Appendix A: Detailed derivations of propositions I & II

The Proposition I derivation begins with equation (6) from the literature:

$$EDF = Trace\left(\frac{\partial \hat{Y}}{\partial Y}\right). \quad (6)$$

Proposition I equation

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace\left(X\left(\hat{\beta}^T \otimes I_p\right)\frac{\partial q}{\partial Y}\right). \quad (7)$$

Equation (4) is substituted into equation (6) resulting in equation (8).

$$EDF = Trace\left(\frac{\partial}{\partial Y} Xq\hat{\beta}\right) \quad (8)$$

Equation (8) is differentiated using equation (10).

$$EDF = Trace\left(X\left(\left(\hat{\beta}^T \otimes I_p\right)\frac{\partial q}{\partial Y} + (I_1 \otimes q)\frac{\partial \hat{\beta}}{\partial Y}\right)\right). \quad (11)$$

$$EDF = Trace\left(Xq\frac{\partial \hat{\beta}}{\partial Y}\right) + Trace\left(X\left(\hat{\beta}^T \otimes I_p\right)\frac{\partial q}{\partial Y}\right). \quad (12)$$

$$EDF = Trace\left(A\frac{\partial \hat{\beta}}{\partial Y}\right) + Trace\left(X\left(\hat{\beta}^T \otimes I_p\right)\frac{\partial q}{\partial Y}\right), \quad (12.1)$$

where A is an n by p block matrix $[X_{m_1} q_{m_1}, \quad \cdots \quad X_{m_M} q_{m_M}]$ of M blocks $X_{m_i} q_{m_i}$ of dimension n by p_{m_i} and where $q_{m_1}, \dots q_{m_M}$ are scalars.

The matrix $\frac{\partial \hat{\beta}}{\partial Y}$ is a p by n block matrix $\begin{bmatrix} \frac{\partial \hat{\beta}_{m_1}}{\partial Y} \\ \vdots \\ \frac{\partial \hat{\beta}_{m_M}}{\partial Y} \end{bmatrix}$ of M blocks $\frac{\partial \hat{\beta}_{m_i}}{\partial Y}$ of dimension p_{m_i} by n.

By the properties of block multiplication, $A\frac{\partial \hat{\beta}}{\partial Y} = \sum_{i=1}^{M} q_{m_i} X_{m_i} \frac{\partial \hat{\beta}_{m_i}}{\partial Y}$,

where the scalar q_i commutes to the left side of X_{m_i}.

$$EDF = Trace\left(\sum_{i=1}^{M} q_{m_i} X_{m_i} \frac{\partial \hat{\beta}_{m_i}}{\partial Y}\right) + Trace\left(X\left(\hat{\beta}^T \otimes I_p\right)\frac{\partial q}{\partial Y}\right). \quad (12.2)$$

$$EDF = Trace \left(\sum_{i=1}^{M} q_{m_i} \frac{\partial X_{m_i} \hat{\beta}_{m_i}}{\partial Y} \right) + Trace \left(X \left(\hat{\beta}^T \otimes I_p \right) \frac{\partial q}{\partial Y} \right). \tag{12.3}$$

Using the properties of the Trace function and substituting $X_{m_i} \hat{\beta}_{m_i}$ $with$ $\hat{Y}_{m_i}$,

$$EDF = \sum_{i=1}^{M} q_{m_i} \, Trace \left(\frac{\partial \hat{Y}_{m_i}}{\partial Y} \right) + Trace \left(X \left(\hat{\beta}^T \otimes I_p \right) \frac{\partial q}{\partial Y} \right). \tag{13}$$

Substituting $Trace \left(\frac{\partial \hat{Y}_{m_i}}{\partial Y} \right)$ with EDF_{m_i} by using equation (6) results in Proposition I.

Proposition II equation

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + EDF_w + Trace \left(X_A \hat{\beta}_{\Delta q} \right). \tag{14}$$

Proposition II is derived from Proposition I.

Starting with equation (7).

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace \left(D \frac{\partial q}{\partial Y} \right), \tag{14.1}$$

where $D = X \left(\hat{\beta}^T \otimes I_p \right)$ is an n by p^2 block matrix $[X\hat{\beta}_1 \quad \cdots \quad X\hat{\beta}_P]$, of P blocks $X\hat{\beta}_i$ of dimensions n by p with scalars $\hat{\beta}_1, \dots, \hat{\beta}_p$.

The matrix $\frac{\partial q}{\partial Y}$ is a p^2 by n matrix as a result of the diagonal p by p matrix q being vectorized and differentiated using the convention articulated in equation (9).

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace\ (E), \tag{14.2}$$

where $E = D \frac{\partial q}{\partial Y}$ is an n by n matrix
$$\begin{bmatrix} \sum_{j=1}^{P} X_{1j} \hat{\beta}_j \frac{\partial q_j}{\partial Y_1} & \cdots & \sum_{i=1}^{P} X_{1j} \hat{\beta}_j \frac{\partial q_j}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{P} X_{nj} \hat{\beta}_j \frac{\partial q_j}{\partial Y_1} & \cdots & \sum_{i=1}^{P} X_{nj} \hat{\beta}_j \frac{\partial q_j}{\partial Y_n} \end{bmatrix} \tag{14.3}$$

and X_{1j} is the $1\,j$ entry of matrix X.

$$EDF = \sum_{i=1}^{M} q_{m_i}\, EDF_{m_i} + Trace\left(\begin{bmatrix} \sum_{j=1}^{P} X_{1j}\hat{\beta}_j \frac{\partial q_j}{\partial Y_1} & \cdots & \sum_{i=1}^{P} X_{1j}\hat{\beta}_j \frac{\partial q_j}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{P} X_{nj}\hat{\beta}_j \frac{\partial q_j}{\partial Y_1} & \cdots & \sum_{i=1}^{P} X_{nj}\hat{\beta}_j \frac{\partial q_j}{\partial Y_n} \end{bmatrix}\right). \tag{15}$$

By grouping terms by underlying auxiliary models and using the fact that $\frac{\partial q_j}{\partial Y_1}$ is constant within the sum for each auxiliary model,

$$\sum_{j=1}^{P} X_{1j}\hat{\beta}_j \frac{\partial q_j}{\partial Y_1} = \sum_{i=1}^{M}\sum_{k=1}^{p_{m_i}} \left(X_{1k}\hat{\beta}_k \frac{\partial q_k}{\partial Y_1}\right) = \sum_{i=1}^{M}\sum_{k=1}^{p_{m_i}} \left(X_{1k}\hat{\beta}_k\right) \frac{\partial q_{m_i}}{\partial Y_1} = \sum_{i=1}^{M} \hat{Y}_{1i} \frac{\partial q_{m_i}}{\partial Y_1}, \tag{15.1}$$

where $\hat{Y}_{1i}$ is a fitted value of the ith auxiliary model.

$$E \text{ can then be restated as } \begin{bmatrix} \sum_{i=1}^{M} \hat{Y}_{1i} \frac{\partial q_{m_i}}{\partial Y_1} & \cdots & \sum_{i=1}^{M} \hat{Y}_{1i} \frac{\partial q_{m_i}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{M} \hat{Y}_{ni} \frac{\partial q_{m_i}}{\partial Y_1} & \cdots & \sum_{i=1}^{M} \hat{Y}_{ni} \frac{\partial q_{m_i}}{\partial Y_n} \end{bmatrix}. \tag{15.2}$$

E can then be restated as $E = X_A \frac{\partial \hat{\beta}_q}{\partial Y}$, (15.3)

where X_A is an n by m matrix of auxiliary model fitted values $\begin{bmatrix} \hat{Y}_{11} & \cdots & \hat{Y}_{1m} \\ \vdots & \ddots & \vdots \\ \hat{Y}_{n1} & \cdots & \hat{Y}_{nm} \end{bmatrix}$, (15.4)

$\hat{\beta}_q$is an m by 1 vector of $q_m's$ $\begin{bmatrix} q_{m_1} \\ \vdots \\ q_{m_M} \end{bmatrix}$, (15.5)

$$EDF = \sum_{i=1}^{M} q_{m_i}\, EDF_{m_i} + Trace\left(\begin{bmatrix} \sum_{i=1}^{M} \hat{Y}_{1i} \frac{\partial q_{m_i}}{\partial Y_1} & \cdots & \sum_{i=1}^{M} \hat{Y}_{1i} \frac{\partial q_{m_i}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{M} \hat{Y}_{ni} \frac{\partial q_{m_i}}{\partial Y_1} & \cdots & \sum_{i=1}^{M} \hat{Y}_{ni} \frac{\partial q_{m_i}}{\partial Y_n} \end{bmatrix}\right), \text{ and} \tag{16}$$

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace\left(X_A \frac{\partial \hat{\beta}_q}{\partial Y}\right). \tag{17}$$

Utilizing the condition that $\hat{\beta}_q$ is estimated by a single equation linear in specification as a function of X_A results in the form $\hat{\beta}_q = F(X_A)Y$, which allows equation (17) to be rewritten as (17.1):

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace\left(X_A \frac{\partial F(X_A)Y}{\partial Y}\right). \tag{17.1}$$

Equation (17.1) is differentiated using equation (10):

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace\left(X_A\left((Y^T \otimes I_m)\frac{\partial F(X_A)}{\partial Y} + (I_1 \otimes F(X_A))I_n\right)\right). \tag{17.2}$$

This simplifies to equation (18):

$$EDF = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + Trace(X_A F(X_A)) + Trace\left(X_A\left((Y^T \otimes I_m)\frac{\partial F(X_A)}{\partial Y}\right)\right). \tag{18}$$

The notation for $(Y^T \otimes I_m)\frac{\partial F(X_A)}{\partial Y}$ is rewritten as $\hat{\beta}_{\Delta q}$, which is an m by n matrix.

$$\hat{\beta}_{\Delta q} = \left(Y^T \otimes I_m\right)\frac{\partial F(X_A)}{\partial Y} = \begin{bmatrix} Y_1 & 0 & 0 & \cdots & Y_n & 0 & 0 \\ 0 & Y_1 & 0 & \cdots & 0 & Y_n & 0 \\ 0 & 0 & Y_1 & \cdots & 0 & 0 & Y_n \end{bmatrix} \begin{bmatrix} \frac{\partial F(X_A)_{11}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{11}}{\partial Y_n} \\ \frac{\partial F(X_A)_{21}}{\partial Y_1} & & \frac{\partial F(X_A)_{21}}{\partial Y_n} \\ \vdots & & \vdots \\ \frac{\partial F(X_A)_{m1}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{m1}}{\partial Y_n} \\ \frac{\partial F(X_A)_{12}}{\partial Y_1} & & \frac{\partial F(X_A)_{12}}{\partial Y_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial F(X_A)_{mn}}{\partial Y_1} & & \frac{\partial F(X_A)_{mn}}{\partial Y_n} \end{bmatrix}. \tag{18.1}$$

$$\hat{\beta}_{\Delta q} = \begin{bmatrix} \sum_{l=1}^{n} \frac{Y_l \partial F(X_A)_{1l}}{\partial Y_1} & \cdots & \sum_{l=1}^{n} \frac{Y_l \partial F(X_A)_{1l}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{l=1}^{n} \frac{Y_l \partial F(X_A)_{ml}}{\partial Y_1} & \cdots & \sum_{l=1}^{n} \frac{Y_l \partial F(X_A)_{ml}}{\partial Y_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial F(X_A)_{11}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{1n}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F(X_A)_{m1}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{mn}}{\partial Y_n} \end{bmatrix} Y. \tag{18.2}$$

$$\hat{\beta}_{\Delta q} = \begin{bmatrix} \frac{\partial F(X_A)_{11}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{1n}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F(X_A)_{m1}}{\partial Y_1} & \cdots & \frac{\partial F(X_A)_{mn}}{\partial Y_n} \end{bmatrix} Y. \tag{18.3}$$

As shown in (18.3), $\hat{\beta}_{\Delta q}$ is the change in $F(X_A)Y$ resulting from X_A being a function of Y.

As a result, the final term in equation (18) can be written as $Trace\left(X_A\hat{\beta}_{\Delta q}\right)$,resulting in Proposition II.

Computing the limit of $\boldsymbol{Trace\left(X_A\hat{\beta}_{\Delta q}\right)}$ $\boldsymbol{as\ \hat{Y}'s \to Y}$:

As $\hat{Y}'s \to Y$, the matrix $X_A = \begin{bmatrix} \hat{Y}_{11} & \cdots & \hat{Y}_{1m} \\ \vdots & \ddots & \vdots \\ \hat{Y}_{n1} & \cdots & \hat{Y}_{nm} \end{bmatrix} \to \begin{bmatrix} Y_1 & \cdots & Y_1 \\ \vdots & \ddots & \vdots \\ Y_n & \cdots & Y_n \end{bmatrix}$. (18.4)

As $X_A \to \begin{bmatrix} Y_1 & \cdots & Y_1 \\ \vdots & \ddots & \vdots \\ Y_n & \cdots & Y_n \end{bmatrix}$, the vector $\hat{\beta}_q = F(X_A)Y \to \phi$, where ϕ is a vector of length m. (18.5)

As $\hat{\beta}_q = F(X_A)Y \to \phi$, the sum $\sum_{i=1}^{m} \hat{\beta}_{q_i} \to 1$. (18.6)

As $\sum_{i=1}^{m} \hat{\beta}_{q_i} \to 1$, any changes in the elements of $\hat{\beta}_q$ are offsetting; therefore,

$$\sum_{i=1}^{m} \frac{\partial \hat{\beta}_{q_i}}{\partial \tau} \to 0 \text{ for an arbitrary } \tau. \tag{18.7}$$

Using the substitution $\hat{\beta}_q = F(X_A)Y$, equation (18.7) can be restated as (18.8):

$$\sum_{i=1}^{m} \frac{\partial \hat{\beta}_{q_i}}{\partial \tau} = \sum_{i=1}^{m} \frac{\partial (F(X_A)Y)_i}{\partial \tau} = \sum_{i=1}^{m} \frac{\partial (\sum_{l=1}^{n} Y_l F(X_A)_{il})_i}{\partial \tau} \to 0 \text{ for an arbitrary } \tau. \tag{18.8}$$

Taking the derivative in equation (18.8) results in (18.9):

$$\sum_{i=1}^{m} \frac{\partial (\sum_{l=1}^{n} Y_l F(X_A)_{il})_i}{\partial \tau} = \sum_{i=1}^{m} \sum_{l=1}^{n} \left(\frac{F(X_A)_{il} \partial Y_l}{\partial \tau} + \frac{Y_l \partial F(X_A)_{il}}{\partial \tau} \right) \to 0 \text{ for an arbitrary } \tau. \tag{18.9}$$

For well behaved cases, $\sum_{i=1}^{m} \sum_{l=1}^{n} \left(\frac{F(X_A)_{il} \partial Y_l}{\partial \tau} \right) \neq -1 \sum_{i=1}^{m} \sum_{l=1}^{n} \left(\frac{Y_l \partial F(X_A)_{il}}{\partial \tau} \right)$, which results

in equations (18.10) and (18.11):

$$\sum_{i=1}^{m}\sum_{l=1}^{n}\left(\frac{F(X_A)_{il}\partial Y_l}{\partial\tau}\right)\rightarrow 0 \text{ for an arbitrary } \tau. \quad (18.10)$$

$$\sum_{i=1}^{m}\sum_{l=1}^{n}\left(\frac{Y_l\partial F(X_A)_{il}}{\partial\tau}\right)\rightarrow 0 \text{ for an arbitrary } \tau. \quad (18.11)$$

Utilizing the (18.2) representation of $\hat{\beta}_{\Delta q} = \begin{bmatrix} \sum_{l=1}^{n}\frac{Y_l\partial F(X_A)_{1l}}{\partial Y_1} & \cdots & \sum_{l=1}^{n}\frac{Y_l\partial F(X_A)_{1l}}{\partial Y_n} \\ \vdots & \ddots & \vdots \\ \sum_{l=1}^{n}\frac{Y_l\partial F(X_A)_{ml}}{\partial Y_1} & \cdots & \sum_{l=1}^{n}\frac{Y_l\partial F(X_A)_{ml}}{\partial Y_n} \end{bmatrix}$

together with equation (18.11) implies that in the matrix $\hat{\beta}_{\Delta q}$, the sum of each column $\rightarrow 0$.

Therefore, as $\sum_{i=1}^{m}\sum_{l=1}^{n}\left(\frac{Y_l\partial F(X_A)_{il}}{\partial\tau}\right)\rightarrow 0$ for an arbitrary τ, $\hat{\beta}_{\Delta q}\rightarrow\lambda$, where λ is an m by n

matrix with columns that sum to zero. (18.12)

Substituting X_A and $\hat{\beta}_{\Delta q}$ with their respective limits from equations (18.4) and (18.12) results in equation (18.13):

$$\lim_{\hat{Y}'s\rightarrow Y} Trace\left(X_A\hat{\beta}_{\Delta q}\right) = Trace\left(\begin{bmatrix} Y_1 & \cdots & Y_1 \\ \vdots & \ddots & \vdots \\ Y_n & \cdots & Y_n \end{bmatrix}\lambda\right). \quad (18.13)$$

Given that the rows of $\begin{bmatrix} Y_1 & \cdots & Y_1 \\ \vdots & \ddots & \vdots \\ Y_n & \cdots & Y_n \end{bmatrix}$ are constant and that the columns of λ sum

to zero, their product is an n by n zero matrix.

$$\lim_{\hat{Y}'s\rightarrow Y} Trace\left(X_A\hat{\beta}_{\Delta q}\right) = Trace\left(\begin{bmatrix} Y_1 & \cdots & Y_1 \\ \vdots & \ddots & \vdots \\ Y_n & \cdots & Y_n \end{bmatrix}\lambda\right) = Trace\begin{bmatrix} 0_{11} & \cdots & 0_{n1} \\ \vdots & \ddots & \vdots \\ 0_{n1} & \cdots & 0_{nn} \end{bmatrix} = 0. \quad (18.14)$$

$$\lim_{\hat{Y}'s\rightarrow Y} Trace\left(X_A\hat{\beta}_{\Delta q}\right) = 0. \quad (18.15)$$

Therefore, the limit of $Trace\left(X_A\hat{\beta}_{\Delta q}\right)$ as $\hat{Y}'s\rightarrow Y$ is zero for well behaved cases.

Approximating $Trace\left(X_A\hat{\beta}_{\Delta q}\right)$ with its limit of zero reduces equation (14) to equation (19):

$$EDF \approx \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + EDF_w + \lim_{\hat{Y}'s\rightarrow Y} Trace\left(X_A\hat{\beta}_{\Delta q}\right) = \sum_{i=1}^{M} q_{m_i} EDF_{m_i} + EDF_w + 0. \quad (18.16)$$

$$EDF \approx \sum_{i=1}^{M} q_{m_i} \, EDF_{m_i} + EDF_w \,. \tag{19}$$

Appendix B: Simulation procedure for section 4

The simulation procedure for section 4 is itemized below.

Step 1: Randomly generate p correlated explanatory variables x_i, where $i = 1, \dots p$, such that each x_i has 100 observations drawn from a mean 0 variance 1 normal distribution, with a correlation of 0.8 across $x_i's$.

Step 2: Randomly generate the dependent variable as $y = \frac{1}{p}x_1 + \cdots + \frac{1}{p}x_p + \varepsilon$, where ε is randomly generated noise from a mean 0 variance 1 normal distribution.

Step 3: Repeat steps 1 and 2 for the four cases $p = 4, 9, 25, 100$.

Step 4: Using observations 1 to 99, estimate a forecast combination model of y for each of the four cases of p, where the auxiliary models are estimated with ordinary least squares of the forms provided below and the weighting scheme is estimated with ordinary least squares under the constraint $\sum_{i=1}^{M} q_{m_i} = 1$:

- for $p = 4$, two auxiliary models of two variables each $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$
- for $p = 9$, three auxiliary models of three variables each $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_3 x_3$
- for $p = 25$, five auxiliary models of five variables each $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_5 x_5$
- for $p = 100$, ten auxiliary models of ten variables each $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_{10} x_{10}$

Step 5: Using Mallows' C_p formula,[1] $C_p = \overline{err} + 2\frac{DF}{n}\hat{\sigma}_\varepsilon^2$, where C_p is the C_p statistic, $\overline{err}$ is the training error, $\hat{\sigma}_\varepsilon^2$ is the noise variance. For each of the four cases, compute two estimates of the out-of-sample mean squared forecast error using $DF = \textit{equation 19}$ and $DF = \textit{count of estimated parameters}$. For both estimates, set $\hat{\sigma}_\varepsilon^2$ equal to its true value of 1.

Step 6: For each of the four cases, use the models from step 4 estimated on observations 1 to 99 and the 100th observation of the $x_i's$, generate out of sample forecasts for the 100th observation of y, and calculate the out-of-sample squared forecast error.

Step 7: Repeat steps 1 to 6 100,000 times and compare the average performance of the two estimates from step 5 with true out-of-sample squared forecast errors resulting from step 6.

[1] Hastie, Tibshirani and Friedman (2001, chapter 7) discuss using this formula to estimate out-of-sample mean squared forecast errors.

Appendix C: EDF computation from Section 5

This appendix computes and decomposed into parts, the EDF value for shrink (0.25) for the US case found in the bottom right cell of Table 1. The three shrink models are specified by equation (20) where $\hat{Y}_{shrink}$ is the model forecast, $\hat{Y}_{ave\ comb}$ and $\hat{Y}_{OLS\ comb}$ are the forecast combination forecasts using simple average weights and linear regression weights respectively, and ω is the weight allocated to the forecast combinations based on sample size and choice of shrink model parameter (1, 0.5, 0.25).

$$\hat{Y}_{shrink} = (1-\omega)\,\hat{Y}_{ave\ comb} + \omega\,\hat{Y}_{OLS comb} \tag{20}$$

As the shrink models are forecast combinations of forecast combinations, the EDF is computed by applying equation (19) in two steps. In step 1, equation (19) is applied to the forecast combination from equation (20), resulting in equation (21). Where $EDF_{W_{shrink}}$ is the EDF of the weighting scheme ω and $EDF_{\hat{Y}_{shrink}}$, $EDF_{\hat{Y}_{ave\ comb}}$, $EDF_{\hat{Y}_{OLS\ comb}}$ are the EDFs of $\hat{Y}_{shrink}$, $\hat{Y}_{ave\ comb}$ and $\hat{Y}_{OLS\ comb}$ respectively. As ω is not estimated but determined by sample size and a fixed parameter, $EDF_{W_{shrink}}$ is zero.

$$EDF_{\hat{Y}_{shrink}} \approx (1-\omega)\,EDF_{\hat{Y}_{ave\ comb}} + \omega\,EDF_{\hat{Y}_{OLS\ comb}} + EDF_{W_{shrink}} \tag{21}$$

In step 2, equation (19) is applied to $EDF_{\hat{Y}_{ave\ comb}}$ and $EDF_{\hat{Y}_{OLS\ comb}}$ resulting in equation (22). Where $q_{ave\ comb_{m_i}}$ and $q_{OLS\ comb_{m_i}}$ are the simple average and linear regression weights, and EDF_{m_i}, $EDF_{W_{ave\ comb}}$ and $EDF_{W_{OLS\ comb}}$ are the EDFs of the auxiliary models, the simple average weighting scheme and linear regression weighting scheme respectively.

$$EDF_{\hat{Y}_{Shrink}} \approx (1-\omega)\left(\sum_{i=1}^{M} q_{ave\ comb_{m_i}} EDF_{m_i} + EDF_{W_{ave\ comb}}\right) + \omega\left(\sum_{i=1}^{M} q_{OLS comb_{m_i}} EDF_{m_i} + EDF_{W_{OLS\ comb}}\right) \tag{22}$$

All of the values in equation (22) are known and are either weights or parameter counts from underlying auxiliary models. To compute the EDF for shrink (0.25) for the US case ($EDF_{\hat{Y}_{Shrink\ 0.25\ (US)}}$), rounded to two decimal places, ω is 0.634, $\sum_{i=1}^{M} q_{ave\ comb_{m_i}} EDF_{m_i}$ is 3.25, $\sum_{i=1}^{M} q_{OLS\ comb_{m_i}} EDF_{m_i}$ is 1.01, $EDF_{W_{ave\ comb}}$ is 0 as the simple average weights have no parameters and $EDF_{W_{OLS\ comb}}$ is 63 from the number of parameters in the linear regression weights.

$$EDF_{\hat{Y}_{Shrink\ 0.25\ (US)}} \approx (1-0.634)(3.25+0) + 0.634(1.01+63) = 41.76 \tag{23}$$

Although this is a complicated forecast combination we can see from equation (23) that in this case almost all of the complexity cost is from the 63 parameters used to compute the linear regression weighting scheme and very little is from the underlying auxiliary models.

Appendix D: Number of ways to group variables into forecast combinations

To compute the number of ways v variables can be grouped into forecast combinations, $\binom{v}{i}$ provides the number of possible groupings of v variables into auxiliary models of i variables, where $\binom{v}{i}$ is the combination operator. Equation (24) computes the total number of possible auxiliary model variable groupings g, where auxiliary models range in size from 1 to v variables.

$$g = \sum_{i=1}^{v} \binom{v}{i}. \tag{24}$$

Then the number of ways the auxiliary model groupings can be arranged into forecast combinations c_0 is arrived at by computing all possible subsets of g by putting this number to a base of 2, provided that forecast combinations of forecast combinations are not included.

$$c_0 = 2^g. \tag{25}$$

To allow for a single generation of forecast combinations of forecast combinations, the initial set is increased from g elements to $g + c_0$ elements, and the total of all possible subsets is c_1.

$$c_1 = 2^{g+c_0}. \tag{26}$$