exploited.

# 3 Multivariate linear combinations

Multivariate linear combinations are of the general form $\mathbf{Bf}+\mathbf{c}$, where $\mathbf{c}\in\mathbb{R}^l$ and $\mathbf{B}\in\mathbb{R}^{l\times kl}$. The versions analyzed here have been introduced in TROSCHKE (2002). They differ with respect to the restrictions imposed on $\mathbf{B}$ and $\mathbf{c}$.
The *strong* version

$$f_{\mathbf{B},\mathbf{c}}=\mathbf{Bf}+\mathbf{c}=\sum_{i=1}^{k}\mathbf{B}_i\mathbf{f}_i+\mathbf{c} \tag{3.1}$$

is based on a full $(l\times kl)$-matrix $\mathbf{B}=(\mathbf{B}_1|\dots|\mathbf{B}_k)$, where

$$\mathbf{B}_i=\begin{pmatrix} B_{i,11} & B_{i,12} & \dots & B_{i,1l} \\ B_{i,21} & B_{i,22} & \dots & B_{i,2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{i,l1} & B_{i,l2} & \dots & B_{i,ll} \end{pmatrix}\in\mathbb{R}^{l\times l} \tag{3.2}$$

for $i=1,\dots,k$. The *medium* version

$$f_{\mathbf{D},\mathbf{c}}=\mathbf{Df}+\mathbf{c}=\sum_{i=1}^{k}\mathbf{D}_i\mathbf{f}_i+\mathbf{c} \tag{3.3}$$

uses an $(l\times kl)$-matrix $\mathbf{D}=(\mathbf{D}_1|\dots|\mathbf{D}_k)$ with diagonal matrices

$$\mathbf{D}_i=\begin{pmatrix} D_{i,11} & 0 & \dots & 0 \\ 0 & D_{i,22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D_{i,ll} \end{pmatrix}\in\mathbb{R}^{l\times l} \tag{3.4}$$

for $i=1,\dots,k$. Finally, the *weak* version

$$f_{\boldsymbol{\alpha},\mathbf{c}}=\sum_{i=1}^{k}\alpha_i\mathbf{f}_i+\mathbf{c} \tag{3.5}$$

can be interpreted as restricting $\mathbf{B}$ to the $(l\times kl)$-matrix $(\alpha_1\mathbf{I}_l|\dots|\alpha_k\mathbf{I}_l)$ consisting of scalar multiples of the $l\times l$-identity matrix. The scalar coefficients are gathered in the vector $\boldsymbol{\alpha}=(\alpha_1,\dots,\alpha_k)^{\mathrm{T}}$.
The respective choices of the matrix $\mathbf{B}$ may be viewed as restrictions on $\mathbf{B}$ with the effect that the number of parameters involved is reduced from $kl^2+l$ over $kl+l$ to $k+l$. Since the number of observations, from which the unknown parameters are to

be estimated in empirical applications, is not so large in general, this reduction of the number of parameters may be reasonable.

Each version strong, medium and weak, may obey additional restrictions similarly to the univariate case (cf. Section 2): The sum of the matrices giving weight to the individual forecasts $\mathbf{f}_i$ may or may not be forced to sum to the identity matrix $\mathbf{I}_l$, and the constant term may or may not be set to the zero vector $\mathbf{0}$. If $\mathbf{c}$ is set to zero this will be indicated by cancelling the subscript $\mathbf{c}$ in the combined forecast; if the summation restriction is required it will be indicated by an additional subscript 'rest'. For example $\mathbf{f}_{\mathbf{D},\text{rest}}$ denotes the medium combined forecast with $\mathbf{c} = \mathbf{0}$ and with the diagonal weight matrices $\mathbf{D}_i$ summing to the identity matrix, i.e.

$$\mathbf{f}_{\mathbf{D},\text{rest}} = \sum_{i=1}^{k} \mathbf{D}_i \mathbf{f}_i \qquad \text{with} \qquad \sum_{i=1}^{k} \mathbf{D}_i = \mathbf{I}_l \; . \tag{3.6}$$

Obviously, the additional restrictions lower the number of unknown parameters even further.

Another reason for considering all three approaches is their relation to well-known methods in the combination of estimators: While strong multivariate linear combination under the restriction $\sum_{i=1}^{k} \mathbf{B}_i = \mathbf{I}_l$ is related to the (strong) covariance adjustment technique introduced by Rao (1966, 1967), weak multivariate linear combination under the restriction $\sum_{i=1}^{k} \alpha_i = 1$ is related to the weak covariance adjustment technique by Trenkler and Ihorst (1995). Medium multivariate linear combination is intermediate with respect to the former two and can be viewed as the univariate treatment of each of the variables involved thus representing the usual treatment in the literature. Confer also Troschke (2002).

In order to facilitate application of multivariate linear forecast combinations it is important to note, that we may regard the problem of finding the respective optimal combination parameters as linear regression problems just like it is the case with the univariate linear combination approaches (cf. Section 2).

Thiele (1993, Section 4.2.1 and 4.2.3) shows that using ordinary least squares estimation in the linear regression problems is equivalent to replacing the true moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by the respective sample moments in the formulae for the optimal combination parameters. For the problems considered here such formulae have been derived in Troschke (2002).

We will now identify the appropriate regression models for the various combinations as well as the corresponding data models.

## 3.1 Strong multivariate linear combinations

Subsequently we will deal with the strong multivariate combinations which have full weight matrices $\mathbf{B}_i$ as coefficients of the individual forecasts $\mathbf{f}_i$. We have to analyze four cases arising from using or neglecting a constant term or using or neglecting the restriction of the matrices $\mathbf{B}_i$ summing up to the identity matrix.

**Case 1a.** *Unrestricted combination with constant term* $\mathbf{f}_{\mathbf{B},\mathbf{c}} = \mathbf{Bf} + \mathbf{c}$:
Setting $\mathbf{f} = (f_{1,1}, \ldots, f_{1,l}, f_{2,1}, \ldots, f_{2,l}, \ldots, f_{k,1}, \ldots, f_{k,l})^{\mathrm{T}} =: (g_1, \ldots, g_{kl})^{\mathrm{T}} = \mathbf{g}$ and

$$\begin{aligned}
\mathbf{B} &= \left(\begin{array}{ccc|c|ccc} B_{1,11} & \ldots & B_{1,1l} & & B_{k,11} & \ldots & B_{k,1l} \\ \vdots & & \vdots & \ldots & \vdots & & \vdots \\ B_{1,l1} & \ldots & B_{1,ll} & & B_{k,l1} & \ldots & B_{k,ll} \end{array}\right) \\
&=: \begin{pmatrix} B_{11} & \ldots & B_{1(kl)} \\ \vdots & & \vdots \\ B_{l1} & \ldots & B_{l(kl)} \end{pmatrix} \in \mathbb{R}^{l \times kl}
\end{aligned} \tag{3.7}$$

we rewrite the SMSPE-function belonging to the strong linear combination without summation restriction and with constant term as:

$$\begin{aligned}
&\mathrm{SMSPE}(\mathbf{f}_{\mathbf{B},\mathbf{c}}, \mathbf{y}) = \mathrm{E}[(\mathbf{y} - \mathbf{f}_{\mathbf{B},\mathbf{c}})^{\mathrm{T}}(\mathbf{y} - \mathbf{f}_{\mathbf{B},\mathbf{c}})] \\
&\quad = \mathrm{E}[(\mathbf{y} - \mathbf{Bf} - \mathbf{c})^{\mathrm{T}}(\mathbf{y} - \mathbf{Bf} - \mathbf{c})] \\
&\quad = \mathrm{E}\left[\sum_{j=1}^{l} (y_j - (\mathbf{Bf})_j - c_j)^2\right] \\
&\quad = \mathrm{E}\left[\sum_{j=1}^{l} \left(y_j - \sum_{s=1}^{kl} B_{js} g_s - c_j\right)^2\right] \\
&\quad = \sum_{j=1}^{l} \mathrm{E}\left[\left(y_j - \sum_{s=1}^{kl} B_{js} g_s - c_j\right)^2\right] .
\end{aligned} \tag{3.8}$$

This function is to be minimized with respect to $\mathbf{B}$ and $\mathbf{c}$. Since $B_{js}$ and $c_j$ occur solely in the $j$-th summand, the sum is minimized by minimizing each summand separately. Minimization of the $j$-th summand, however, corresponds to the linear regression problem of regressing the target variable $y_j$ on the vector $\mathbf{f}$ of all forecasts for all components using a constant term $c_j$ for $j = 1, \ldots, l$, cf. Rao (1965, pp. 222f.). Consequently, minimization of the SMSPE-function may be regarded as $l$ linear regression problems.

The regression representation facilitates the application of the strong linear plus quadratic combination to empirical data: For the $j$-th regression we consider the regression model $\mathbf{y}^j = \mathbf{X}\boldsymbol{\omega}^j + \mathbf{u}^j$. Here we construct the regression matrix $\mathbf{X}$ from a column of ones (for the constant term) and $kl$ columns with the observations on the vector $\mathbf{f}$ of the individual forecasts: The first observation on $\mathbf{f}$ is written into the first row of $\mathbf{X}$, the second observation on $\mathbf{f}$ is written into the second row of $\mathbf{X}$, and so on. The observations on the target variable $y_j$ yield the vector $\mathbf{y}^j$. The coefficients $\boldsymbol{\omega}^j = (\omega_0^j, \omega_1^j, \dots, \omega_{kl}^j)^{\mathrm{T}}$ are the combination parameters $c_j$ and $(B_{j1}, \dots, B_{j(kl)})^{\mathrm{T}} = (B_{1,j1}, \dots, B_{1,jl}, B_{2,j1}, \dots, B_{2,jl}, \dots, B_{k,j1}, \dots, B_{k,jl})^{\mathrm{T}}$, i.e. the $j$-th component of $\mathbf{c}$ together with the $j$-th row of $\mathbf{B}$. Then we may apply any estimator from linear regression theory, e.g. the ordinary least squares estimator $\widehat{\boldsymbol{\omega}}^j = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}^j$, to estimate the regression parameters and thus the combination parameters.
Since the regression matrix $\mathbf{X}$ is the same in each of the above $l$ univariate linear regressions, they may be incorporated into a multivariate regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Omega} + \mathbf{U}$. Here $\mathbf{Y} = (\mathbf{y}^1 | \dots | \mathbf{y}^l)$ contains the observations on the target variable $\mathbf{y}$ with the first observation in the first row, the second observation in the second row, and so on. Analogously we get the error matrix $\mathbf{U}$ from the vectors $\mathbf{u}^j$. As already indicated the regression matrix $\mathbf{X}$ is the same as above. Finally the parameter matrix $\boldsymbol{\Omega}$ is

$$\boldsymbol{\Omega} = (\boldsymbol{\omega}^1 | \dots | \boldsymbol{\omega}^l) = \begin{pmatrix} c_1 & \dots & c_l \\ B_{11} & \dots & B_{l1} \\ \vdots & & \vdots \\ B_{1(kl)} & \dots & B_{l(kl)} \end{pmatrix} = \begin{pmatrix} \mathbf{c}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} \end{pmatrix}, \tag{3.9}$$

and its ordinary least squares estimator is given by $\widehat{\boldsymbol{\Omega}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$.
As mentioned above using the ordinary least squares estimator leads to the same results as replacing the true moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by the respective sample moments in the formulae for the optimal combination parameters derived in TROSCHKE (2002). Compared to the univariate treatment of each of the $l$ variables, where $y_j$ would solely be regressed on the corresponding forecasts $f_{1,j}, \dots, f_{k,j}$, additional explanatory variables have been included in each of the $l$ regressions. Consequently, the strong multivariate linear combination is superior in theory and, thus, has the potential to outperform the univariate treatment of each variable in empirical applications.

**Case 1b.** *Unrestricted combination without constant term* $\mathbf{f_B} = \mathbf{Bf}$:
This case is completely analogous to the previous one. Minimizing the SMSPE-function can be viewed as the composition of $l$ univariate linear regressions, where $y_j$ is regressed on the vector $\mathbf{f}$ *without* a constant term for $j = 1, \dots, l$.

In the data model $\mathbf{y}^j = \mathbf{X}\boldsymbol{\omega}^j + \mathbf{u}^j$ the regression matrix $\mathbf{X}$ does not contain a column of ones, but is the same otherwise. Moreover, $\boldsymbol{\omega}^j = (B_{j1}, \ldots, B_{j(kl)})^{\mathrm{T}}$ such that in the multivariate model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Omega} + \mathbf{U}$ we have $\boldsymbol{\Omega} = \mathbf{B}^{\mathrm{T}}$.

**Case 1c.** *Restricted combination with constant term* $\mathbf{f}_{\mathbf{B},\mathbf{c},\mathrm{rest}} = \mathbf{B}\mathbf{f} + \mathbf{c} = \sum_{i=1}^{k} \mathbf{B}_i \mathbf{f}_i + \mathbf{c}$ with $\sum_{i=1}^{k} \mathbf{B}_i = \mathbf{I}_l$:
Similarly to the previous cases we see that minimizing the SMSPE-function under the restriction $\sum_{i=1}^{k} \mathbf{B}_i = \mathbf{I}_l$ is equivalent to $l$ univariate linear regressions with appropriate restrictions: In the $j$-th regression $y_j$ is regressed on $\mathbf{f}$ with constant term $c_j$ under the restrictions:

$$\begin{aligned} B_{1,jj} + \ldots + B_{k,jj} &= 1 \,, \\ B_{1,ji} + \ldots + B_{k,ji} &= 0 \qquad \text{for } i = 1, \ldots, l, \quad i \neq j \,. \end{aligned} \tag{3.10}$$

The restrictions (3.10) may be written in matrix form as $\mathbf{R}^j \boldsymbol{\omega}^j = \mathbf{r}^j$, where $\boldsymbol{\omega}^j = (\omega_0^j, \omega_1^j, \ldots, \omega_{kl}^j)^{\mathrm{T}} = (c_j, B_{j1}, \ldots, B_{j(kl)})^{\mathrm{T}}$, $\mathbf{R}^j = (\mathbf{0}_{l\times 1} | \mathbf{I}_l | \ldots | \mathbf{I}_l) =: \mathbf{R} \in \mathbb{R}^{l \times kl+1}$ is row regular and independent of $j$ and $\mathbf{r}^j = \mathbf{e}_j$ is the $j$-th unit vector in $\mathbb{R}^l$.
The corresponding data model is $\mathbf{y}^j = \mathbf{X}\boldsymbol{\omega}^j + \mathbf{u}^j$, where $\mathbf{y}^j$ contains the observations on $y_j$, $\mathbf{X}$ contains a vector of ones in the first column and the observations on $\mathbf{f}$ in the next $kl$ columns. $\boldsymbol{\omega}^j$ as well as the restrictions $\mathbf{R}\boldsymbol{\omega}^j = \mathbf{r}^j$ are given above.
For regression parameter and hence combination parameter estimation we may use the restricted least squares estimator

$$\widehat{\boldsymbol{\omega}}_R^j = \widehat{\boldsymbol{\omega}}^j - (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{R}^{\mathrm{T}}[\mathbf{R}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{R}^{\mathrm{T}}]^{-1}(\mathbf{R}\widehat{\boldsymbol{\omega}}^j - \mathbf{r}^j) \,, \tag{3.11}$$

where $\widehat{\boldsymbol{\omega}}^j = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}^j$ is the ordinary least squares estimator.
Since not only $\mathbf{X}$, but also $\mathbf{R}$ is independent of $j$, we may write the $l$ univariate regressions in a more compact way as a multivariate linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Omega} + \mathbf{U}$ under the restriction $\mathbf{R}\boldsymbol{\Omega} = \mathbf{I}_l$. Here the right hand side stems from $\mathbf{I}_l = (\mathbf{e}_1 | \ldots | \mathbf{e}_l) = (\mathbf{r}^1 | \ldots | \mathbf{r}^l)$. Again

$$\mathbf{Y} = (\mathbf{y}^1 | \ldots | \mathbf{y}^l) \quad \text{and} \quad \boldsymbol{\Omega} = (\boldsymbol{\omega}^1 | \ldots | \boldsymbol{\omega}^l) = \begin{pmatrix} \mathbf{c}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} \end{pmatrix} \tag{3.12}$$

and the corresponding multivariate restricted least squares estimator is

$$\widehat{\boldsymbol{\Omega}}_R = \widehat{\boldsymbol{\Omega}} - (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{R}^{\mathrm{T}}[\mathbf{R}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{R}^{\mathrm{T}}]^{-1}(\mathbf{R}\widehat{\boldsymbol{\Omega}} - \mathbf{I}_l) \,, \tag{3.13}$$

with $\widehat{\boldsymbol{\Omega}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$.

**Case 1d.** *Restricted combination without constant term* $\mathbf{f}_{\mathbf{B},\text{rest}} = \mathbf{Bf} = \sum_{i=1}^{k} \mathbf{B}_i \mathbf{f}_i$ with $\sum_{i=1}^{k} \mathbf{B}_i = \mathbf{I}_l$:

Like $\mathbf{f}_\mathbf{B}$ is derived from $\mathbf{f}_{\mathbf{B},\mathbf{c}}$ we may now derive $\mathbf{f}_{\mathbf{B},\text{rest}}$ from $\mathbf{f}_{\mathbf{B},\mathbf{c},\text{rest}}$: Minimizing the SMSPE-function can be viewed as the composition of $l$ univariate linear regressions, where $y_j$ is regressed on the vector $\mathbf{f}$ *without* a constant term but obeying the restrictions (3.10) for $j = 1, \ldots, l$. Since the constant term is missing in this case the restrictions (3.10) translate to matrix form as $\mathbf{R}^j \boldsymbol{\omega}^j = \mathbf{r}^j$, where $\boldsymbol{\omega}^j = (\omega_1^j, \ldots, \omega_{kl}^j)^{\mathrm{T}} = (B_{j1}, \ldots, B_{j(kl)})^{\mathrm{T}}$, $\mathbf{R}^j = (\mathbf{I}_l | \ldots | \mathbf{I}_l) =: \mathbf{R}$ is row regular and independent of $j$ and $\mathbf{r}^j = \mathbf{e}_j$ is the $j$-th unit vector in $\mathbb{R}^l$.

In the data model $\mathbf{y}^j = \mathbf{X}\boldsymbol{\omega}^j + \mathbf{u}^j$ the regression matrix $\mathbf{X}$ does not contain a column of ones, but otherwise is the same as in Case 1c. Moreover, we have $\boldsymbol{\Omega} = \mathbf{B}^{\mathrm{T}}$ in the multivariate model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Omega} + \mathbf{U}$ with the restriction $\mathbf{R}\boldsymbol{\Omega} = \mathbf{I}_l$.

## 3.2 Medium multivariate linear combinations

The medium multivariate linear approaches emerge from restricting the full matrices $\mathbf{B}_i \in \mathbb{R}^{l \times l}$ in the strong approach to diagonal matrices $\mathbf{D}_i$. In Troschke (2002) it is shown that this restriction results in the univariate consideration of each target variable $y_j$ and its corresponding forecasts $f_{1,j}, \ldots, f_{k,j}$. Consequently, the regression models suitable for empirical applications are obvious from Table 1, and the following considerations can be kept short.

We use the additional notation $\mathbf{f}^j := (f_{1,j}, \ldots, f_{k,j})^{\mathrm{T}}$, i.e. $\mathbf{f}^j$ is the $k \times 1$ vector with all forecasts for the target variable $y_j$.

**Case 2a.** *Unrestricted combination with constant term* $\mathbf{f}_{\mathbf{D},\mathbf{c}} = \mathbf{Df} + \mathbf{c}$:

As stated above the medium multivariate approach is equivalent to the univariate treatment of each of the $l$ variables. This can also be seen from the SMSPE-function for $f_{\mathbf{D},\mathbf{c}}$ which is to be minimized with respect to $\mathbf{D}$ and $\mathbf{c}$:

$$
\begin{aligned}
&\text{SMSPE}(\mathbf{f}_{\mathbf{D},\mathbf{c}}, \mathbf{y}) = \mathrm{E}[(\mathbf{y} - \mathbf{f}_{\mathbf{D},\mathbf{c}})^{\mathrm{T}}(\mathbf{y} - \mathbf{f}_{\mathbf{D},\mathbf{c}})] \\
&\quad = \mathrm{E}[(\mathbf{y} - \mathbf{Df} - \mathbf{c})^{\mathrm{T}}(\mathbf{y} - \mathbf{Df} - \mathbf{c})] \\
&\quad = \mathrm{E}\left[\sum_{j=1}^{l} (y_j - (\mathbf{Df})_j - c_j)^2\right] \\
&\quad = \mathrm{E}\left[\sum_{j=1}^{l} (y_j - (\mathbf{D}_1\mathbf{f}_1 + \ldots + \mathbf{D}_k\mathbf{f}_k)_j - c_j)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{E}\left[\sum_{j=1}^{l}(y_j - (\mathbf{D}_{1,jj}f_{1,j} + \ldots + \mathbf{D}_{k,jj}f_{k,j}) - c_j)^2\right] \\
&= \mathrm{E}\left[\sum_{j=1}^{l}\left(y_j - \sum_{i=1}^{k}\mathbf{D}_{i,jj}f_{i,j} - c_j\right)^2\right] \\
&= \sum_{j=1}^{l}\mathrm{E}\left[\left(y_j - \sum_{i=1}^{k}\mathbf{D}_{i,jj}f_{i,j} - c_j\right)^2\right] . \qquad (3.14)
\end{aligned}
$$

The combination parameters $D_{i,jj}$ and $c_j$ only occur in the $j$-th summand, so that the sum is minimized by minimizing each summand separately. Furthermore, the $j$-th summand only involves the target variable $y_j$ and the corresponding forecasts $f_{i,j}$ for $i = 1, \ldots, k$. Consequently, minimization of the $j$-th summand corresponds to the linear regression problem of regressing the target variable $y_j$ on the vector $\mathbf{f}^j$ using a constant term $c_j$ (OLSCO linear combination in Table 1). Minimization of the SMSPE-function is achieved by solving these $l$ linear regression problems.
When applying the medium approach to empirical data, for the $j$-th regression we consider the model $\mathbf{y}^j = \mathbf{X}^j\boldsymbol{\omega}^j + \mathbf{u}^j$, $j = 1, \ldots, l$. The regression matrix $\mathbf{X}^j$ contains a column of ones (for the constant term) and $k$ columns with the observations on the individual forecasts for the $j$-th variable. Again, the observations on the target variable $y_j$ yield the vector $\mathbf{y}^j$. The coefficients $\boldsymbol{\omega}^j = (\omega_0^j, \omega_1^j, \ldots, \omega_k^j)^{\mathrm{T}}$ are the combination parameters $c_j$ and $(D_{1,jj}, \ldots, D_{k,jj})^{\mathrm{T}}$, i.e. the $j$-th component of $\mathbf{c}$ together with the respective $j$-th diagonal elements of the submatrices $\mathbf{D}_i$. Then any estimator from linear regression theory may serve to estimate the regression parameters and thus the combination parameters, e.g. the ordinary least squares estimator $\widehat{\boldsymbol{\omega}}^j = (\mathbf{X}^{j\mathrm{T}}\mathbf{X}^j)^{-1}\mathbf{X}^{j\mathrm{T}}\mathbf{y}^j$. Since the regression matrices in each of the $l$ regressions are different, we cannot incorporate them into a multivariate regression model.

Likewise we can identify the regression models for the other three medium multivariate approaches. How the ingredients of the respective data models $\mathbf{y}^j = \mathbf{X}^j\boldsymbol{\omega}^j + \mathbf{u}^j$ should be chosen is obvious. How the restrictions for Cases 2c and 2d can be written in matrix form $\mathbf{R}^j\boldsymbol{\omega}^j = \mathbf{r}^j$ is obvious as well. Like in Case 2a representation via a multivariate regression model is not possible.
The *unrestricted combination set-up without constant term* (**Case 2b**) $\mathbf{f_D} = \mathbf{Df}$ is equivalent to the $l$ univariate linear regressions of regressing $y_j$ on $\mathbf{f}^j$ without constant term (OLS linear combination in Table 1).
The *restricted combination set-up with constant term* (**Case 2c**) $\mathbf{f}_{\mathbf{D},\mathbf{c},\mathrm{rest}} = \mathbf{Df} + \mathbf{c}$

with $\sum_{i=1}^{k} \mathbf{D}_i = \mathbf{I}_l$ is equivalent to the $l$ univariate linear regressions of regressing $y_j$ on $\mathbf{f}^j$ with constant term and with the restriction that the coefficients $D_{i,jj}$, $i = 1, \ldots, k$, of the $f_{i,j}$ should sum up to 1 (ERLSCO linear combination in Table 1).

Finally, the *restricted combination set-up without constant term* (**Case 2d**) $\mathbf{f}_{\mathbf{D},\text{rest}} = \mathbf{Df}$ with $\sum_{i=1}^{k} \mathbf{D}_i = \mathbf{I}_l$ is equivalent to the $l$ univariate linear regressions of regressing $y_j$ on $\mathbf{f}^j$ without constant term and with the restriction that the coefficients $D_{i,jj}$, $i = 1, \ldots, k$, of the $f_{i,j}$ should sum up to 1 (ERLS linear combination in Table 1).

## 3.3 Weak multivariate linear combinations

In the weak linear plus quadratic approaches the full matrices $\mathbf{B}_i$ from the strong approach are restricted to scalar multiples of the $l \times l$ identity matrix $\alpha_i \mathbf{I}_l$. It should be pointed out again, that as a consequence the weak multivariate linear combination reduces the number of combination parameters substantially with respect to the medium and strong approaches. Considering the fairly small amount of data generally available in empirical applications this might be an important advantage. Naturally, there is a price to pay for this reduction in number of parameters as we will discuss in the following. In order to present our results conveniently we will consider the variants without constant term first.

**Case 3b.** *Unrestricted combination without constant term* $\mathbf{f}_{\boldsymbol{\alpha}} = \sum_{i=1}^{k} \alpha_i \mathbf{f}_i$:
The scalar mean square prediction error of such a combined forecast is given as

$$\begin{aligned}
&\text{SMSPE}(\mathbf{f}_{\boldsymbol{\alpha}}, \mathbf{y}) = \text{E}[(\mathbf{y} - \mathbf{f}_{\boldsymbol{\alpha}})^{\mathrm{T}}(\mathbf{y} - \mathbf{f}_{\boldsymbol{\alpha}})] \\
&= \text{E}\left[\left(\mathbf{y} - \sum_{i=1}^{k} \alpha_i \mathbf{f}_i\right)^{\mathrm{T}} \left(\mathbf{y} - \sum_{i=1}^{k} \alpha_i \mathbf{f}_i\right)\right] \\
&= \text{E}\left[\sum_{j=1}^{l} \left(y_j - \left(\sum_{i=1}^{k} \alpha_i \mathbf{f}_i\right)_j\right)^2\right] \\
&= \text{E}\left[\sum_{j=1}^{l} \left(y_j - \sum_{i=1}^{k} \alpha_i f_{i,j}\right)^2\right] && (3.15) \\
&= \sum_{j=1}^{l} \text{E}\left[\left(y_j - \sum_{i=1}^{k} \alpha_i f_{i,j}\right)^2\right] && (3.16)
\end{aligned}$$

and this function has to be minimized with respect to $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)^{\mathrm{T}}$. Since the coefficients $\alpha_i$ occur in each of the $l$ summands it is not possible to reduce the

problem by minimizing each summand separately like it could be done in the two preceding subsections. The same coefficients $\alpha_i$ are used to simultaneously fit the forecasts $f_{1,j},\dots,f_{k,j}$ to the target variable $y_j$ for $j=1,\dots,l$.
From representation (3.15) it can be seen, however, that minimization of this function corresponds to a linear regression problem nevertheless: Regarding all components $y_1,\dots,y_l$ as one variable, $y^*$ say, and all components $f_{i,1},\dots,f_{i,l}$ as one forecast $f_i^*$ for that variable ($i=1,\dots,k$), we have to regress $y^*$ on the vector $\mathbf{f}^*=(f_1^*,\dots,f_k^*)^{\mathrm{T}}$ without constant term. I.e. we perform only one univariate regression. Considering all variables as one is the announced price we have to pay for the reduction in the number of combination parameters (cf. the introduction to this subsection). A discussion of this feature follows at the end of this subsection.
For empirical applications we consider the data model $\mathbf{y}^*=\mathbf{X}^*\boldsymbol{\omega}+\mathbf{u}^*$, where we obtain the regression vector $\mathbf{y}^*$ from stacking all observations on the target vector variable $\mathbf{y}$ atop of each other; i.e. below the first observation on $\mathbf{y}$ we write the second observation on $\mathbf{y}$, then the third observation on $\mathbf{y}$ and so on. Likewise we obtain the $i$-th column of the regression matrix $\mathbf{X}^*$ from stacking the observations on $\mathbf{f}_i$ for $i=1,\dots,k$. The vector $\boldsymbol{\omega}$ of regression coefficients equals the vector of combination parameters $\boldsymbol{\alpha}$. It may be estimated e.g. by the ordinary least squares estimator $\widehat{\boldsymbol{\omega}}=(\mathbf{X}^{*\mathrm{T}}\mathbf{X}^*)^{-1}\mathbf{X}^{*\mathrm{T}}\mathbf{y}^*$.

**Case 3d.** *Restricted combination without constant term* $\mathbf{f}_{\boldsymbol{\alpha},\mathrm{rest}}=\sum_{i=1}^k\alpha_i\mathbf{f}_i$ with $\sum_{i=1}^k\alpha_i=1$:
The SMSPE-function for this case is the same as that given in (3.15) for Case 3b. Here it has to be minimized with respect to $\boldsymbol{\alpha}=(\alpha_1,\dots,\alpha_k)^{\mathrm{T}}$ under the restriction $\sum_{i=1}^k\alpha_i=1$. The corresponding univariate linear regression problem is that of regressing $y^*$ on $f_1^*,\dots,f_k^*$ under the above restriction. Using $\boldsymbol{\omega}=\boldsymbol{\alpha}$ we may write the restriction in matrix form $\mathbf{R}\boldsymbol{\omega}=\mathbf{r}$ with $\mathbf{R}=\mathbf{1}_k^{\mathrm{T}}$ and $\mathbf{r}=1\in\mathbb{R}$.
For empirical applications we may use the data model $\mathbf{y}^*=\mathbf{X}^*\boldsymbol{\omega}+\mathbf{u}^*$ with $\mathbf{y}^*$, $\mathbf{X}^*$ and $\boldsymbol{\omega}$ as in Case 3b. To take the restriction into account we have to use the restricted least squares estimator for parameter estimation:

$$\widehat{\boldsymbol{\omega}}_R=\widehat{\boldsymbol{\omega}}-(\mathbf{X}^{*\mathrm{T}}\mathbf{X}^*)^{-1}\mathbf{1}_k[\mathbf{1}_k^{\mathrm{T}}(\mathbf{X}^{*\mathrm{T}}\mathbf{X}^*)^{-1}\mathbf{1}_k]^{-1}(\mathbf{1}_k^{\mathrm{T}}\widehat{\boldsymbol{\omega}}-1)\ , \tag{3.17}$$

where $\widehat{\boldsymbol{\omega}}=(\mathbf{X}^{*\mathrm{T}}\mathbf{X}^*)^{-1}\mathbf{X}^{*\mathrm{T}}\mathbf{y}^*$.

It might be conjectured that the regression models corresponding to the remaining two cases could be obtained from those of the preceding two cases by adding a constant term to the identified models. But this is not true, as we will see in the following.

**Case 3a.** *Unrestricted combination with constant term* $\mathbf{f}_{\boldsymbol{\alpha},\mathbf{c}} = \sum_{i=1}^{k} \alpha_i \mathbf{f}_i + \mathbf{c}$:
Like in Case 3b we can derive the SMSPE-function as

$$
\begin{aligned}
\mathrm{SMSPE}(\mathbf{f}_{\boldsymbol{\alpha},\mathbf{c}}, \mathbf{y}) &= \mathrm{E}[(\mathbf{y} - \mathbf{f}_{\boldsymbol{\alpha},\mathbf{c}})^{\mathrm{T}}(\mathbf{y} - \mathbf{f}_{\boldsymbol{\alpha},xc})] \\
&= \mathrm{E}\left[\left(\mathbf{y} - \sum_{i=1}^{k} \alpha_i \mathbf{f}_i - \mathbf{c}\right)^{\mathrm{T}} \left(\mathbf{y} - \sum_{i=1}^{k} \alpha_i \mathbf{f}_i - \mathbf{c}\right)\right] \qquad (3.18) \\
&= \mathrm{E}\left[\sum_{j=1}^{l} \left(y_j - \sum_{i=1}^{k} \alpha_i f_{i,j} - c_j\right)^2\right] \\
&= \sum_{j=1}^{l} \mathrm{E}\left[\left(y_j - \sum_{i=1}^{k} \alpha_i f_{i,j} - c_j\right)^2\right] , \qquad (3.19)
\end{aligned}
$$

which has to be minimized with respect to $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)^{\mathrm{T}}$ and $\mathbf{c}$. Obviously, the situation is different from that in Cases 3b and 3d, since $\alpha_1, \ldots, \alpha_k$ occur in each of the $l$ summands whereas $c_j$ only occurs in the $j$-th summand. For this reason, we can *not* interpret the above minimization problem as the univariate linear regression of $y^*$ on $f_1^*, \ldots, f_k^*$ including a constant term. (We will, however, take this idea into account in our empirical investigations in Section 4.)
From Troschke (2002) it is known that the SMSPE-optimal choice for the constant term $\mathbf{c}$ in this case is

$$
\mathbf{c}_{\mathrm{opt}} = \boldsymbol{\mu}_0 - \sum_{i=1}^{k} \alpha_{i,\mathrm{opt}} \boldsymbol{\mu}_i , \qquad (3.20)
$$

where $\alpha_{i,\mathrm{opt}}$ denotes the optimal choice for $\alpha_i$, $\boldsymbol{\mu}_i$ denotes the expectation of the $i$-th forecast $\mathbf{f}_i$ and $\boldsymbol{\mu}_0$ denotes the expectation of the target variable $\mathbf{y}$. Consequently, the above minimization problem is equivalent to minimizing the function

$$
\begin{aligned}
&\mathrm{E}\left[\left((\mathbf{y} - \boldsymbol{\mu}_0) - \sum_{i=1}^{k} \alpha_i (\mathbf{f}_i - \boldsymbol{\mu}_i)\right)^{\mathrm{T}} \left((\mathbf{y} - \boldsymbol{\mu}_0) - \sum_{i=1}^{k} \alpha_i (\mathbf{f}_i - \boldsymbol{\mu}_i)\right)\right] \\
&= \mathrm{E}\left[\sum_{j=1}^{l} \left((y_j - \mu_{0,j}) - \sum_{i=1}^{k} \alpha_i (f_{i,j} - \mu_{i,j})\right)^2\right] \\
&= \sum_{j=1}^{l} \mathrm{E}\left[\left((y_j - \mu_{0,j}) - \sum_{i=1}^{k} \alpha_i (f_{i,j} - \mu_{i,j})\right)^2\right] \qquad (3.21)
\end{aligned}
$$

with respect to $\boldsymbol{\alpha}$. This can be seen by inserting $\mathbf{c} = \boldsymbol{\mu}_0 - \sum_{i=1}^{k} \alpha_i \boldsymbol{\mu}_i$ into Equation (3.18).

As in Case 3b we may now conclude that the univariate linear regression problem for this case is that of regressing $(\mathbf{y} - \boldsymbol{\mu}_0)^*$ onto $(\mathbf{f}_1 - \boldsymbol{\mu}_1)^*, \dots, (\mathbf{f}_k - \boldsymbol{\mu}_k)^*$, where the $^*$ indicates that all components of the vector are regarded as a single variable. No constant term is included in this regression and no restriction is placed on the regression coefficients $\alpha_1, \dots, \alpha_k$. The remaining combination parameter vector $\mathbf{c}$ is obtained from Equation (3.20).

For empirical applications we use the data model $\widetilde{\mathbf{y}} = \widetilde{\mathbf{X}}\boldsymbol{\omega} + \widetilde{\mathbf{u}}$. To obtain $\widetilde{\mathbf{y}}$ we first calculate the estimate $\widehat{\boldsymbol{\mu}}_0$ of the expectation of $\mathbf{y}$ as the mean of all observations on $\mathbf{y}$. Then we adjust all observations on $\mathbf{y}$ by subtracting $\widehat{\boldsymbol{\mu}}_0$ and stack these vectors atop of each other thus arriving at $\widetilde{\mathbf{y}}$. Likewise we adjust the observations on the forecasts $\mathbf{f}_i$ by subtracting the corresponding mean $\widehat{\boldsymbol{\mu}}_i$. If we stack these adjusted vectors we obtain the $i$-th column of $\widetilde{\mathbf{X}}$. The regression coefficients $\boldsymbol{\omega}$ equal the combination parameters $\boldsymbol{\alpha}$ and they may be estimated by the ordinary least squares estimator $\widehat{\boldsymbol{\omega}} = \widehat{\boldsymbol{\alpha}} = (\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{y}}$. Finally, the parameter vector $\mathbf{c}$ is estimated by

$$\widehat{\mathbf{c}} = \widehat{\boldsymbol{\mu}_0} - \sum_{i=1}^{k} \widehat{\alpha}_i \widehat{\mu}_i \; . \tag{3.22}$$

**Case 3c.** *Restricted combination with constant term* $\mathbf{f}_{\boldsymbol{\alpha},\mathbf{c},\mathrm{rest}} = \sum_{i=1}^k \alpha_i \mathbf{f}_i + \mathbf{c}$ with $\sum_{i=1}^k \alpha_i = 1$:

This case follows from Case 3a like Case 3d follows from Case 3b. Again we may use $\mathbf{c} = \boldsymbol{\mu}_0 - \sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i$ and, thus, we have to regress $(\mathbf{y} - \boldsymbol{\mu}_0)^*$ onto $(\mathbf{f}_1 - \boldsymbol{\mu}_1)^*, \dots, (\mathbf{f}_k - \boldsymbol{\mu}_k)^*$ but now under the restriction that the regression parameters sum up to 1.

The corresponding data model $\widetilde{\mathbf{y}} = \widetilde{\mathbf{X}}\boldsymbol{\omega} + \widetilde{\mathbf{u}}$ is built in exactly the same way as in Case 3a. In order to heed the restrictions we now have to use the restricted least squares estimator

$$\widehat{\boldsymbol{\omega}}_R = \widehat{\boldsymbol{\omega}} - (\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}})^{-1}\mathbf{1}_k[\mathbf{1}_k^{\mathrm{T}}(\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}})^{-1}\mathbf{1}_k]^{-1}(\mathbf{1}_k^{\mathrm{T}}\widehat{\boldsymbol{\omega}} - 1) \tag{3.23}$$

with $\widehat{\boldsymbol{\omega}} = (\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{y}}$. Thus the estimate for the combination parameters $\boldsymbol{\alpha}$ is obtained: $\widehat{\boldsymbol{\alpha}} = \widehat{\boldsymbol{\omega}}_R$. Again the parameter vector $\mathbf{c}$ is estimated by

$$\widehat{\mathbf{c}} = \widehat{\boldsymbol{\mu}_0} - \sum_{i=1}^{k} \widehat{\alpha}_i \widehat{\mu}_i \; . \tag{3.24}$$

Since the sum of the $\alpha_i$ equals 1 we may write

$$\boldsymbol{\mu}_0 = \sum_{i=1}^{k} \alpha_i \boldsymbol{\mu}_0 \tag{3.25}$$

and hence we may alternatively interpret this combination as the regression of $y^*$ on the bias-corrected $(\mathbf{f}_i - \boldsymbol{\mu}_i + \boldsymbol{\mu}_0)^*$ under the then reasonable restriction of the regression parameters summing up to 1, compare Equation (3.21).

In the weak combinations *without* constant term (Cases 3b and 3d) the target variables $y_1, \ldots, y_l$ are treated as one single target variable $y^*$ and the forecasts $f_{i,1}, \ldots, f_{i,l}$ are treated as a single forecast $f_i^*$ for $y^*$. In the case of weak combinations *with* constant term (Cases 3a and 3c) the possibly different expectations of the $y_j$ and the $f_{i,j}$ are corrected for in the first place.
It is obvious that in general the weak approach is only reasonable if the (possibly corrected) target variables are similar enough so that their treatment as one single target variable seems justified. This may be the case if, e.g., all components of $\mathbf{y}$ are relative changes of certain quantities with respect to the preceding period. If the above treatment is justified, i.e. if all components follow the same regression equation, application of the weak multivariate linear combination approach is very advantageous: We do not try to estimate several regression relationships where there only is one. Furthermore, we have $l$ times the number of observations to estimate that one relationship, so that the relationship can be determined more precisely.

## 3.4 Multivariate linear adjustment

In Troschke (2002) the above combination models are also applied to the special case of only $k = 1$ forecast. The results are improvements of this individual forecast, named *adjustments*. The corresponding regression models can be derived from the results of the preceding subsections. Again the medium adjustments are equivalent to the corresponding univariate adjustments, whereas for the weak adjustments we have to consider each component of the target variable as one variable and each component of the forecast as forecasts for that one variable.
It should be noted that there are only seven different adjustments since for $k = 1$ Cases 1c, 2c and 3c coincide, while Cases 1d, 2d and 3d coincide with the individual forecast.

In Section 4 we will report an empirical example comparing the performance of the above linear adjustments and combinations in the case of $k = 2$ forecasts for $l = 2$ variables.

| Year | 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | 5.6 | 2.8 | 3.2 | 4.6 | 1.9 | 0.1 | -1.0 | 1.0 | 2.6 | 2.6 | 2.6 |
| $y_2$ | 3.6 | 3.1 | 3.4 | 3.2 | 1.7 | -1.2 | -2.2 | 1.1 | 0.6 | 1.8 | 4.3 |
| $f_{1,1}$ | 5 | 5.5 | 3 | 4 | 2 | -1 | 0.5 | -0.5 | 2 | 2 | 3 |
| $f_{1,2}$ | 3 | 4.5 | 3 | 3.5 | 2 | 1 | -0.5 | -0.5 | 0 | 1.5 | 3.5 |
| $f_{2,1}$ | 4 | 4.5 | 3 | 4 | 2 | -1 | 1 | -0.5 | 2.25 | 2.25 | 3 |
| $f_{2,2}$ | 2.5 | 4.5 | 3 | 3.5 | 1.5 | 1 | 0 | -0.5 | 1 | 1.5 | 3 |
| Year | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | |
| $y_1$ | 1.9 | 3.7 | 3.3 | 4.7 | 3.7 | 1.6 | -1.7 | 2.4 | 1.9 | 1.4 | |
| $y_2$ | 3.5 | 2.7 | 1.7 | 4.7 | 3.6 | 1.7 | 0.2 | 0.6 | 1.8 | 1.3 | |
| $f_{1,1}$ | 1.5 | 1 | 2.5 | 3.5 | 3.5 | 1 | -1 | -0.5 | 2 | 1 | |
| $f_{1,2}$ | 3 | 3 | 2 | 3.5 | 3.5 | 2 | 0 | -1.5 | 0.5 | 2 | |
| $f_{2,1}$ | 2.25 | 1 | 2.25 | 3 | 3.25 | 1.5 | -0.5 | 1 | 3 | 1.75 | |
| $f_{2,2}$ | 3.5 | 2.5 | 2.5 | 4 | 3 | 2 | 0 | -1 | 0.5 | 2.5 | |

Table 2: Real change of German gross national product ($y_1$) and real change of German private consumption ($y_2$) together with corresponding DIW and Ifo forecasts ($\mathbf{f}_1$, $\mathbf{f}_2$) for the period from 1976 to 1996

# 4 An empirical example

In this section we will present an empirical example illustrating the various adjustments of single forecasts as well as the combination of $k = 2$ forecasts for $l = 2$ variables on the basis of the new methods. It should be pointed out, however, that this example is only meant to provide a first impression of the possible usefulness of the multivariate linear approaches. Detailed analyses are bound to follow, and they will be presented in a future paper.

The data for the numerical example are taken from a larger data set of German macro economic variables and corresponding forecasts investigated by Klapper (1998). See Klapper (1999) for a multivariate treatment of these data by means of rank based methods and Wenzel (1999) for a treatment on the basis of a multivariate Pitman closeness criterion. We selected the DIW (Deutsches Institut für Wirtschaftsforschung, $\mathbf{f}_1 = (f_{1,1}, f_{1,2})^{\mathrm{T}}$) and Ifo (Ifo-Institut für Wirtschaftsforschung, $\mathbf{f}_2 = (f_{2,1}, f_{2,2})^{\mathrm{T}}$) forecasts for the target variables 'real change of gross national product' ($y_1$) and 'real change of private consumption' ($y_2$). These yearly data are available for a period of 21 years from 1976 to 1996. They are given in Table 2.

When evaluating the data it is important to take their availability into account: The forecasts $\mathbf{f}_1$ and $\mathbf{f}_2$ for year $t$, say, are made at the end of year $t-1$ and the true value of the target variable $\mathbf{y}$ for the year $t-2$ is not published by the Statistisches Bundesamt before the end of year $t-1$. Consequently, at the time when the individual forecasts for year $t$ are to be combined, namely at the end of year $t-1$, we can only use the past data up to year $t-2$.
These past data serve to estimate the optimal combination parameters at each time point on the basis of the regression models for empirical data from the previous sections. Due to structural changes in the data set the optimal combination parameters may not be stable over time. A common procedure in this situation is to use only the latest observations for parameter estimation. Of course the amount of past data should not be too small either so that the regression fit is at least fairly reasonable. As a compromise we chose a history of 10 data points for parameter estimation.
Altogether we will use the data from 1976 to 1985 to estimate the combination parameters for the 1987 forecasts, the data from 1977 to 1986 to estimate the combination parameters for the 1988 forecasts, and so on. This leads to a time span of 10 years (1987 to 1996) in which the performance of the various methods is evaluated with the average sum of the squared forecast errors

$$\widehat{\mathrm{SMSPE}}(\mathbf{f}_{\cdot}, \mathbf{y}) = \frac{1}{10} \sum_{t=1987}^{1996} (\mathbf{y}_t - \mathbf{f}_{\cdot,t})^{\mathrm{T}} (\mathbf{y}_t - \mathbf{f}_{\cdot,t}) \ , \tag{4.1}$$

where the subscript $t$ indicates the considered year. This is the empirical counterpart of the scalar mean square prediction error

$$\mathrm{SMSPE}(\mathbf{f}_{\cdot}, \mathbf{y}) = \mathrm{E}[(\mathbf{y} - \mathbf{f}_{\cdot})^{\mathrm{T}} (\mathbf{y} - \mathbf{f}_{\cdot})] \ , \tag{4.2}$$

hence the denotation $\widehat{\mathrm{SMSPE}}$.

When dealing with the weak multivariate linear combinations we will not only consider the combinations without constant term (Cases 3b and 3d) and with individual constant terms for each component (Cases 3a and 3c) but also the intermediate choice of one single constant term for all $l$ components. It arises from regressing $y^*$ on $f_1^*, \dots, f_k^*$ including a constant term $c^*$ either with or without the restriction that the coefficients of the $f_i^*$ sum up to one. Then the combined forecasts are given as

$$\mathbf{f}_{\boldsymbol{\alpha}, c^* \mathbf{1}} = \sum_{i=1}^{k} \alpha_i \mathbf{f}_i + c^* \mathbf{1}_l \tag{4.3}$$

or

$$\mathbf{f}_{\boldsymbol{\alpha},c^*\mathbf{1},\text{rest}} = \sum_{i=1}^{k} \alpha_i \mathbf{f}_i + c^* \mathbf{1}_l \quad \text{with} \quad \sum_{i=1}^{k} \alpha_i = 1 \ , \tag{4.4}$$

respectively. Compare Section 3.3.

**Projections:** A feature of the linear approaches to the univariate combination of forecasts is that the linearly combined forecasts can lie outside the range of the individual forecasts. This feature is desirable because even if every single individual forecast underestimates, say, the target variable, the combined forecast can still hit the target variable. On the other hand, in some empirical applications it has been observed that linear combinations do not perform too well, because they produce combined forecasts far out of the range of the individual forecasts. Consequently, the mentioned feature is desirable as well as critical for the linear approaches at the same time. These facts apply to the linear combination of multivariate forecasts as well.
We will try to find a compromise between the good and the bad side of this feature by considering three *projections* of each combined forecast: The *first* projection will restrict each component $f_{\text{comb},j}$ of the combined forecast to the range of the corresponding individual forecasts

$$[\min_{i\in\{1,\dots,k\}} f_{i,j}, \max_{i\in\{1,\dots,k\}} f_{i,j}] =: [f_{\min,j}, f_{\max,j}] \ ,$$

i.e. if $f_{\text{comb},j}$ is smaller than the minimum it is set to the minimum and if $f_{\text{comb},j}$ is larger than the maximum it is set to the maximum. Likewise the *second* projection restricts the $j$-th component of the combined forecast to the interval

$$[f_{\min,j} - 0.1(f_{\max,j} - f_{\min,j}), f_{\max,j} + 0.1(f_{\max,j} - f_{\min,j})]$$

such that the combined forecast is allowed to lie slightly (by ten percent of the range) outside the range of the individual forecasts. Finally, the *third* projection restricts the $j$-th component of the combined forecast to the interval

$$[f_{\min,j} - 0.3(f_{\max,j} - f_{\min,j}), f_{\max,j} + 0.3(f_{\max,j} - f_{\min,j})]$$

thus allowing the combined forecast to leave the range of the individual forecasts by thirty percent of the range.

A very simple strategy for the combination of the single forecasts is their arithmetic mean. Since it is easy to apply and also quite successful in empirical investigations, any other combination technique is measured against the arithmetic mean. Therefore we decided to present all $\widehat{\mathrm{SMSPE}}$-values relative to the $\widehat{\mathrm{SMSPE}}$-value of the arithmetic mean, which is 2.3900 in the considered time period. All decimals have been deleted following the fourth decimal such that methods outperforming the arithmetic mean can be identified immediately.

The results of this evaluation are presented in Table 3. Values in brackets are those belonging to the first, second and third projection, respectively. It can be seen that in this example the weak combination including a full constant term but without restriction on the $\alpha_i$ is the best of all combination methods. It outperforms the arithmetic mean by 10% and the medium approaches which are equivalent to the classical univariate treatment by 13% or more. Also both individual forecasts are outperformed. The strong combinations all perform worse than the arithmetic mean and also worse than both individual forecasts. The strong combinations without restriction on the combination weights perform far worst of all methods. A projection substantially improves their results but they are still worse than the arithmetic mean. We can observe that all weak combinations perform better than their medium counterparts which in turn outperform their strong counterparts.

The adjustment $\mathbf{Bf}_i + \mathbf{c}$ of individual forecasts is very effective for the DIW forecast but not advantageous for the Ifo forecast. For the DIW forecast the medium and weak versions produce the best forecasts among all considered methods outperforming the arithmetic mean by 24% and the DIW forecast by even 34%. For the Ifo forecast the strong and weak versions at least do not attenuate the result of the Ifo forecast very far.

It is interesting to note that the combinations without constant term and with restriction on the sum of the forecast weights perform relatively bad. These are the standard combinations under the often assumed unbiasedness of the individual forecasts. The linear unrestricted adjustments and combinations without constant term perform worst in their respective groups. The latter is especially remarkable, since these methods exhibit an undesired theoretical behaviour as outlined in Troschke (2002).

Concerning the three types of projections we may state that the projection works the better the narrower the projection interval is. In general, restricting the combined forecast to the range of the individual forecasts works best, but the differences between the results of the projections are not very large in this example. Compared to the original combination we may say that for a bad original combination it is advantageous to project, whereas for a good original combination it is not advantageous to project.

| Forecast $f_{\cdot}$ | | $\widehat{\text{SMSPE}}(f_{\cdot}, y)$ |
|---|---|---|
| DIW forecast | $\mathbf{f}_1 = \mathbf{f}_{\text{DIW}}$ | 1.1548 |
| Adjustments: | $f_{\text{DIW},\widehat{\mathbf{B}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.8248 |
| | $f_{\text{DIW},\widehat{\mathbf{B}}_{\text{opt}}}$ | 1.2170 |
| | $f_{\text{DIW},\mathbf{I},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.9457 |
| | $f_{\text{DIW},\widehat{\mathbf{D}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.7631 |
| | $f_{\text{DIW},\widehat{\mathbf{D}}_{\text{opt}}}$ | 1.2352 |
| | $f_{\text{DIW},\widehat{\alpha}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.7669 |
| | $f_{\text{DIW},\widehat{\alpha}_{\text{opt}}}$ | 1.2399 |
| Ifo forecast | $\mathbf{f}_2 = \mathbf{f}_{\text{Ifo}}$ | 0.9916 |
| Adjustments: | $f_{\text{Ifo},\widehat{\mathbf{B}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.9950 |
| | $f_{\text{Ifo},\widehat{\mathbf{B}}_{\text{opt}}}$ | 1.0775 |
| | $f_{\text{Ifo},\mathbf{I},\widehat{\mathbf{c}}_{\text{opt}}}$ | 1.0383 |
| | $f_{\text{Ifo},\widehat{\mathbf{D}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 1.1038 |
| | $f_{\text{Ifo},\widehat{\mathbf{D}}_{\text{opt}}}$ | 1.1358 |
| | $f_{\text{Ifo},\widehat{\alpha}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 1.0212 |
| | $f_{\text{Ifo},\widehat{\alpha}_{\text{opt}}}$ | 1.1138 |
| Strong combinations: | $f_{\widehat{\mathbf{B}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 1.8465 (1.1805, 1.2287, 1.3403) |
| | $f_{\widehat{\mathbf{B}}_{\text{opt}}}$ | 1.8980 (1.1727, 1.2216, 1.3160) |
| | $f_{\widehat{\mathbf{B}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}},\text{rest}}$ | 1.1746 (1.1511, 1.1599, 1.1691) |
| | $f_{\widehat{\mathbf{B}}_{\text{opt}},\text{rest}}$ | 1.2344 (1.2106, 1.2308, 1.2513) |
| Medium combinations: | $f_{\widehat{\mathbf{D}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 1.0300 (1.0784, 1.0790, 1.0881) |
| | $f_{\widehat{\mathbf{D}}_{\text{opt}}}$ | 1.2010 (1.1627, 1.1680, 1.1734) |
| | $f_{\widehat{\mathbf{D}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}},\text{rest}}$ | 1.0834 (1.0720, 1.0762, 1.0913) |
| | $f_{\widehat{\mathbf{D}}_{\text{opt}},\text{rest}}$ | 1.1399 (1.1314, 1.1290, 1.1317) |
| Weak combinations: | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ | 0.9015 (0.9644, 0.9695, 0.9754) |
| | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}}}$ | 1.1808 (1.1034, 1.1148, 1.1251) |
| | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}},\text{rest}}$ | 0.9653 (0.9784, 0.9858, 1.0004) |
| | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\text{rest}}$ | 1.0577 (1.0626, 1.0590, 1.0577) |
| Weak combinations with scalar constant term $c^*$: | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\widehat{c}^*_{\text{opt}}\mathbf{1}}$ | 0.9204 (0.9825, 0.9861, 0.9971) |
| | $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\widehat{c}^*_{\text{opt}}\mathbf{1},\text{rest}}$ | 0.9917 (0.9813, 0.9750, 0.9708) |

Table 3: $\widehat{\text{SMSPE}}$-values of adjusted and combined forecasts in an empirical application (all values relative to the $\widehat{\text{SMSPE}}$ of the arithmetic mean)

The weak combinations utilizing the constant $c^*\mathbf{1}_l$ work better than the arithmetic mean but a little worse than their respective counterparts utilizing a full constant vector $\mathbf{c}$.

The forecasts for the years 1987 to 1996 produced by the weak combination $f_{\widehat{\boldsymbol{\alpha}}_{\text{opt}},\widehat{\mathbf{c}}_{\text{opt}}}$ of $\mathbf{f}_1$ and $\mathbf{f}_2$ are given by 1.5511, 1.2330, 2.9409, 4.2146, 4.0524, 1.5098, -0.4429, 0.4665, 2.6490 and 2.1205 for target variable $y_1$ and by 2.3581, 2.5069, 1.7916, 2.9704, 3.5148, 2.1468, 0.4856, -0.2157, 0.6412 and 2.6437 for target variable $y_2$. Together with the respective target variable, individual forecasts and their arithmetic mean they are visualized in Figure 1.

In TROSCHKE (2002) we have conducted an analysis of the potential of the various combination techniques in terms of expected scalar squared error loss (SMSPE). That analysis was based on known first and second order moments of the joint distribution of $\mathbf{y}$ and $\mathbf{f} = (\mathbf{f}_1^{\mathrm{T}}, \mathbf{f}_2^{\mathrm{T}})^{\mathrm{T}}$. It constitutes the theoretical counterpart of the current empirical example in that the moments considered in that analysis are the sample moments obtained from all 21 observations in Table 2.

Comparing the results of the current empirical investigation to the theoretical counterparts we observe that only the results for the weak combinations and adjustments come close to their expected values. The medium and especially the strong approaches perform far worse than could have been expected for known first and second order moments. The extraordinary results for the medium and weak adjustments of the DIW forecast are far better than could have been expected, especially when taking into consideration that the necessity to estimate the optimal combination parameters leads to an even worse theoretical SMSPE.

A possible reason for the results observed is the very small amount of data available for parameter estimation: only 10 data points seem to be a small basis and favor the methods with only few parameters.

It can be seen that some of the new approaches are able to outperform the arithmetic mean. How good their performance is in general will depend on how good the regression reflects the future relationship between target variable $\mathbf{y}$ and forecasts $\mathbf{f}_i$. Clearly, the more suitable data are available for that regression, the better. Also the data should not be subject to extreme structural changes during the considered period. Consequently, the multivariate linear approaches should be more valuable for monthly, weekly or even daily data (e.g. from the stock market) than they are for yearly data.

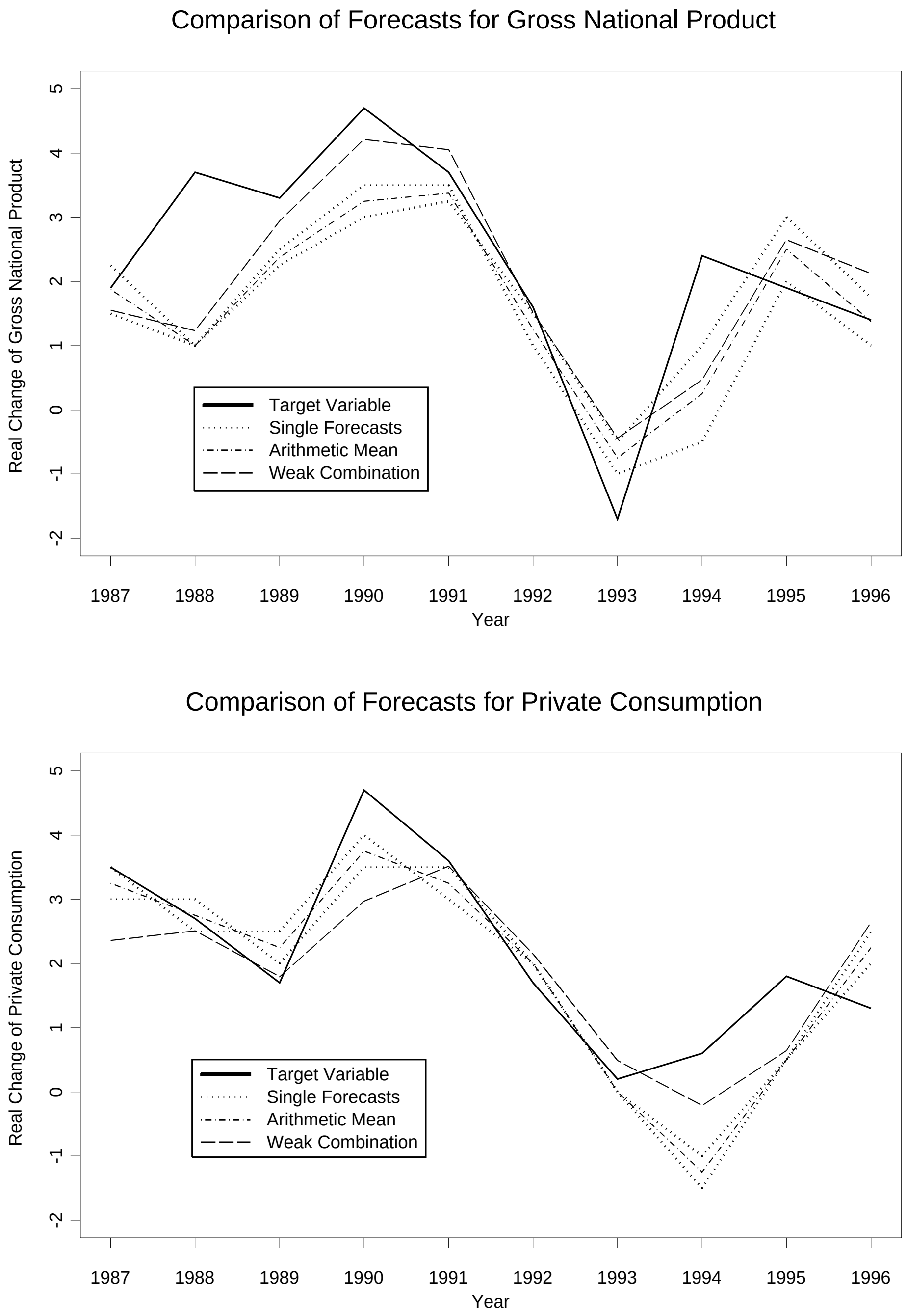


Figure 1: Target variables *real change of gross national product* (top) and *real change of private consumption* (bottom), together with their DIW and Ifo forecasts, their arithmetic mean and their weak multivariate linear combination $f_{\widehat{\boldsymbol{\alpha}}_{\rm opt},\widehat{\mathbf{c}}_{\rm opt}}$.

## 5 Conclusions

In this paper we have introduced the linear regression approach for the multivariate linear combination of forecasts. We have also considered the classical univariate linear approaches as competitors to the new set-up as well as adjustments of individual forecasts which emerge from the special case $k = 1$. The advantage of the regression approach is that it facilitates empirical application of the new methods by making use of standard software packages possible. Moreover, the identification of appropriate regression models constitutes an analogy to the univariate case.

We found out that the strong multivariate linear combination may be interpreted as the $l$ regressions of the target variables $y_j$ on the vector of all forecasts for all variables. These regressions may be incorporated into a multivariate linear regression model. For the medium multivariate linear combination the $y_j$ are solely regressed on the vector of all forecasts for the $j$-th variable. This approach corresponds to the standard univariate treatment of each of the components. Finally, the weak multivariate linear combination results in the treatment of all components $y_j$ as a single variable $y^*$. Then in two cases $y^*$ is regressed on the $f_i^*$, while in the other two cases $(\mathbf{y} - \boldsymbol{\mu}_0)^*$ is regressed on the $(\mathbf{f}_i - \boldsymbol{\mu}_i)^*$.

Furthermore, we have reported an empirical example comparing the classical and the new approaches. We have seen that employing multivariate linear adjustments and combinations may be beneficial, but also that this is not always the case. Due to the smaller number of parameters involved the weak multivariate linear combinations and adjustments seem to be suitable if only a small amount of data is available for parameter estimation.

A much more detailed analysis of the possible benefits of the multivariate linear approaches has to follow, as was explained in Section 4. It will be carried out in the dissertation thesis of the author. A point of special interest would be to find a guideline for potential users identifying situations beforehand in which multivariate linear combination of forecasts is promising. Especially the question of how much data should be available is interesting. Another point is to find out how large the number of forecasts $k$ and the number of variables $l$ should be chosen depending on the amount of data available for the multivariate linear approaches.

**Acknowledgements:** The author wishes to thank Götz Trenkler and Jürgen Groß for their helpful comments and suggestions. The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of complexity in multivariate data structures") is gratefully acknowledged.

# References

CLEMEN, R.T. (1989): 'Combining forecasts: A review and annotated bibliography'. *International Journal of Forecasting* **5**, 559-583.

GRANGER, C.W.J. (1989): 'Combining forecasts – Twenty years later'. *Journal of Forecasting* **8**, 167-173.

GRANGER, C.W.J. and RAMANATHAN, R. (1984): 'Improved methods of combining forecasts'. *Journal of Forecasting* **3**, 197-204.

KLAPPER, M. (1998): Combining German Macro Economic Forecasts Using Rank-Based Techniques. *Technical Report 19/1998, Sonderforschungsbereich 475, University of Dortmund.*

KLAPPER, M. (1999): Multivariate Rank-Based Forecast Combining Techniques. *Technical Report 2/1999, Sonderforschungsbereich 475, University of Dortmund.*

RAO, C.R. (1965): Linear Statistical Inference and Its Applications. *Wiley, New York.*

RAO, C.R. (1966): 'Covariance adjustment and related problems in multivariate analysis'. In: KRISHNAIAH, P.R. (ed.), Multivariate Analysis. *Academic Press, New York*, 87-103.

RAO, C.R. (1967): 'Least squares theory using an estimated dispersion matrix and its application to measurement of signals'. In: LECAM, L.M. and NEYMAN, J. (eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1. *University of California Press, Berkeley, CA*, 355-372.

THIELE, J. (1993): Kombination von Prognosen. (Wirtschaftswissenschaftliche Beiträge: Band 74). *Physica, Heidelberg.*

TRENKLER, G. and IHORST, G. (1995): 'Improved estimation by weak covariance adjustment technique'. *Discussiones Mathematicae Algebra and Stochastic Methods* **15**, 189-201.

TROSCHKE, S.O. (2002): Scalar Mean Square Error Optimal Linear Combination of Multivariate Forecasts. *Technical Report 3/2002, Sonderforschungsbereich 475, University of Dortmund.*

WENZEL, T. (1999): Using Different Pitman-Closeness Techniques for the Linear Combination of Multivariate Forecasts. *Technical Report 18/1999, Sonderforschungsbereich 475, University of Dortmund.*