Internet Appendix for

"Cross-Sectional Expected Returns: New Fama-MacBeth Regressions in the Era of Machine Learning"

Not for Publication

April 28, 2023

A1. Forecast Construction

This section details the construction of the cross-sectional out-of-sample return forecasts in the Fama-MacBeth framework described in Section 2 of the paper.

A1.1. Smoothed WLS Forecast

Step 1 Estimate a month-t cross-sectional multiple regression via WLS:

$$r_{i,t} = a_t + \sum_{j=1}^{J_{t-1}} b_{j,t} z_{i,j,t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t. \tag{A.1}$$

Step 2 Compute the month-$(t+1)$ smoothed WLS forecast:

$$\hat{r}^{\text{S-WLS}}_{i,t+1|t} = \bar{a}^{\text{WLS}}_t + \sum_{j=1}^{J_t} \bar{b}^{\text{WLS}}_{j,t} z_{i,j,t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.2}$$

where

$$\bar{a}_t^{\text{WLS}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{a}_{t-m}^{\text{WLS}}, \tag{A.3}$$

$$\bar{b}_{j,t}^{\text{WLS}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{b}_{j,t-m}^{\text{WLS}} \quad \text{for } j = 1, \ldots, J_t, \tag{A.4}$$

M is the size of the window for smoothing the coefficient estimates over time, and $\hat{a}_t^{\text{WLS}}$ and $\hat{b}_{j,t}^{\text{WLS}}$ are the WLS estimates of a_t and $b_{j,t}$, respectively, in Equation (A.1). For a characteristic to be included in Equation (A.2), it needs to be available in months $t-1$ and t.

A1.2. Smoothed LASSO Forecast

Step 1 Estimate a month-t cross-sectional multiple regression via the weighted LASSO:

$$r_{i,t} = a_t + \sum_{j=1}^{J_{t-1}} b_{j,t} z_{i,j,t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t. \tag{A.5}$$

Step 2 Compute the month-$(t+1)$ smoothed LASSO forecast:

$$\hat{r}_{i,t+1|t}^{\text{S-LASSO}} = \bar{a}_t^{\text{LASSO}} + \sum_{j=1}^{J_t} \bar{b}_{j,t}^{\text{LASSO}} z_{i,j,t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.6}$$

where

$$\bar{a}_t^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{a}_{t-m}^{\text{LASSO}}, \tag{A.7}$$

$$\bar{b}_{j,t}^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{b}_{j,t-m}^{\text{LASSO}} \quad \text{for } j = 1, \ldots, J_t, \tag{A.8}$$

M is the size of the window for smoothing the coefficient estimates over time, and $\hat{a}_t^{\text{LASSO}}$ and $\hat{b}_{j,t}^{\text{LASSO}}$ are the weighted LASSO estimates of a_t and $b_{j,t}$, respectively, in

Equation (A.5). For a characteristic to be included in Equation (A.6), it needs to be available in months $t-1$ and t.

A1.3. Combination LASSO Forecast

Step 1 For each characteristic, estimate a month-$(t-1)$ cross-sectional univariate regression via WLS:

$$r_{i,t-1} = c_{j,t-1} + d_{j,t-1} z_{i,j,t-2} + \varepsilon_{i,t-1} \quad \text{for } 1, \ldots, n_{t-1};\ j = 1, \ldots, J_{t-2}. \tag{A.9}$$

Step 2 For each characteristic, compute a month-t univariate regression forecast:

$$\hat{r}^{(j)}_{i,t|t-1} = \hat{c}^{\text{WLS}}_{j,t-1} + \hat{d}^{\text{WLS}}_{j,t-1} z_{i,j,t-1} \quad \text{for } 1, \ldots, n_t;\ j = 1, \ldots, J_{t-1}, \tag{A.10}$$

where $\hat{c}^{\text{WLS}}_{j,t-1}$ and $\hat{d}^{\text{WLS}}_{j,t-1}$ are the WLS estimates of $c_{j,t-1}$ and $d_{j,t-1}$, respectively, in Equation (A.9). For a characteristic to be included in Equation (A.10), it needs to be available in months $t-2$ and $t-1$.

Step 3 Estimate a month-t cross-sectional Granger and Ramanathan (1984) multiple regression via the weighted LASSO:

$$r_{i,t} = \xi_t + \sum_{j=1}^{J_{t-1}} \phi_{j,t} \hat{r}^{(j)}_{i,t|t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t, \tag{A.11}$$

where we impose the restrictions that $\phi_{j,t} \geq 0$ for $j = 1, \ldots, J_{t-1}$ in Equation (A.11). Let $\hat{\mathcal{J}}_t \subseteq \{1, \ldots, J_{t-1}\}$ denote the index set of univariate forecasts selected by the LASSO in Equation (A.11).

Step 4 For each characteristic, estimate a month-t cross-sectional univariate regression via WLS:

$$r_{i,t} = c_{j,t} + d_{j,t} z_{i,j,t-1} + \varepsilon_{i,t} \quad \text{for } 1, \ldots, n_t;\ j = 1, \ldots, J_{t-1}. \tag{A.12}$$

Step 5 For each characteristic, compute a month-$(t+1)$ cross-sectional univariate regression forecast:

$$\hat{r}^{(j)}_{i,t+1|t} = \hat{c}^{\text{WLS}}_{j,t} + \hat{d}^{\text{WLS}}_{j,t} z_{i,j,t} \quad \text{for } i = 1, \ldots, n_{t+1};\ j = 1, \ldots, J_t, \tag{A.13}$$

where $\hat{c}^{\text{WLS}}_{j,t}$ and $\hat{d}^{\text{WLS}}_{j,t}$ are the WLS estimates of $c_{j,t}$ and $d_{j,t}$, respectively, in Equation (A.12). For a characteristic to be included in Equation (A.13), it needs to be available in months $t-1$ and t.

Step 6 Compute the month-$(t+1)$ combination LASSO forecast:

$$\hat{r}^{\text{C-LASSO}}_{i,t+1|t} = \frac{1}{\left|\hat{\mathcal{J}}_t\right|} \sum_{j \in \hat{\mathcal{J}}_t} \hat{r}^{(j)}_{i,t+1|t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.14}$$

where $|\hat{\mathcal{J}}_t|$ is the cardinality of $\hat{\mathcal{J}}_t$.

A1.4. Encompassing LASSO Forecast

Step 1 For each characteristic, estimate a month-$(t-2)$ cross-sectional univariate regression via WLS:

$$r_{i,t-2} = c_{j,t-2} + d_{j,t-2} z_{i,j,t-3} + \varepsilon_{i,t-2} \quad \text{for } 1, \ldots, n_{t-2};\ j = 1, \ldots, J_{t-3}. \tag{A.15}$$

Step 2 For each characteristic, compute a month-$(t-1)$ univariate regression forecast:

$$\hat{r}^{(j)}_{i,t-1|t-2} = \hat{c}^{\text{WLS}}_{j,t-2} + \hat{d}^{\text{WLS}}_{j,t-2} z_{i,j,t-2} \quad \text{for } 1, \ldots, n_{t-1};\ j = 1, \ldots, J_{t-2}, \tag{A.16}$$

where $\hat{c}_{j,t-2}^{\text{WLS}}$ and $\hat{d}_{j,t-2}^{\text{WLS}}$ are the WLS estimates of $c_{j,t-2}$ and $d_{j,t-2}$, respectively, in Equation (A.15). For a characteristic to be included in Equation (A.16), it needs to be available in months $t-3$ and $t-2$.

Step 3 Estimate a month-$(t-1)$ cross-sectional Granger and Ramanathan (1984) multiple regression via the weighted LASSO:

$$r_{i,t-1} = \xi_{t-1} + \sum_{j=1}^{J_{t-2}} \phi_{j,t-1} \hat{r}_{i,t-1|t-2}^{(j)} + \varepsilon_{i,t-1} \quad \text{for } i = 1, \ldots, n_{t-1}, \tag{A.17}$$

where we impose the restrictions that $\phi_{j,t-1} \geq 0$ for $j = 1, \ldots, J_{t-2}$ in Equation (A.26). Let $\hat{\mathcal{J}}_{t-1} \subseteq \{1, \ldots, J_{t-2}\}$ denote the index set of univariate forecasts selected by the LASSO in Equation (A.17).

Step 4 For each characteristic, estimate a month-$(t-1)$ cross-sectional univariate regression via WLS:

$$r_{i,t-1} = c_{j,t-1} + d_{j,t-1} z_{i,j,t-2} + \varepsilon_{i,t-1} \quad \text{for } 1, \ldots, n_{t-1};\ j = 1, \ldots, J_{t-2}. \tag{A.18}$$

Step 5 For each characteristic, compute a month-t univariate regression forecast:

$$\hat{r}_{i,t|t-1}^{(j)} = \hat{c}_{j,t-1}^{\text{WLS}} + \hat{d}_{j,t-1}^{\text{WLS}} z_{i,j,t-1} \quad \text{for } 1, \ldots, n_t;\ j = 1, \ldots, J_{t-1}, \tag{A.19}$$

where $\hat{c}_{j,t-1}^{\text{WLS}}$ and $\hat{d}_{j,t-1}^{\text{WLS}}$ are the WLS estimates of $c_{j,t-1}$ and $d_{j,t-1}$, respectively, in Equation (A.18). For a characteristic to be included in Equation (A.19), it needs to be available in months $t-2$ and $t-1$.

Step 6 Compute the month-t combination LASSO forecast:

$$\hat{r}_{i,t|t-1}^{\text{C-LASSO}} = \frac{1}{\left|\hat{\mathcal{J}}_{t-1}\right|} \sum_{j \in \hat{\mathcal{J}}_{t-1}} \hat{r}_{i,t|t-1}^{(j)} \quad \text{for } i = 1, \ldots, n_t, \tag{A.20}$$

where $|\hat{\mathcal{J}}_{t-1}|$ is the cardinality of $\hat{\mathcal{J}}_{t-1}$.

Step 7 Estimate a month-$(t-1)$ cross-sectional multiple regression via the weighted LASSO:

$$r_{i,t-1} = a_{t-1} + \sum_{j=1}^{J_{t-2}} b_{j,t-1} z_{i,j,t-2} + \varepsilon_{i,t-1} \quad \text{for } i = 1, \ldots, n_{t-1}. \tag{A.21}$$

Step 8 Compute the month-t smoothed LASSO forecast:

$$\hat{r}_{i,t|t-1}^{\text{S-LASSO}} = \bar{a}_{t-1}^{\text{LASSO}} + \sum_{j=1}^{J_{t-1}} \bar{b}_{j,t-1}^{\text{LASSO}} z_{i,j,t-1} \quad \text{for } i = 1, \ldots, n_t, \tag{A.22}$$

where

$$\bar{a}_{j,t-1}^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{a}_{t-1-m}^{\text{LASSO}}, \tag{A.23}$$

$$\bar{b}_{j,t-1}^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{b}_{j,t-1-m}^{\text{LASSO}} \quad \text{for } j = 1, \ldots, J_{t-1}, \tag{A.24}$$

M is the size of the window for smoothing the coefficient estimates over time, and $\hat{a}_{t-1}^{\text{LASSO}}$ and $\hat{b}_{j,t-1}^{\text{LASSO}}$ are the weighted LASSO estimates of a_{t-1} and $b_{j,t-1}$, respectively, in Equation (A.21). For a characteristic to be included in Equation (A.22), it needs to be available in months $t-2$ and $t-1$.

Step 9 Estimate a month-t cross-sectional encompassing regression via WLS:

$$\hat{e}_{i,t|t-1}^{\text{S-LASSO}} = \eta_t + \theta_t \left(\hat{e}_{i,t|t-1}^{\text{S-LASSO}} - \hat{e}_{i,t|t-1}^{\text{C-LASSO}} \right) + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t, \tag{A.25}$$

where $\hat{e}_{i,t|t-1}^{\text{S-LASSO}} = r_{i,t} - \hat{r}_{i,t|t-1}^{\text{S-LASSO}}$ and $\hat{e}_{i,t|t-1}^{\text{C-LASSO}} = r_{i,t} - \hat{r}_{i,t|t-1}^{\text{C-LASSO}}$.

Step 10 Estimate a month-t cross-sectional Granger and Ramanathan (1984) multiple regression via the weighted LASSO:

$$r_{i,t} = \xi_t + \sum_{j=1}^{J_{t-1}} \phi_{j,t} \hat{r}^{(j)}_{i,t|t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t, \tag{A.26}$$

where we impose the restrictions that $\phi_{j,t} \geq 0$ for $j = 1, \ldots, J_{t-1}$ in Equation (A.26). Let $\hat{\mathcal{J}}_t \subseteq \{1, \ldots, J_{t-1}\}$ denote the index set of univariate forecasts selected by the LASSO in Equation (A.26).

Step 11 For each characteristic, estimate a month-t cross-sectional univariate regression via WLS:

$$r_{i,t} = c_{j,t} + d_{j,t} z_{i,j,t-1} + \varepsilon_{i,t} \quad \text{for } 1, \ldots, n_t;\ j = 1, \ldots, J_{t-1}. \tag{A.27}$$

Step 12 For each characteristic, compute a month-$(t+1)$ univariate regression forecast:

$$\hat{r}^{(j)}_{i,t+1|t} = \hat{c}^{\text{WLS}}_{j,t} + \hat{d}^{\text{WLS}}_{j,t} z_{i,j,t} \quad \text{for } 1, \ldots, n_{t+1};\ j = 1, \ldots, J_t, \tag{A.28}$$

where $\hat{c}^{\text{WLS}}_{j,t}$ and $\hat{d}^{\text{WLS}}_{j,t}$ are the WLS estimates of $c_{j,t}$ and $d_{j,t}$, respectively, in Equation (A.27). For a characteristic to be included in Equation (A.28), it needs to be available in months $t-1$ and t.

Step 13 Compute the month-$(t+1)$ combination LASSO forecast:

$$\hat{r}^{\text{C-LASSO}}_{i,t+1|t} = \frac{1}{\left|\hat{\mathcal{J}}_t\right|} \sum_{j \in \hat{\mathcal{J}}_t} \hat{r}^{(j)}_{i,t+1|t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.29}$$

where $|\hat{\mathcal{J}}_t|$ is the cardinality of $\hat{\mathcal{J}}_t$.

Step 14 Estimate a month-t cross-sectional multiple regression via the weighted LASSO:

$$r_{i,t} = a_t + \sum_{j=1}^{J_{t-1}} b_{j,t} z_{i,j,t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t. \tag{A.30}$$

Step 15 Compute the month-$(t+1)$ smoothed LASSO forecast:

$$\hat{r}_{i,t+1|t}^{\text{S-LASSO}} = \bar{a}_t^{\text{LASSO}} + \sum_{j=1}^{J_t} \bar{b}_{j,t}^{\text{LASSO}} z_{i,j,t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.31}$$

where

$$\bar{a}_{j,t}^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{a}_{t-m}^{\text{LASSO}}, \tag{A.32}$$

$$\bar{b}_{j,t}^{\text{LASSO}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{b}_{j,t-m}^{\text{LASSO}} \quad \text{for } j = 1, \ldots, J_t, \tag{A.33}$$

M is the size of the window for smoothing the coefficient estimates over time, and $\hat{a}_t^{\text{LASSO}}$ and $\hat{b}_{j,t}^{\text{LASSO}}$ are the weighted LASSO estimates of a_t and $b_{j,t}$, respectively, in Equation (A.30). For a characteristic to be included in Equation (A.31), it needs to be available in months $t-1$ and t.

Step 16 Compute the month-$(t+1)$ encompassing LASSO forecast:

$$\bar{r}_{i,t+1|t}^{\text{E-LASSO}} = \left(1 - \bar{\theta}_t^{\text{WLS}}\right) \hat{r}_{i,t+1|t}^{\text{S-LASSO}} + \bar{\theta}_t^{\text{WLS}} \hat{r}_{i,t+1|t}^{\text{C-LASSO}} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.34}$$

where

$$\bar{\theta}_t^{\text{WLS}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{\theta}_{t-m}^{\text{WLS}}, \tag{A.35}$$

M is the size of the window for smoothing the coefficient estimates over time, $\hat{\theta}_t^{\text{WLS}}$ is the WLS estimate of θ_t in Equation (A.25), and we set $\hat{\theta}_t^{\text{WLS}}$ to zero (one) if it is less (greater) than zero (one).

A1.5. Smoothed Random Features Forecast

Step 1 Create the $S \times J$ weight matrix $\boldsymbol{\Omega}$, where each element $\omega_{s,j}$ is drawn independently from $\mathcal{N}(0, 1/S)$; also create the S-dimensional vector $\boldsymbol{\tau}$, where each element τ_s is drawn independently from $\mathcal{N}(0, 1/S)$.

Step 2 Compute the random features for month $t-1$:

$$g_{i,s,t-1} = \chi\left(\tau_s + \sum_{j=1}^{J_{t-1}} \omega_{s,j} z_{i,j,t-1}\right) \quad \text{for } i = 1, \ldots, n_t;\ s = 1, \ldots, S, \tag{A.36}$$

where $\chi(x) = \max\{x, 0\}$.

Step 3 Estimate a month-t cross-sectional multiple regression via weighted ridge regression:

$$r_{i,t} = \kappa_t + \sum_{s=1}^{S} \psi_{s,t} g_{i,s,t-1} + \varepsilon_{i,t} \quad \text{for } i = 1, \ldots, n_t, \tag{A.37}$$

where the regularization hyperparameter is selected via ten-fold cross validation.

Step 4 Compute the random features for month t:

$$g_{i,s,t} = \chi\left(\tau_s + \sum_{j=1}^{J_t} \omega_{s,j} z_{i,j,t}\right) \quad \text{for } i = 1, \ldots, n_{t+1};\ s = 1, \ldots, S. \tag{A.38}$$

Step 5 Compute the month-$(t+1)$ smoothed random features forecast:

$$\hat{r}_{i,t+1|t}^{\text{S-RanFeat}} = \bar{\kappa}_t^{\text{Ridge}} + \sum_{s=1}^{S} \bar{\psi}_{s,t}^{\text{Ridge}} g_{i,s,t} \quad \text{for } i = 1, \ldots, n_{t+1}, \tag{A.39}$$

where

$$\bar{\kappa}_t^{\text{Ridge}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{\kappa}_{t-m}^{\text{Ridge}}, \tag{A.40}$$

$$\bar{\psi}_{s,t}^{\text{Ridge}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{\psi}_{s,t-m}^{\text{Ridge}} \quad \text{for } s = 1, \ldots, S, \tag{A.41}$$

M is the size of the window for smoothing the coefficient estimates over time, and $\hat{\kappa}_t^{\text{Ridge}}$ and $\hat{\psi}_{s,t}^{\text{Ridge}}$ are the weighted ridge estimates of κ_t and $\psi_{s,t}$, respectively, in Equation (A.37).

A2. Forecast Evaluation

This section derives Equations (37) and (41) in Sections 3.2 and 3.3, respectively, of the paper. To derive Equation (37), first note that the value-weighted total sum of squares for Equation (33) in the paper is given by

$$\text{TSS}_t = \sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)^2. \tag{A.42}$$

The value-weighted residual sum of squares can be expressed as

$$\begin{aligned} \text{RSS}_t &= \sum_{i=1}^{n_t} w_{i,t}\Big[(r_{i,t} - \bar{r}_t) - \hat{\delta}_t^{\text{WLS}}\big(\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1}\big)\Big]^2 \\ &= \sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)^2 - 2\hat{\delta}_t^{\text{WLS}} \sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)\big(\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1}\big) \\ &\quad + \Big(\hat{\delta}_t^{\text{WLS}}\Big)^2 \sum_{i=1}^{n_t} w_{i,t}\big(\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1}\big)^2. \end{aligned} \tag{A.43}$$

Using Equations (30) and (34) in the paper and Equations (A.42) and (A.43), we can write the value-weighted R^2 statistic for Equation (33) in the paper as

$$
\begin{aligned}
R^2_{\text{CSMZ},t} &= 1 - \frac{\text{RSS}_t}{\text{TSS}_t} \\
&= 2\hat{\delta}_t^{\text{WLS}} \frac{\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)\left(\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1}\right)}{\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)^2} \\
&\quad - \left(\hat{\delta}_t^{\text{WLS}}\right)^2 \frac{\frac{1}{n_t}\sum_{i=1}^{n_t}\left(\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1}\right)^2}{\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(r_{i,t} - \bar{r}_t)^2} \\
&= 2\hat{\delta}_t^{\text{WLS}}\left(\hat{\delta}_t^{\text{WLS}}\frac{\hat{\sigma}^2_{\hat{r},t}}{\hat{\sigma}^2_{r,t}}\right) - \left(\hat{\delta}_t^{\text{WLS}}\right)^2\frac{\hat{\sigma}^2_{\hat{r},t}}{\hat{\sigma}^2_{r,t}} \\
&= \left(\hat{\delta}_t^{\text{WLS}}\right)^2\frac{\hat{\sigma}^2_{\hat{r},t}}{\hat{\sigma}^2_{r,t}} \quad \text{for } t = 1, \ldots, T,
\end{aligned}
\tag{A.44}
$$

where $\hat{\sigma}^2_{\hat{r},t}$ is the value-weighted cross-sectional forecast variance in Equation (38) in the paper.

Observe that Equation (A.44) implies that

$$
\left(\hat{\delta}_t^{\text{WLS}}\right)^2 = R^2_{\text{CSMZ},t}\frac{\hat{\sigma}^2_{r,t}}{\hat{\sigma}^2_{\hat{r},t}}. \tag{A.45}
$$

Furthermore, we can write the squared forecast bias as

$$
\underbrace{\left(\hat{\delta}_t^{\text{WLS}} - 1\right)^2}_{\text{bias}_t} = \left(\hat{\delta}_t^{\text{WLS}}\right)^2 - 2\hat{\delta}_t^{\text{WLS}} + 1 \;\Rightarrow\; 2\hat{\delta}_t^{\text{WLS}} - 1 = \left(\hat{\delta}_t^{\text{WLS}}\right)^2 - \left(\hat{\delta}_t^{\text{WLS}} - 1\right)^2. \tag{A.46}
$$

Using Equation (34) in the paper and Equations (A.45) and (A.46), we can write the value-weighted cross-sectional MSFE in Equation (25) in the paper as

$$
\begin{aligned}
\text{MSFE}_t &= \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(r_{i,t}-\bar{r}_t)^2 - 2\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(r_{i,t}-\bar{r}_t)\big(\hat{r}_{i,t|t-1}-\bar{\hat{r}}_{t|t-1}\big) \\
&\quad + \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{r}_{i,t|t-1}-\bar{\hat{r}}_{t|t-1}\big)^2 \\
&= \hat{\sigma}_{r,t} - 2\hat{\delta}_t^{\text{WLS}}\hat{\sigma}_{\hat{r},t} + \hat{\sigma}_{\hat{r},t} \\
&= \hat{\sigma}_{r,t} - \hat{\sigma}_{\hat{r},t}\Big(2\hat{\delta}_t^{\text{WLS}} - 1\Big) \qquad\qquad (\text{A.47})\\
&= \hat{\sigma}_{r,t} - \hat{\sigma}_{\hat{r},t}\bigg[\Big(\hat{\delta}_t^{\text{WLS}}\Big)^2 - \Big(\hat{\delta}_t^{\text{WLS}} - 1\Big)^2\bigg] \\
&= \Big(\hat{\delta}_t^{\text{WLS}} - 1\Big)^2\hat{\sigma}_{\hat{r},t} + \hat{\sigma}_{r,t} - \hat{\sigma}_{\hat{r},t}\left(R^2_{\text{CSMZ}}\frac{\hat{\sigma}^2_{r,t}}{\hat{\sigma}^2_{\hat{r},t}}\right) \\
&= \Big(\hat{\delta}_t^{\text{WLS}} - 1\Big)^2\hat{\sigma}_{\hat{r},t} + \big(1 - R^2_{\text{CSMZ}}\big)\hat{\sigma}^2_{r,t}.
\end{aligned}
$$

which corresponds to Equation (37) in the paper.

To derive Equation (41) in the paper, we first define

$$
\hat{u}^*_{i,t|t-1} = (r_{i,t}-\bar{r}_t) - \big(\hat{r}^*_{i,t|t-1} - \bar{\hat{r}}^*_{t|t-1}\big), \qquad (\text{A.48})
$$

so we can express Equation (40) in the paper as

$$
\text{MSFE}^*_t = \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{u}^*_{i,t|t-1}\big)^2 \quad \text{for } t = 1,\ldots,T. \qquad (\text{A.49})
$$

We also define

$$
\hat{u}^k_{i,t|t-1} = \big(r_{i,t} - \hat{r}^k_{i,t|t-1}\big) - \big(\bar{r}_t - \bar{\hat{r}}^k_{t|t-1}\big) \quad \text{for } k = \text{A}, \text{B}, \qquad (\text{A.50})
$$

where

$$\bar{\hat{r}}^{k}_{t|t-1} = \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\hat{r}^{k}_{i,t|t-1} \quad \text{for } k = \text{A}, \text{B}. \tag{A.51}$$

Using Equation (39) in the paper and Equation (A.50), we can write Equation (A.48) as

$$\begin{aligned}
\hat{u}^{*}_{i,t|t-1} &= (r_{i,t} - \bar{r}_t) - \big(\hat{r}^{*}_{i,t|t-1} - \bar{\hat{r}}^{*}_{t|t-1}\big) \\
&= (r_{i,t} - \bar{r}_t) - \big\{\big[(1-\zeta_t)\hat{r}^{\text{A}}_{i,t|t-1} + \zeta_t\hat{r}^{\text{B}}_{i,t|t-1}\big] - \big[(1-\zeta_t)\bar{\hat{r}}^{\text{A}}_{i,t|t-1} + \zeta_t\bar{\hat{r}}^{\text{B}}_{i,t|t-1}\big]\big\} \\
&= (r_{i,t} - \bar{r}_t) - \big\{\hat{r}^{\text{A}}_{i,t|t-1} + \zeta_t\big(\hat{r}^{\text{B}}_{i,t|t-1} - \hat{r}^{\text{A}}_{i,t|t-1}\big) - \big[\bar{\hat{r}}^{\text{A}}_{t|t-1} + \zeta_t\big(\bar{\hat{r}}^{\text{B}}_{t|t-1} - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big)\big]\big\} \\
&= (r_{i,t} - \bar{r}_t) - \big\{\big(\hat{r}^{\text{A}}_{i,t|t-1} - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big) + \zeta_t\big[\big(\hat{r}^{\text{B}}_{i,t|t-1} - \bar{\hat{r}}^{\text{B}}_{t|t-1}\big) - \big(\hat{r}^{\text{A}}_{i,t|t-1} - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big)\big]\big\} \\
&= \big(r_{i,t} - \hat{r}^{\text{A}}_{i,t|t-1}\big) - \big(\bar{r}_t - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big) - \zeta_t\big[\big(\hat{r}^{\text{B}}_{i,t|t-1} - \bar{\hat{r}}^{\text{B}}_{t|t-1}\big) - \big(\hat{r}^{\text{A}}_{i,t|t-1} - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big)\big] \\
&\quad + \zeta_t[(r_{i,t} - \bar{r}_t) - (r_{i,t} - \bar{r}_t)] \\
&= \underbrace{\big(r_{i,t} - \hat{r}^{\text{A}}_{i,t|t-1}\big) - \big(\bar{r}_t - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big)}_{\hat{u}^{\text{A}}_{i,t|t-1}} + \zeta_t\big\{\underbrace{\big(r_{i,t} - \hat{r}^{\text{B}}_{i,t|t-1}\big) - \big(\bar{r}_t - \bar{\hat{r}}^{\text{B}}_{t|t-1}\big)}_{\hat{u}^{\text{B}}_{i,t|t-1}} \\
&\quad - \underbrace{\big[\big(r_{i,t} - \hat{r}^{\text{A}}_{i,t|t-1}\big) - \big(\bar{r}_t - \bar{\hat{r}}^{\text{A}}_{t|t-1}\big)\big]}_{\hat{u}^{\text{A}}_{i,t|t-1}}\big\} \\
&= \hat{u}^{\text{A}}_{i,t|t-1} - \zeta_t\big(\hat{u}^{\text{A}}_{i,t|t-1} - \hat{u}^{\text{B}}_{i,t|t-1}\big).
\end{aligned} \tag{A.52}$$

We can then write the value-weighted cross-sectional MSFE in Equation (A.49) as

$$\begin{aligned}
\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{u}^{*}_{i,t|t-1}\big)^2 &= \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big[\hat{u}^{\text{A}}_{i,t|t-1} - \zeta_t\big(\hat{u}^{\text{A}}_{i,t|t-1} - \hat{u}^{\text{B}}_{i,t|t-1}\big)\big]^2 \\
&= \frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{u}^{\text{A}}_{i,t|t-1}\big)^2 - 2\zeta_t\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\hat{u}^{\text{A}}_{i,t|t-1}\big(\hat{u}^{\text{A}}_{i,t|t-1} - \hat{u}^{\text{B}}_{i,t|t-1}\big) \\
&\quad + \zeta_t^2\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{u}^{\text{A}}_{i,t|t-1} - \hat{u}^{\text{B}}_{i,t|t-1}\big)^2.
\end{aligned} \tag{A.53}$$

Taking the derivative with respect to ζ_t, we have

$$\frac{d}{d\,\zeta_t}\left[\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}(\hat{u}^{*}_{i,t|t-1})^2\right] = 2\zeta_t\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\big(\hat{u}^{\mathrm{A}}_{i,t|t-1} - \hat{u}^{\mathrm{B}}_{i,t|t-1}\big)^2 - 2\frac{1}{n_t}\sum_{i=1}^{n_t} w_{i,t}\hat{u}^{\mathrm{A}}_{i,t|t-1}\big(\hat{u}^{\mathrm{A}}_{i,t|t-1} - \hat{u}^{\mathrm{B}}_{i,t|t-1}\big). \tag{A.54}$$

Setting the derivative to zero and solving for ζ_t yields

$$\zeta_t^* = \frac{\sum_{i=1}^{n_t} w_{i,t}\hat{u}^{\mathrm{A}}_{i,t|t-1}\Big(\hat{u}^{\mathrm{A}}_{i,t|t-1} - \hat{u}^{\mathrm{B}}_{i,t|t-1}\Big)}{\sum_{i=1}^{n_t} w_{i,t}\Big(\hat{u}^{\mathrm{A}}_{i,t|t-1} - \hat{u}^{\mathrm{B}}_{i,t|t-1}\Big)^2}. \tag{A.55}$$

Inspection of ζ_t^* reveals that it is equivalent to the WLS slope coefficient estimate in a regression of the demeaned forecast error for A on the difference between the demeaned forecast errors for A and B. By the Frisch-Waugh-Lovell theorem, ζ_t^* is thus identical to the WLS estimate of θ_t in Equation (41) in the paper.

References

Chen, A. Y. and T. Zimmermann (2022). Open Source Cross-Sectional Asset Pricing. *Critical Finance Review* 27:2, 207–264.

Granger, C. W. J. and R. Ramanathan (1984). Improved Methods of Combining Forecasts. *Journal of Forecasting* 3:2, 197–204.

Hou, K., C. Xue, and L. Zhang (2020). Replicating Anomalies. *Review of Financial Studies* 33:5, 2019–2133.

Table A1. Characteristics

The table provides abbreviations and names for the 207 firm characteristics from Chen and Zimmermann (2022) used in the empirical application in the paper. The characteristics are grouped according to six economic categories from Hou, Xue, and Zhang (2020). Detailed descriptions of the characteristics and data are available from Open Source Asset Pricing.

(1) Abbreviation	(2) Name	(3) Abbreviation	(4) Name
Panel A: Momentum (36)			
AnnouncementReturn	Earnings announcement return	MomOffSeason11YrPlus	Off season reversal years 11 to 15
CustomerMomentum	Customer momentum	MomOffSeason16YrPlus	Off season reversal years 16 to 20
EarnSupBig	Earnings surprise of big firms	MomRev	Momentum and long-term reversal
EarningsStreak	Earnings surprise streak	MomSeason	Return seasonality years 2 to 5
EarningsSurprise	Earnings surprise	MomSeason06YrPlus	Return seasonality years 6 to 10
FirmAgeMom	Firm age—momentum	MomSeason11YrPlus	Return seasonality years 11 to 15
High52	52-week high	MomSeason16YrPlus	Return seasonality years 16 to 20
IndMom	Industry momentum	MomSeasonShort	Return seasonality last year
IndRetBig	Industry return of big firms	MomVol	Momentum in high-volume stocks
IntMom	Intermediate momentum	NumEarnIncrease	Earnings streak length
LReversal	Long-term reversal	REV6	Earnings forecast revisions
MReversal	Momentum reversal	ResidualMomentum	Momentum based on FF3 residuals
Mom12m	Momentum (12 months)	RevenueSurprise	Revenue surprise
Mom12mOffSeason	Momentum without the seasonal part	ShortTermReversal	Short-term reversal
Mom6m	Momentum (6 months)	iomom_cust	Customers momentum
Mom6mJunk	Junk stock momentum	iomom_supp	Suppliers momentum
MomOffSeason	Off season long-term reversal	retConglomerate	Conglomerate return
MomOffSeason06YrPlus	Off season reversal years 6 to 10	TrendFactor	Trend factor
Panel B: Value vs. growth (28)			
AM	Total assets to market	EquityDuration	Equity duration
BM	Book to market using most recent ME	FEPS	Analyst earnings per share
BMdec	Book to market using December ME	IntanBM	Intangible return using BM
BPEBM	Leverage component of BM	IntanCFP	Intangible return using CFtoP
BookLeverage	Book leverage (annual)	IntanEP	Intangible return using EP
CF	Cash flow to market	IntanSP	Intangible return using Sale2P
Cash	Cash to assets	logSize	Log size
DivInit	Dividend initiation	MeanRankRevGrowth	Revenue growth rank
DivOmit	Dividend omission	NetDebtPrice	Net debt to price
DivSeason	Dividend seasonality	NetPayoutYield	Net payout yield
DivYieldST	Predicted dividend yield next month	PayoutYield	Payout yield
EBM	Enterprise component of BM	SP	Sales to price
EP	Earnings-to-price ratio	cfp	Operating cash flows to price
EntMult	Enterprise multiple	sfe	Earnings forecast to price

Table A1 (continued)

(1) Abbreviation	(2) Name	(3) Abbreviation	(4) Name
Panel C: Investment (34)			
AbnormalAccruals	Abnormal accruals	GrLTNOA	Growth in long-term op assets
Accruals	Accruals	InvGrowth	Inventory growth
AccrualsBM	Book to market and accruals	InvestPPEInv	Change in PPE and inv/assets
AssetGrowth	Asset growth	Investment	Investment to revenue
BrandInvest	Brand capital investment	NOA	Net operating assets
ChInv	Change in capital investment	NetDebtFinance	Net debt financing
ChInvIA	Change in capital inv (ind adj)	NetEquityFinance	Net equity financing
ChNNCOA	Change in net noncurrent op assets	PctAcc	Percent operating accruals
CompEquIss	Composite equity issuance	PctTotAcc	Percent total accruals
CompositeDebtIssuance	Composite debt issuance	ShareIss1Y	Share issuance (1 year)
DebtIssuance	Debt issuance	ShareIss5Y	Share issuance (5 years)
DelCOA	Change in current op assets	ShareRepurchase	Share repurchases
DelCOL	Change in current op liabilities	TotalAccruals	Total accruals
DelEqu	Change in equity to assets	XFIN	Net external financing
DelFINL	Change in financial liabilities	dNoa	Change in net operating assets
DelLTI	Change in long-term investment	grcapx	Change in capex (2 years)
DelNetFin	Change in net financial assets	grcapx3y	Change in capex (3 years)
Panel D: Profitability (15)			
CBOperProf	Cash-based operating profitability	OScore	O score
ChAssetTurnover	Change in asset turnover	OperProf	Operating profits / book equity
ChNWC	Change in net working capital	OperProfRD	Operating profitability R&D adj
CredRatDG	Credit rating downgrade	PS	Piotroski F-score
DelDRC	Deferred revenue	RoE	Net income / book equity
GP	Gross profits / total assets	Tax	Taxable income to income
Leverage	Market leverage	roaq	Return on assets (quarterly)
MS	Mohanram G-score		

Table A1 (continued)

(1) Abbreviation	(2) Name	(3) Abbreviation	(4) Name
Panel E: Intangibles (59)			
AOP	Analyst optimism	HerfAsset	Industry concentration (assets)
Activism1	Takeover vulnerability	HerfBE	Industry concentration (equity)
Activism2	Active shareholders	IO_ShortInterest	Inst own among high short interest
AdExp	Advertising expense	IndIPO	Initial public offerings
AgeIPO	IPO and age	OPLeverage	Operating leverage
AnalystRevision	EPS forecast revision	OrderBacklog	Order backlog
AnalystValue	Analyst value	OrderBacklogChg	Change in order backlog
CashProd	Cash Productivity	OrgCap	Organizational capital
ChEQ	Growth in book equity	PatentsRD	Patents to RD expenses
ChForecastAccrual	Change in forecast and accrual	PredictedFE	Predicted analyst forecast error
ChNAnalyst	Decline in analyst coverage	RD	R&D over market cap
ChTax	Change in taxes	RDAbility	R&D ability
ChangeInRecommendation	Change in recommendation	RDIPO	IPO and no R&D spending
CitationsRD	Citations to RD expenses	RDS	Real dirty surplus
ConsRecomm	Consensus Recommendation	RDcap	R&D capital to assets
ConvDebt	Convertible debt indicator	RIO_Disp	Inst Own and Forecast Dispersion
DelBreadth	Breadth of ownership	RIO_MB	Inst Own and Market to Book
DownRecomm	Down forecast EPS	RIO_Turnover	Inst Own and Turnover
EarningsConsistency	Earnings consistency	RIO_Volatility	Inst Own and Idio Vol
EarningsForecastDisparity	Long-vs-short EPS forecasts	Recomm_ShortInterest	Analyst Recomm and Short Interest
ExclExp	Excluded expenses	Spinoff	Spinoffs
FR	Pension funding status	SurpriseRD	Unexpected R&D increase
FirmAge	Firm age based on CRSP	UpRecomm	Up forecast
ForecastDispersion	EPS forecast dispersion	VarCF	Cash flow to price variance
Frontier	Efficient frontier index	fgr5yrLag	Long-term EPS forecast
Governance	Governance index	hire	Employment growth
GrAdExp	Growth in advertising expenses	realestate	Real estate holdings
GrSaleToGrInv	Sales growth over inventory growth	sinAlgo	Sin stock (selection criteria)
GrSaleToGrOverhead	Sales growth over overhead growth	tang	Tangibility
Herf	Industry concentration (sales)		

Table A1 (continued)

(1) Abbreviation	(2) Name	(3) Abbreviation	(4) Name
Panel F: Trading frictions (35)			
Beta	CAPM beta	PriceDelaySlope	Price delay coefficient
BetaFP	Frazzini-Pedersen beta	PriceDelayTstat	Price delay SE adjusted
BetaLiquidityPS	Pastor-Stambaugh liquidity beta	ProbInformedTrading	Probability of informed trading
BetaTailRisk	Tail risk beta	ReturnSkew	Return skewness
BidAskSpread	Bid-ask spread	ReturnSkew3F	Idiosyncratic skewness (3F model)
CoskewACX	Coskewness using daily returns	ShareVol	Share volume
Coskewness	Coskewness	ShortInterest	Short interest
DolVol	Past trading volume	SmileSlope	Put vol minus call vol
ExchSwitch	Exchange switch	VolMkt	Volume to market equity
IdioRisk	Idiosyncratic risk	VolSD	Volume variance
IdioVol3F	Idiosyncratic risk (3 factors)	VolumeTrend	Volume trend
IdioVolAHT	Idiosyncratic risk (AHT)	betaVIX	Systematic volatility
Illiquidity	Amihud's illiquidity	skew1	Volatility smirk near the money
logPrice	Log price	std_turn	Share turnover volatility
MaxRet	Maximum return over month	zerotrade	Days with zero trades
OptionVolume1	Option to stock volume	zerotradeAlt1	Days with zero trades (alt 1)
OptionVolume2	Option volume to average	zerotradeAlt12	Days with zero trades (alt 2)
PriceDelayRsq	Price delay R^2		